


## Ongoing effects of pandemic-imposed learning environment disruption on student attitudes

Teemu Hynninen<sup>1</sup>,\* Henna Pesonen<sup>1</sup>, Olli Lintu<sup>1</sup>, and Petriina Paturi<sup>1</sup>

*Department of Physics and Astronomy, University of Turku, FI-20014, Turku, Finland*

 (Received 29 September 2022; accepted 5 December 2022; published 5 January 2023)

We present a study on the development of Finnish first-year physics majors' attitudes towards physics, as measured by the Colorado Learning Attitudes about Science Survey, before, during, and after the period of mandatory remote learning due to the coronavirus pandemic. We find that in the years before the pandemic, these attitudes did not change, but the period of extended remote learning due to the pandemic had a negative effect on the students' expertlike views. Similarly, the students who experienced the remote learning period in high school displayed a lower level of expertlike thinking as they entered university. As contact teaching resumed, moderate positive gains were seen, bridging some but not all of the gap in student attitudes left by the pandemic.

DOI: [10.1103/PhysRevPhysEducRes.19.010101](https://doi.org/10.1103/PhysRevPhysEducRes.19.010101)

### I. INTRODUCTION

Many societies shut down in 2020 due to the coronavirus pandemic. In many countries, this included the complete suspension of contact teaching. Online courses and distance education were known to be socially demanding with high dropout rates already before the pandemic [1–3] and the sudden shift to exclusively remote learning is believed to have had and continues to have a profound negative impact on social interactions, engagement, and well-being of students, due to the effect it had on both the quality of teaching and the social environment in which learning took place [4–6]. It is also yet unclear how far reaching in the future the impacts of this disruption will be.

While COVID has had a negative impact on multiple aspects of the lives of students, results on how emergency distance education has affected their academic achievement are somewhat controversial: according to some studies, some groups of university students have performed better in distance education caused by the pandemic than they did before the pandemic [7–10]. It seems that this positive shift is related to a change in students' learning habits. In remote teaching, studying is done more continuously rather than just before the exam. On the other hand, many studies imply that emergency distance education has had a negative impact on student performance [11–13]. It seems that while some students have had the necessary skills or access to appropriate support in order to manage in a remote learning

environment, other groups, like students at the beginning of their studies, students with low self-organization and communication skills, students with anxiety or depressive symptomatology, or students that are lacking in academic social interactions, are the most vulnerable to the negative effects of remote learning [13–16].

The impacts of the coronavirus pandemic have been studied from many viewpoints including students' habits [17], wellbeing [4,5,13], integration [18], satisfaction to teaching [19], and academic performance [8,14,20]. However, research on epistemological impacts is almost absent. Only few studies have looked into the changes in students' attitudes towards physics during the pandemic [21,22]. Much of the COVID-related research is also focusing on younger students, while higher education level is getting less attention.

While student well-being and the development of their understanding are important, we are more interested in how extended remote learning has affected the attitudes and views of the students towards physics as a science and themselves as physicists. In our view, the central goal of undergraduate studies in physics is to guide the students' personal growth towards the mindset of a physicist, and this may arguably be better represented by their views instead of academic merits. We study these views using the Colorado Learning Attitudes about Science Survey (CLASS) [23], a widely used and thoroughly verified tool with which instructors can evaluate how closely the views of their students resemble those of experts, and how teaching affects the development of these views. The views measured by CLASS may even predict the likelihood that a student goes on to become a physicist [24].

We administer the CLASS survey yearly to new physics students at the University of Turku, Finland, and study the development of student views during the first year of studies. Changing attitudes is often a slow process affected

\*teemu.hynninen@utu.fi

*Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.*

by many factors besides teaching, and development towards the mindset of a professional physicist will often continue even after graduation. Therefore it is realistic to expect only modest development of views during the students' first year of studies. Furthermore, not only is it difficult to develop expertlike thinking, physics teaching often has a negative effect on the students' views on physics [25–28], and the pandemic may have increased the risk for that to happen.

In this work, we set out to generally investigate how the views and attitudes of new physics students have evolved during their first year of studies at the university level before, during, and after the pandemic. We are specifically interested in comparing the development of expertlike thinking in three student populations:

- (i) the students who began their studies in 2019 or earlier and were able to complete their first year of studies in normal conditions,
- (ii) the students who began their studies in 2020 and had to study remotely during their first year of studies, and
- (iii) the students who began their studies in 2021 and had to study remotely in high school before starting their studies at the university.

We wish to find out (1) if expertlike thinking has developed differently in these groups and (2) if these differences correlate with academic performance. By comparing the first two groups, we get to compare the effects of contact and remote teaching directly. By comparing the third group to the first two, it is possible to investigate the long-term effects a previous teaching disruption has on future learning even after contact teaching has become available again. This kind of information may prove important in the immediate future, as new generations of students who have lived through the pandemic begin their studies.

## II. COURSE STRUCTURE

Some 40–60 physics students begin their studies each year, following a fairly fixed curriculum during their first year of studies. They study the basics of mechanics, thermodynamics, electromagnetism, and wave dynamics during a two-semester calculus-based course series called *Introduction to Physics*. These courses were originally redesigned and built in 2015–2017 based on ideas such as team-based learning [29,30], flipped learning [31], and gamification [32]. The goal was to foster a learning environment that supports active learning and student engagement, promoting both learning and inclusion in the campus community. The courses were not designed to specifically develop expertlike thinking.

The *Introduction* courses are worth about 1/3 of the credits the students earn during their first study year. The rest of the curriculum consists of introduction to studies, tutorials on basic computer use, experimental labs, foreign languages, and mathematics courses.

### A. Learning before the pandemic

Before the pandemic, there were no restrictions on contact teaching. We will call this the *full contact teaching* period even though the courses had online components also during this time.

