The use of HTML-based dynamic content for the learning of statistical test

Takuya Kitamoto[1,*,†], Takeo Noda[2,†] and Masataka Kaneko[3,†]

[1] *Yamaguchi University, 1677-1 Yoshida, Yamaguchi, 7538513, Japan*
[2] *Toho University, 2-11-1 Miyama, Funabashi, 2748510, Japan*
[3] *Toho University, 2-11-1 Miyama, Funabashi, 2748510, Japan*

Abstract
With the wide-spread availability of high-speed internet, an HTML-based system is one of the most promising options for mathematical software in this post COVID-19 era. While many web-based software libraries are available for specific purposes, including ones for symbolic calculations and dynamic visualizations, extra elaboration is needed to integrate their capabilities so that users can perform mathematical activities on a single HTML content. Moreover, it seems that the realities and benefits derived from the use of those systems have not yet been fully investigated. Especially in the complex case of statistics learning, more precise knowledge about them will likely help maximize the effect of using newly developed systems. This research is mainly concerned with the use of our newly developed system for the learning of the χ^2 test for independence. Through the comparative study of the log of simulations, which learners performed on the HTML content between experimental and control groups, it has been shown that some restrictions to the context of its use can have a large influence on learners' thinking.

Keywords
HTML5, CindyJS, dynamic visualization, Jspreadsheet, Apache ECharts, Numerical Recipes

1. Introduction

Although statistical methods are widely applied to analyze massive data derived from our daily lives, it is not easy for average university students to understand the theoretical background of those methods. Among other methods, the concept of statistical tests is one of the most difficult in statistics. Heavy cognitive load is imposed on learners' working memory because various modalities, including the verbal formulation of the argument, the symbolic calculation of the statistical quantity, and related visualizations, are involved in the completion of statistical tests ([1]). Moreover, in a traditional paper-and-pencil-based environment, it is not easy for learners to perform the interactive simulation in which they change the input data and examine the corresponding test output. To overcome these difficulties, we produced an interactive content for the learning of χ^2 test for independence which enables learners to perform the

MathUI'24: 15th MathUI Workshop 2024, August 5–9, 2024, Montreal, Canada
*Corresponding author.
†These authors contributed equally.
kitamoto@yamaguchi-u.ac.jp (T. Kitamoto); noda@c.sci.toho-u.ac.jp (T. Noda); masataka.kaneko@phar.toho-u.ac.jp (M. Kaneko)

interactive simulation. From the authors' experiences, it is not easy for them to understand the interrelation between the verbal meaning of the "independence" and the proportional allocation of numerical entries in the cross table. As a conceptual metaphor, the content is used to search for the condition in which a drug is regarded effective (i.e. the recovery rate is improved by using the drug). Since the content is HTML5-based and imports the relevant libraries via the internet, it can be used on a browser which is implemented not only on PCs but also on tablets and smartphones, without any installation of applications. To lessen learners' cognitive load, our newly developed system ([2]) uses various JavaScript libraries for functions including spreadsheet-style data input, natural expressions of mathematical expressions, and rendering of 3D bar graphs and 2D graphs, which are subsequently integrated into a single HTML content. Though some dynamic geometry systems like GeoGebra (https://www.geogebra.org) have been equipped with the function of exporting its geometric output in HTML format, some extra effort is needed for the bidirectional exchange of data between dynamic geometry component and other components in a single HTML content. Our system provides an easily accessible interface to unify various types of components in a single HTML content and to set up the neccessary data exchange between them.

In this paper, we will show the features of the HTML content created by using this system together with a sample of its use. The great emphasis is laid on the influence of the context of its use on learners' thinking.