Students were distributed in small teams at the start of each *Introduction* course. These teams were assigned by the teacher and they were constructed to be as heterogeneous as possible with respect to factors such as incoming grades and gender. All contact teaching was set to make students actively work with their team. Evaluation was designed so that a student could earn the top grade without taking an exam. Instead, it was enough to participate actively and show enough effort and mastery in course assignments. This was also the preferred way for the students, as only very few students ended up taking the exam.

As the courses started, students first engaged with new information in reading assignments, where the material was distributed online on the social learning platform *Perusall* [33]. This was followed by contact teaching where student teams worked on questions and tutorials, the teacher acting as an advisor.

Once students became more familiar with the most important physical concepts, the teams were given a few fairly broad and sometimes open-ended problems. The solutions to these problems were turned in on paper at a later date, and given to other teams for peer assessment. While the problems were solved in teams, each problem was also accompanied by questions about the solution process, and every student had to answer these individually.

Traditional physics problems focusing on relatively simple calculations with little or no real physical context were not a major part of these courses. There were examples of such problems in the reading material, and some were also solved as part of the activities during contact teaching, but much more focus was given on conceptual understanding. Still, we provided the students with an online collection of automatically assessed physics problems to complement the concept-focused material. Solving these problems was not mandatory but could earn credit towards higher grades.

All in all, the new course structure was markedly different from traditional introductory physics teaching, since these new courses had practically no lectures, recitations or exams, which are typical activities on traditional courses.

### B. Learning during lockdown

Finland entered lockdown in March 2020. This was about a month before the end of the spring semester, but after the postsurvey data had been collected in Spring 2020. However, as the lockdown continued and contact teaching became impossible during the semesters of Fall 2020 and Spring 2021, all activity had to be taken online. We will call this the time of *no contact teaching*.

At this point we knew that the challenges of keeping students engaged and fostering a feeling of inclusion would

become even more crucial than before. However, since the courses were originally designed to promote exactly these factors, we decided that the basic structure of the courses should remain unchanged. This meant that students were still assigned to teams, and they still began their studies by reading the given material online. The questions, tutorials and problems were mostly the same or similar to those given to students during normal contact teaching, only now they were distributed online. The same online collection of extra problems was made available.

The courses also followed the same schedule as before lockdown. Whenever there would have been contact teaching in normal conditions, the students would now log on to the online platform *Discord* [34], where they would discuss the problems in teams with access to the teacher for advice. If a tutorial would have been completed during the contact teaching session, the students would now submit their work online at the end of the remote study session. The few teacher-driven sessions were replaced by online videos with embedded questions for the teams to answer. Peer assessment was also organized online.

The most important difference compared to previous years was therefore not in what the students were supposed to do but how and where they would do it. Most importantly, all team work and social activities would happen remotely instead of in person.

### C. Learning after lockdown

During the semesters of Fall 2021 and Spring 2022, the pandemic was ongoing but most people had access to vaccines. This allowed the university to start a gradual return to normalcy. Contact teaching was allowed in small study groups, but larger gatherings of several tens of people were still banned. This ban was lifted midsemester in the fall, but reinstated in January 2022 for several more months. In practice, this meant that for half of the academic year, roughly 1/3 of teaching could be given in person, as before the pandemic, while 2/3 was still online. We call this *limited contact teaching*.

## III. METHODOLOGY

We have administered the Finnish translation of the CLASS survey [35] to our students as part of the *Introduction* courses for four consecutive years, starting in Fall 2018. The students take the survey as a presurvey on the first lesson of the course, at the very beginning of their studies. They take the survey again the next spring near the end of the spring semester as a postsurvey. Normally, the survey is given on paper to the students present at the time during a lesson. During remote learning, the survey was made available online and course credit was given for completing the survey in order to ensure a high response rate, as recommended [36]. In our analysis, we only consider the students who completed both surveys.

The survey consists of 42 statements answered on a 5-point Likert scale (strongly disagree to strongly agree). The results are analyzed by calculating expertlike (favorable) and nonexpertlike (unfavorable) responses in the pre- and postsurveys for each student. These results are compiled in 8 categories (*real world connection*, *personal interest*, *sense making or effort*, *conceptual connections*, *applied conceptual understanding*, *problem-solving general*, *problem-solving confidence*, and *problem-solving sophistication*) and an overall score for expertlike thinking. Scores are calculated for the whole class of students by averaging the scores of individual students [23].

The difference, or effect size, between the score averages of two surveys,  $x_0$  and  $x_1$ , are evaluated using Cohen's  $d$ ,

$$d = \frac{x_1 - x_0}{\sqrt{\frac{1}{2}(s_1^2 + s_0^2)}}, \quad (1)$$

where  $s_0^2$  and  $s_1^2$  are the corresponding sample variances. A large  $d$  means a more pronounced difference between the means of the two distributions. Although there is no absolute scale for  $d$ , a common rule of thumb says that  $d = 0.5$  corresponds to moderate difference.

In addition to averages, we also compare distributions of expertlike thinking scores for entire classes. To test if two distributions are statistically different, we perform the Mann-Whitney U test [37]. For this test, we report the effect size

$$r = 1 - \frac{2U}{n_1 n_2}, \quad (2)$$

where  $U$  is the Mann-Whitney statistic and  $n_1$  and  $n_2$  are the sample sizes from both distributions. Since the maximum value of  $U$  is  $n_1 n_2$ , this effect size will always be between  $-1$  and  $1$ .

Although the *Introduction* courses are meant for first year physics students, some students retake the courses later during their studies. Also a few students majoring in other natural sciences or mathematics decide to enroll on these courses instead of the somewhat easier physics courses meant for nonphysicists. In this study, we wish to concentrate on physics majors who are beginning their studies, and therefore we remove responses from other students from the data. This leaves us with a total of 140 valid responses. The breakdown of responses per academic year is shown in Table I. This table also lists the limitations

TABLE I. Number of responses each academic year.

Semesters	Contact teaching	$N$
Fall 2018, Spring 2019	Full	38
Fall 2019, Spring 2020	Full	32
Fall 2020, Spring 2021	None	30
Fall 2021, Spring 2022	Limited	40

on contact teaching for each year. Because we have data from two years of full contact teaching, academic years 2018–2019 and 2019–2020, we combine the survey results from these years in our analysis (*full contact teaching*,  $N = 70$ ).

## IV. RESULTS

### A. Effects of teaching on expertlike thinking

Survey results are collected in Fig. 1 for all three levels of contact teaching. Results from full contact teaching is shown on the top, no contact teaching in the middle, and limited contact teaching at the bottom. Each pair of columns represents a category of the CLASS survey, the rightmost one being the overall score. In each pair of columns, presurvey results are on the left and postsurvey on the right. Bottom bars show the percentage of answers displaying expertlike thinking, and top bars show non-expertlike thinking. The gap in between is the portion of neutral answers. Error bars show the calculated standard error of mean.

Overall, the students display a fairly high level of expertlike thinking. This is no surprise, as we are surveying physics majors, who very likely have a higher than average level of interest in physics. This is seen especially in the categories *real world connection* and *personal interest*, where scores of over 80% favorable answers are reached. The lowest scores are seen in the categories applied conceptual understanding and problem-solving sophistication, which measure the

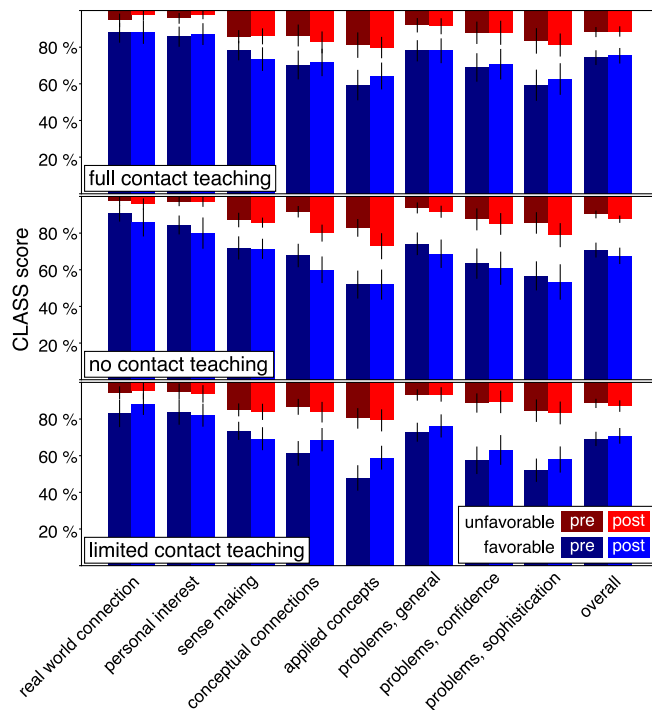


FIG. 1. Portion of expertlike (favorable) and nonexpertlike (unfavorable) answers in both pre- and postsurvey for different levels of contact teaching.

students' attitudes when applying their knowledge of physics. These categories do not measure just interest but also strategies, which are often still unsophisticated when students begin their studies at the university level.

The effect of the first year of studies is seen by comparing the pre- and postsurvey results. During the prepandemic years where full contact teaching was possible, shifts in expertlike thinking are fairly moderate. Some gains are made in the categories where the presurvey scores were the lowest, but there is practically no gain in the overall score. Still, this is not a bad result, since the overall score is fairly high in both pre- and postsurvey ( $74\% \pm 4\%$  and  $75\% \pm 4\%$  favorable, respectively), and it is not uncommon for CLASS scores to drop during the first year of studies [25–28].

The results are very different when no contact teaching was possible, as the portion of favorable answers are seen to decrease while unfavorable answers become more common. Although in some categories the shifts are not very pronounced, what is most striking is that this happens consistently. Not a single category shows any positive gains. Fortunately, when limited contact teaching was again possible, we again see increases in favorable answers, and in some cases these gains seem to be even stronger than when full contact teaching was possible.

The changes from pre- to postsurvey can be better understood by examining the effect sizes as shown in Fig. 2. In this plot, a positive value means a higher value in the postsurvey. Therefore we wish to see positive values for

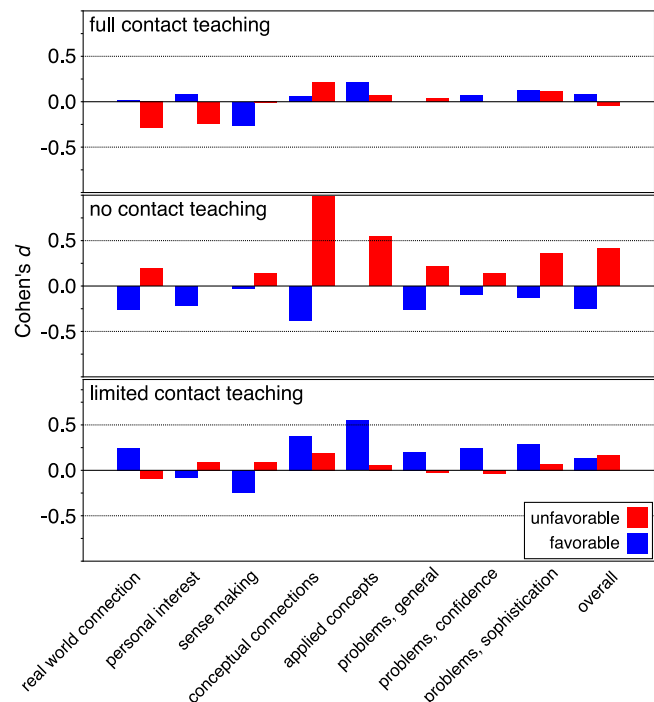


FIG. 2. Effect size from presurvey to postsurvey for different levels of contact teaching.

favorable answers (more expertlike answers after teaching) and negative values for unfavorable answers (less non-expertlike answers after teaching).

As we already discussed, full contact teaching seems to have had little effect on student thinking. We see positive and negative values for both favorable and unfavorable answers, but the effect sizes are typically close to zero and at most  $|d| = 0.26$ , which is still considered to be a small effect.

The case is different for the students who received no contact teaching. In all categories, the effect on favorable answers is either zero or negative while the effect on unfavorable answers is zero or positive. This negative effect is small or moderate on the favorable answers, but in some categories there are strong effects on the unfavorable answers. The strongest increase in unfavorable answers is seen for conceptual connections where the effect size is 1.0. Upon further analysis we find that although there is an increase in unfavorable answers to all the statements in this category, this effect is most strongly driven by a great increase of unfavorable answers in statement 1, “A significant problem in learning physics is being able to memorize all the information I need to know.” In the presurvey, only 10% of students agreed with this statement while 40% did so in the postsurvey, showing a significant increase in nonexpertlike thinking. For comparison, 19% and 17% of students in full contact teaching agreed with this statement in pre- and postsurveys, respectively.

Besides statement 1, unfavorable answers increase markedly also in statement 5, “After I study a topic in physics and feel that I understand it, I have difficulty solving problems on the same topic” (from 35% to 57%), and 40, “If I get stuck on a physics problem, there is no chance I’ll figure it out on my own” (from 20% to 30%). These contribute to the applied conceptual understanding as well the problem solving categories, and as a result these categories also show a moderate effect in increased non-expertlike thinking.

Much more positive results are seen with limited contact teaching. Although overall gains in favorable answers are small and there is roughly equal increase in unfavorable answers, most categories show small to moderate effects in increased favorable answers. In fact, the highest gains in expertlike thinking are seen in conceptual connections and applied conceptual understanding, where we saw the highest increase in nonexpertlike thinking for the students with no contact teaching. All problem-solving categories also show a small positive effect in expertlike thinking. These results can again be attributed to statements 1, 5, and 40, only now we see a significant increase in favorable answers (from 43%, 8%, and 25% to 58%, 18%, and 48%, respectively).

## B. Distributions of expertlike thinking scores

Next, we turn our attention from mere averages to the distribution of expertlike thinking in the student population. Figure 3 show these distributions for overall favorable

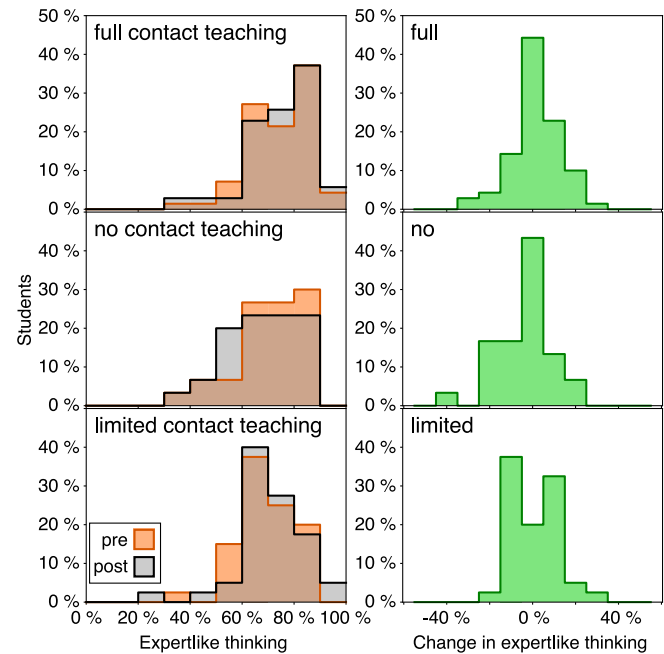


FIG. 3. Distribution of overall favorable CLASS scores in pre- and postsurveys, and the change in this score.

answers in both pre- and postsurveys, as well as the distribution of change in favorable answers.

In classes that received full contact teaching, most students demonstrate expertlike thinking in 60%–90% of their answers in both pre- and postsurveys, as only a few students fall outside of this range. The distribution peaks in the 80%–90% range showing that many students have developed a high level of expertlike thinking already in high school. The distribution for the change in expertlike thinking peaks symmetrically at 0% showing that, on average, student views did not change.

The class that had to study with no contact teaching shows a similar distribution in the presurvey with most students in the 60%–90% range. However, this is not true anymore for the post-survey distribution, which has broadened so that the bulk of students now fall in the 50%–90% range. That is, the number of students scoring between 60% and 90% has decreased while those scoring between 50% and 60% has increased. This drift towards lower scores in expertlike thinking is seen also in the distribution for change, which again peaks at 0% but is now asymmetric, displaying more negative than positive changes.

The situation is reversed in the class that studied under limited contact teaching. Now, the bulk of students are distributed between 50% and 90% already in the presurvey. Furthermore, the distribution peaks between 60% and 70%. Clearly, this student population demonstrated a lower level of expertlike thinking compared to previous years at the start of their studies.

In the postsurvey most students are again found in the 60%–90% range with only few students in the 50%–60%

interval, showing a positive shift due to teaching. However, the distribution still peaks between 60% and 70% instead of 80% and 90% as in the full contact teaching data. Also, the distribution for changes in expertlike thinking now shows two peaks at around  $\pm 10\%$ . Curiously, further analysis reveals these peaks consist entirely of female students. For male students this distribution peaks at 0%. However, we could not demonstrate any particular explaining factor for this.

We check if these changes in distributions are statistically significant using the Mann-Whitney U test. We find that the difference between pre- and postsurvey distributions is not significant in any of the groups ( $|r| \approx 0.1$  and  $p \approx 0.2$  in all cases). This is expected in the full contact teaching group since we see little difference between pre- and postsurvey distributions in Fig. 3. For the other groups, Fig. 3 suggests there may be a difference between pre- and postsurveys, but our data is insufficient to reach a statistically significant conclusion.

However, we do find statistically significant differences between distributions of *different* groups. The presurvey results of full and no contact teaching groups are similar ( $|r| = 0.16$  and  $p < 0.1$ ), but the postsurvey distributions are statistically different ( $|r| = 0.34$  and  $p < 0.005$ ). That is, these groups began their studies showing a similar distribution of expertlike thinking but displayed genuinely different levels of expertlike thinking after the first year of studies.

Similar analysis reveals that the full and limited contact teaching groups are statistically different in both presurvey ( $|r| = 0.26$  and  $p < 0.02$ ) and postsurvey ( $|r| = 0.25$  and  $p < 0.02$ ), in agreement with our previous discussion. The distribution of expertlike thinking of students was significantly different (less mature) in the limited contact teaching group compared to previous years at the beginning of their first year of studies. Furthermore, compared to the full contact teaching group, a significant difference remained after one year, although this difference was smaller than the difference between the groups that received full or no contact teaching.

To further quantify the differences between the full and limited contact teaching classes, we calculate Cohen's  $d$  for their presurvey and postsurvey results in all CLASS categories, as shown in Fig. 4. Note the difference to Fig. 2: We are not comparing the same set of students before and after teaching, so this is not a measure of the effect of teaching. We are measuring statistical difference between distributions involving different students. The plot on the top shows the difference in presurvey between full and limited contact teaching groups, and the bottom plot similarly for postsurvey. Positive values mean a higher score was recorded for the limited contact teaching group.

Comparing favorable answers in presurveys, we find that the group that received limited contact teaching scored lower in all categories compared to the group in full contact

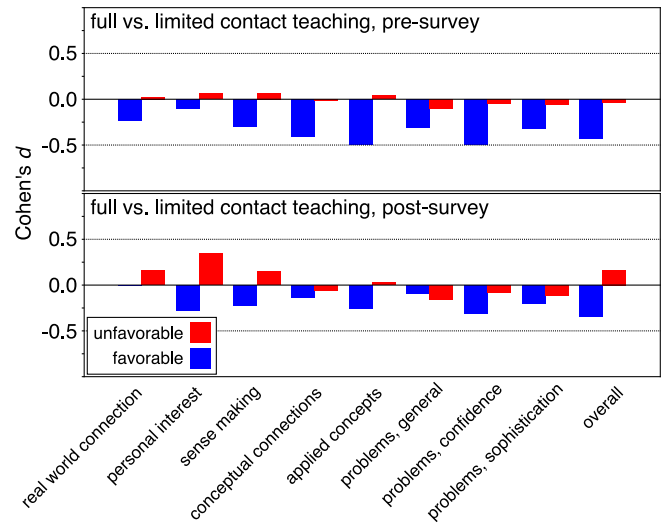


FIG. 4. Comparison of survey results for classes with full and limited contact teaching. Differences are calculated in such a way that a positive value means the class with limited contact teaching scored higher.

teaching. In several categories, including the overall CLASS score,  $|d| \approx 0.5$ , meaning that the difference is moderately significant. Furthermore, even though we saw generally positive gains in expertlike thinking during limited contact teaching, these gains were not enough to close the gap, as the limited contact teaching group still scored lower in the postsurvey compared to the full contact teaching group.

### C. Course completion

Our survey data has shown that there are moderate differences in the level of expertlike thinking in students who studied under different conditions. Finally, we compare these results to student performance.

Since students pass the *Introduction* courses by participating in course activities, grades do not reflect mastery of the subject, as displayed effort is a fairly significant factor in several assignments. Still, the final grade is a measure of some combination of motivation, effort and mastery, and therefore a meaningful reference nonetheless.

To test if there is correlation between CLASS results and course performance, we calculate Pearson's correlation coefficient  $R$  between the CLASS category scores and the final grade (given on an integer scale of 0–5). As a result, we do find positive correlation, but usually this correlation is fairly weak. For classes in full or limited contact teaching, all  $R$  values are about 0.35 or less. So even though high-performing students tend to have higher CLASS scores on average, CLASS is not a good predictor of individual student performance in these groups.

However, for students that received no contact teaching, we see stronger correlation between CLASS and grades. The highest  $R$  is found for postsurvey overall CLASS

TABLE II. Number of physics majors who passed the *Introduction* courses in the fall and in the spring and the dropout rate.

Semesters	Passed fall	Passed spring	Dropout rate
Fall 2018, Spring 2019	56	52	7%
Fall 2019, Spring 2020	44	40	9%
Fall 2020, Spring 2021	45	37	17%
Fall 2021, Spring 2022	52	49	6%

score, for which  $R = 0.65$ . There is moderately strong correlation also between all the problem-solving categories and the final grade, as they have  $R$  values between 0.5 and 0.6 in both pre- and postsurvey ( $p < 0.001$  for all).

Finally, in Table II, we list the total number of physics majors who passed the *Introduction* courses each fall and spring, and the corresponding dropout rate calculated as the difference between these two numbers, divided by the number of students who pass the fall course. In the years 2018–2019 and 2019–2020, with full contact teaching, the dropout rate during the first study year was slightly less than 10%. In 2020–2021 and no contact teaching, the dropout rate doubled to 17%. This is in accordance with negative changes in expertlike thinking during this year. Lastly in 2021–2022 and limited contact teaching, a low dropout rate of 6% was recovered.

## V. DISCUSSION

We set out to compare the development of expertlike views in three different groups of new physics majors, and our results found differences between all of them.

Before the pandemic, students displayed a fairly high level of expertlike thinking already at the beginning of their studies, and their views did not significantly change during their first year of studies. The students who began their studies in 2020 also demonstrated similarly high levels of expertlike views at the beginning of their studies, but their results deteriorated during a year of remote learning, showing a systematic decrease in expertlike answers and an increase in nonexpertlike answers.

This result is not surprising but it is in direct contrast to previous results that showed no change in the students' views on experimental physics after emergency remote teaching [21]. Similarly in a study conducted before the pandemic, there were no differences in attitudes between students in in-person and online physics laboratory teaching [38]. Note though that these studies were based on a survey focusing explicitly on experimental physics, whereas CLASS focuses on theoretical and more general aspects of physics. These surveys probe clearly different aspects of physical thinking and therefore it is not surprising that different results are observed. It is also not straightforward to compare results from before and during the pandemic, as the lockdowns did not only disrupt the way people studied but many other aspects of their lives as

well. Places such as libraries, theaters, restaurants, and sports facilities were all shut down in Finland during the pandemic, and, for instance, student clubs were not allowed to organize in-person activities.

Although student views developed negatively in all categories during the pandemic, the most severe effect was on a statement about physics as a collection of facts to remember. There is much to learn during undergraduate studies, but physics forms a logical whole which must be studied through understanding instead of memorization. A student who sees physics as a collection of facts is not in a position to construct a coherent understanding of physical theories and will likely face severe difficulties in his or her studies. Therefore, it is a worrying observation that the pandemic affected this view so negatively. This result is also in agreement with some earlier studies where emergency remote teaching had a negative impact on student performance [11–13].

The flipped learning model applied in the *Introduction* courses makes the student responsible for getting acquainted with new information, but the logical structure of this information is explored together with the teacher during lessons. In contact teaching, this was successful, but even though students had access to teacher assistance during the pandemic, many students failed to understand these connections when this aid was provided remotely. This is possibly due to it being easier for the students to miss or even purposefully avoid communicating with the teacher and their peers in an online environment, while it is more difficult for the teacher to reach the students in order to guide them. In such an environment, self-organization and communication skills of the students become increasingly important [14]. Also, learning to understand the logical structure of physics can require resilience which many students may have lacked in the stressful environment created by the pandemic [39,40].

We saw that high CLASS scores correlated modestly with academic performance in the full and limited contact teaching groups. This is in agreement with previous studies where similar correlations have been observed [25,41]. In the group that received no contact teaching, however, we found the correlation to be much stronger than in the other groups and also much stronger than what is typically reported. Especially the postsurvey overall CLASS score was a good indicator of success. In addition, the presurvey scores in the problem-solving categories were fairly good predictors of academic performance. This suggests that a contact teaching environment can help overcome some of the nonexpertlike views students have regarding problem solving. Perhaps this is because students seek help more frequently face to face than online [42,43]. In a remote learning environment, the students need to develop fairly sophisticated expertlike thinking in order to achieve high academic performance. The necessity for well-developed study skills is likely reflected also in the

dropout rate [44–46], which was clearly higher in remote learning compared to other groups. It seems that contact learning could provide enough aid and social support so that also most of the low-performing students were able to complete the courses, whereas remote learning could not.

As a result of these challenges, a higher percentage of students did not finish their first year of studies during the pandemic compared to other years. We do not expect this to affect our conclusions, though. Since correlation between academic performance and CLASS scores was particularly strong in this group, it is likely that the students who did not finish their studies would have scored low in the CLASS postsurvey, had they taken it. Therefore, if these missing students were included in the dataset, we would expect even lower scores in the no contact teaching postsurvey, i.e., an even more pronounced difference to the other groups.

The students in the limited contact teaching group began their studies a year later after having experienced the emergency remote learning during their final year in high school. Subsequently, they scored consistently lower in expertlike thinking in the presurvey compared to previous classes. We cannot claim for certain nor confirm that this lower score is due to the learning disruption they experienced earlier, as we have no CLASS data for this group from before the pandemic. This difference can be due to other factors or even just chance. Also, a few students in this group had graduated from high school earlier than the previous spring and entered university after a gap year. They had not experienced the pandemic during high school. Therefore it is an oversimplification to claim that the drop in expertlike thinking is solely due to remote teaching. It is plausible, however, that earlier remote teaching was an important contributing factor.

In any case, as we saw in the group that received no contact teaching, a disruption in the learning environment can cause deterioration in expertlike thinking. Therefore we may consider the group in limited contact teaching as a *model* for a population whose views have been affected by the pandemic, even if we are not certain if that the pandemic was the ultimate reason for the observed difference.

For this group of students, participation in the flipped-learning-based *Introduction* courses resulted in an overall increase in both expertlike and nonexpertlike thinking. Importantly though, this group displayed positive development in expertlike thinking in exactly the same statements that were most adversely affected by the pandemic in the

previous class. The dropout risk was also reduced back to prepandemic level despite possible skill deficits compared to earlier classes. This suggests that negative effects of earlier learning disruptions could be gradually mended once contact learning can be resumed. Still, overall non-expertlike thinking increased as much as expertlike thinking in this group, and their postsurvey CLASS score was still moderately lower than that of the full contact learning group. Apparently, the negative effects from one year of disrupted studies can be long lasting and not easily mended in a similar timeframe.

On a more positive note, these results were obtained in limited contact teaching. This indicates that even though many students do not cope well in a learning environment based only on remote learning, providing some regular in-person lessons in addition to online activities may be enough to mitigate the negative effects of exclusive remote learning.

Naturally, these results are likely dependent on the details of the learning environment and provided communication possibilities, and one cannot directly generalize the conclusions to other courses organized in a different fashion. Many activities on the *Introduction* courses require teamwork, which forces the students to communicate at least to some degree. This may have provided the students with valuable social connections during the pandemic, but it may also have increased cognitive load and lead to the students focusing on how to use online collaborative tools instead of learning physics. Courses focusing on individual effort surely had to deal with very different kinds of difficulties during the lockdown.

Also note that the *Introduction* courses account for only one-third of the academic workload these students experience during their first year of studies. Although these courses likely have the most direct impact on the students' attitudes towards physics, other courses and extracurricular activities (or lack thereof) also influence student thinking. We cannot separate these effects from our data, but on the other hand, all of these activities were disrupted by the pandemic simultaneously and our results may be interpreted as a sum effect of the disruption.

## ACKNOWLEDGMENTS

This work is supported by the Jenny and Antti Wihuri Foundation and the Magnus Ehrnrooth Foundation (H. P.).



- [1] V. Rolfe, A systematic review of the socio-ethical aspects of massive online open courses, *Eur. J. Open, Distance E-Learning* **18**, 52 (2015).
- [2] O. Simpson, Student retention in distance education: Are we failing our students?, *Open Learning* **28**, 105 (2013).
- [3] S. J. Greenland and C. Moore, Large qualitative sample and thematic analysis to redefine student dropout and retention strategy in open online education, *Br. J. Educ. Technol.* **53**, 647 (2022).
- [4] T. Elmer, K. Mepham, and C. Stadtfeld, Students under lockdown: Comparisons of students' social networks and mental health before and during the COVID-19 crisis in Switzerland, *PLoS One* **15**, e0236337 (2020).
- [5] S. Rudenstine, K. Mcneal, T. Schulder, C. K. Ettman, M. Hernandez, K. Gvozdieva, and S. Galea, Depression and anxiety during the COVID-19 pandemic in an urban, low-income public university sample, *Journal of traumatic stress* **34**, 12 (2021).
- [6] I. Marzoli, A. Colantonio, C. Fazio, M. Giliberti, U. Scotti di Uccio, and I. Testa, Effects of emergency remote instruction during the COVID-19 pandemic on university physics students in Italy, *Phys. Rev. Phys. Educ. Res.* **17**, 020130 (2021).
- [7] E. Burkholder and C. Wieman, Absence of a COVID-induced academic drop in high-school physics learning, *Phys. Rev. Phys. Educ. Res.* **18**, 023102 (2022).
- [8] S. Iglesias-Pradas, Á. Hernández-García, J. Chaparro-Peláez, and J. L. Prieto, Emergency remote teaching and students' academic performance in higher education during the COVID-19 pandemic: A case study, *Comput. Human Behav.* **119**, 106713 (2021).
- [9] T. Gonzalez, M. A. de la Rubia, K. P. Hincz, M. Comas-Lopez, L. Subirats, S. Fort, and G. M. Sacha, Influence of COVID-19 confinement on students' performance in higher education, *PLoS One* **15**, e0239490 (2020).
- [10] G. M. Alam and M. Parvin, Can online higher education be an active agent for change?—Comparison of academic success and job-readiness before and during COVID-19, *Technological Forecasting and Social Change* **172**, 121008 (2021).
- [11] L. Elsalem, N. Al-Azzam, A. A. Jum'ah, and N. Obeidat, Remote E-exams during Covid-19 pandemic: A cross-sectional study of students' preferences and academic dishonesty in faculties of medical sciences, *Ann. Med. Surgery* **62**, 326 (2021).
- [12] C. Tan, The impact of COVID-19 on student motivation, community of inquiry and learning performance, *Asian Educ. Develop. Studies* **10**, 308 (2021).
- [13] L. Giusti, S. Mammarella, A. Salza, S. D. Vecchio, D. Ussorio, M. Casacchia, and R. Roncone, Predictors of academic performance during the COVID-19 outbreak: Impact of distance education on mental health, social cognition and memory abilities in an Italian university student sample, *BMC Psychol.* **9**, 142 (2021).
- [14] P. Klein, L. Ivanjek, M. Dahlkemper, K. Jelicic, M.-A. Geyer, S. Küchemann, and A. Susac, Studying physics during the COVID-19 pandemic: Student assessments of learning achievement, perceived effectiveness of online recitations, and online laboratories, *Phys. Rev. Phys. Educ. Res.* **17**, 010117 (2021).
- [15] T. P. Ivanec, The lack of academic social interactions and students' learning difficulties during COVID-19 faculty lockdowns in Croatia: The mediating role of the perceived sense of life disruption caused by the pandemic and the adjustment to online studying, *Social Sci.* **11**, 42 (2022).
- [16] Y. M. Tang, P. C. Chen, K. M. Law, C. H. Wu, Y. yip Lau, J. Guan, D. He, and G. T. Ho, Comparative analysis of student's live online learning readiness during the coronavirus (COVID-19) pandemic in the higher education sector, *Comput. Educ.* **168**, 104211 (2021).
- [17] F. Gallè, E. A. Sabella, G. Da Molin, O. De Giglio, G. Caggiano, V. D. Onofrio, S. Ferracuti, M. T. Montagna, G. Liguori, G. B. Orsi, and C. Napoli, Understanding knowledge and behaviors related to COVID-19 epidemic in Italian undergraduate students: The EPICO study, *Int. J. Environ. Res. Public Health* **17**, 3481 (2020).
- [18] K. Resch, G. Alnahdi, and S. Schwab, Exploring the effects of the COVID-19 emergency remote education on students' social and academic integration in higher education in Austria, *Higher Educ. Res. Dev.* **1** (2022).
- [19] C. Y. Su and Y. Guo, Factors impacting university students' online learning experiences during the COVID-19 epidemic, *J. Comput. Assist. Learn.* **37**, 1578 (2021).
- [20] M. Kuhfeld, J. Soland, B. Tarasawa, A. Johnson, E. Ruzek, and J. Liu, Projecting the potential impact of COVID-19 school closures on academic achievement, *Educ. Res.* **49**, 549 (2020).
- [21] M. F. Fox, J. R. Hoehn, A. Werth, and H. Lewandowski, Lab instruction during the COVID-19 pandemic: Effects on student views about experimental physics in comparison with previous years, *Phys. Rev. Phys. Educ. Res.* **17**, 010148 (2021).
- [22] V. Borish, A. Werth, N. Sulaiman, M. F. Fox, J. R. Hoehn, and H. Lewandowski, Undergraduate student experiences in remote lab courses during the COVID-19 pandemic, *Phys. Rev. Phys. Educ. Res.* **18**, 020105 (2022).
- [23] W. K. Adams, K. K. Perkins, N. S. Podolefsky, M. Dubson, N. D. Finkelstein, and C. E. Wieman, New instrument for measuring student beliefs about physics and learning physics: The Colorado Learning Attitudes about Science Survey, *Phys. Rev. ST Phys. Educ. Res.* **2**, 010101 (2006).
- [24] K. K. Perkins and M. Gratny, Who becomes a physics major? A long-term longitudinal study examining the roles of pre-college beliefs about physics and learning physics, interest, and academic achievement, *AIP Conf. Proc.* **1289**, 253 (2010).
- [25] K. Perkins, W. Adams, S. Pollock, N. Finkelstein, C. Wieman, R. Beichner, R. Chabay, B. Sherwood, E. Gire, A. Hines, R. Lindell, and Z. Hrepic, Correlating student beliefs with student learning using the Colorado Learning Attitudes about Science Survey, *AIP Conf. Proc.* **790**, 61 (2005).
- [26] A. Madsen, S. B. McKagan, and E. C. Sayre, How physics instruction impacts students' beliefs about learning physics: A meta-analysis of 24 studies, *Phys. Rev. ST Phys. Educ. Res.* **11**, 010115 (2015).

- [27] P. Zhang, L. Ding, and E. Mazur, Peer Instruction in introductory physics: A method to bring about positive changes in students' attitudes and beliefs, *Phys. Rev. ST Phys. Educ. Res.* **13**, 010104 (2017).
- [28] W. Z. Shi, L. Ma, and J. Wang, Effects of inquiry-based teaching on Chinese university students' epistemologies about experimental physics and learning performance, *J. Baltic Sci. Educ.* **19**, 289 (2020).
- [29] L. K. Michaelsen, A. B. Knight, and L. D. Fink, *Team-based Learning: A Transformative Use of Small Groups* (Greenwood Publishing Group, Westport, CT, 2002).
- [30] D. W. Johnson and R. T. Johnson, An educational psychology success story: Social interdependence theory and cooperative learning, *Educ. Res.* **38**, 365 (2009).
- [31] C. H. Crouch and E. Mazur, Peer Instruction: Ten years of experience and results, *Am. J. Phys.* **69**, 970 (2001).
- [32] M. Sailer and L. Homner, The gamification of learning: A meta-analysis, *Educ. Psychol. Rev.* **32**, 77 (2020).
- [33] K. Miller, B. Lukoff, G. King, and E. Mazur, Use of a social annotation platform for pre-class reading assignments in a flipped introductory physics class, *Front. Educ.* **3** (2018).
- [34] Discord, Discord Inc. (2020).
- [35] M. Asikainen, Colorado Learning Attitudes about Science Survey, Finnish translation (2020), <https://www.physport.org/assessments/assessment.cfm?I=3&A=CLASS>.
- [36] A. Madsen and S. McKagan, Administering research-based assessments online (2020), <https://www.physport.org/recommendations/Entry.cfm?ID=93329>.
- [37] H. B. Mann and D. R. Whitney, On a test of whether one of two random variables is stochastically larger than the other, *Ann. Math. Stat.* **18**, 50 (1947).
- [38] D. J. Rosen and A. M. Kelly, Epistemology, socialization, help seeking, and gender-based views in in-person and online, hands-on undergraduate physics laboratories, *Phys. Rev. ST Phys. Educ. Res.* **16**, 020116 (2020).
- [39] R. Baltà-Salvador, N. Olmedo-Torre, M. Peña, and A.-I. Renta-Davids, Academic and emotional effects of online learning during the COVID-19 pandemic on engineering students, *Educ. Inf. Technol. (Dordr)* **26**, 7407 (2021).
- [40] M. L. Pertegal-Felices, D. A. Valdivieso-Salazar, A. Espín-León, and A. Jimeno-Morenilla, Resilience and academic dropout in Ecuadorian University students during COVID-19, *Sustainability (Switzerland)* **14**, 8066 (2022).
- [41] A. Robinson, J. H. Simonetti, K. Richardson, and M. Wawro, Positive attitudinal shifts and a narrowing gender gap: Do expertlike attitudes correlate to higher learning gains for women in the physics classroom?, *Phys. Rev. Phys. Educ. Res.* **17**, 010101 (2021).
- [42] R. A. Mahasneh, A. K. Sowan, and Y. H. Nassar, Academic help-seeking in online and face-to-face learning environments, *E-Learning Dig. Media* **9**, 196 (2012).
- [43] P. M. Reeves and R. A. Sperling, A comparison of technologically mediated and face-to-face help-seeking sources, *Br. J. Educ. Psychol.* **85**, 570 (2015).
- [44] S. B. Robbins, H. Le, D. Davis, K. Lauver, R. Langley, and A. Carlstrom, Do psychosocial and study skill factors predict college outcomes? A meta-analysis, *Psychol. Bull.* **130**, 261 (2004).
- [45] N. D. Bas, R. Heininger, M. C. Utesch, and H. Krcmar, Influence of study skills on the dropout rate of universities: Results from a literature study, *Adv. Intel. Syst. Comp.* **715**, 133 (2018).
- [46] J. Polansky, J. J. Horan, and C. Hanish, Experimental construct validity of the outcomes of study skills training and career counseling as treatments for the retention of at-risk students, *J. Counsel. Develop.* **71**, 488 (1993).