2. Features of the teaching material created

Figure 1 shows a screenshot of the created content. The 2×2 cross table is displayed in the upper left area with the attributes "drug group" and "control group" in rows and "cured" and "not cured" in columns respectively. While it is possible to set all four entries to be variable, we set some restrictions in this case. Entries in B3 and C3 are fixed to be 60 and 40, and C2 is automatically decided when users input some number into B2 so that the sum of B2 and C2 is 100. First, when learners enter some number into B2, the table of expected frequencies is displayed in the upper right area. Then, by pushing the button "Graph Combining", the data in these two tables are visualized by 3D bar graphs in the middle left area. Since the difference between these two tables is important for the calculation of the χ^2 statistic, these two 3D bar graphs are superposed in the middle right area. Additionally, the processes of calculating the χ^2 statistic are rendered. Although Yate's correction for continuity is needed to correct the error introduced by assuming that the discrete probabilities of frequencies in the contingency table can be approximated by a continuous χ^2 distribution, we did not use it because the explanation of this mechanism might impose some excessive cognitive load on learners' working memory. Moreover, a graph of the density function for the χ^2 distribution is generated in the lower area. The χ^2 value and the corresponding p-value are computed and rendered in that graph. So that learners can identify the conclusion of the hypothesis test only by watching the figure. When new data are reentered into the cross table, the whole process is automatically repeated. Through the repetition of this simulation, students are expected to observe that large residuals of the entries in cross table from their expected frequencies implies the strong correlation between the two attributes.

Go to edit mode

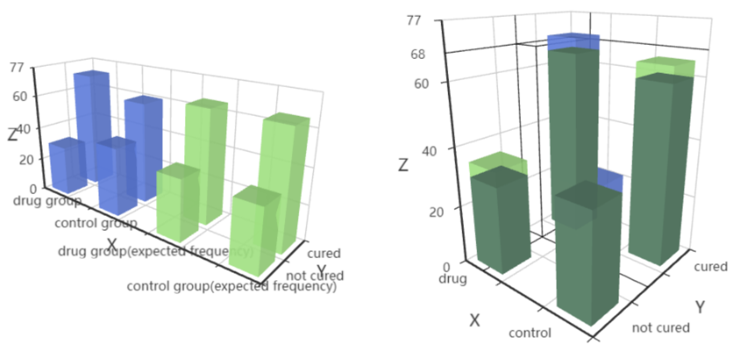
Enter data in cells B2, B3, C2, and C3 in the table below.

	A	B	C	D
1		cured	not cured	sum
2	drug grou	70	30	100
3	control gr	60	40	100
4	sum	130	70	200

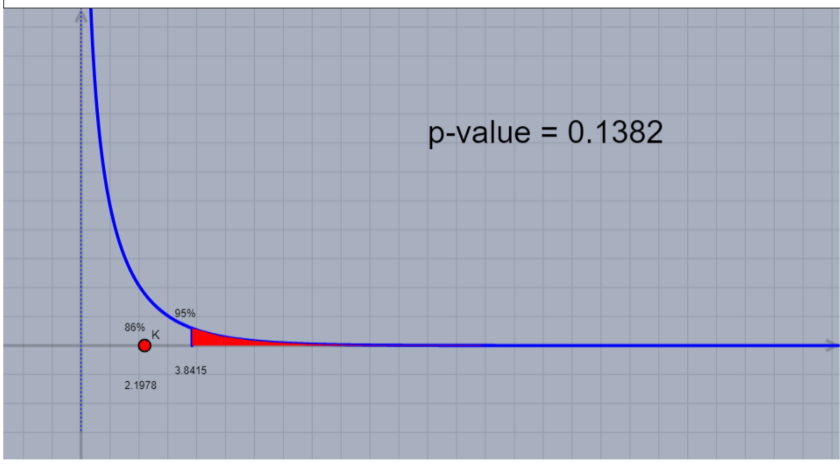
Table of expected frequency

	A	B	C	D
1		cured	not cured	sum
2	drug grou	65	35	100
3	control gr	65	35	100
4	sum	130	70	200

Graph Combining



$$
\begin{aligned}
\chi^2 &= \sum_{i,j} \frac{(a_{ij} - e_{ij})^2}{e_{ij}} \\
&= \frac{(a_{11} - e_{11})^2}{e_{11}} + \frac{(a_{12} - e_{12})^2}{e_{12}} + \frac{(a_{21} - e_{21})^2}{e_{21}} + \frac{(a_{22} - e_{22})^2}{e_{22}} \\
&= \frac{(70 - 65)^2}{65} + \frac{(30 - 35)^2}{35} + \frac{(60 - 65)^2}{65} + \frac{(40 - 35)^2}{35} \\
&= 2.2
\end{aligned}
$$



Send

Fig. 1: Screenshot of the created content

The overall picture of the system to create this content is sumarized in the top diagram of Figure 2. The figure below it shows the interface used to create HTML content in our system.

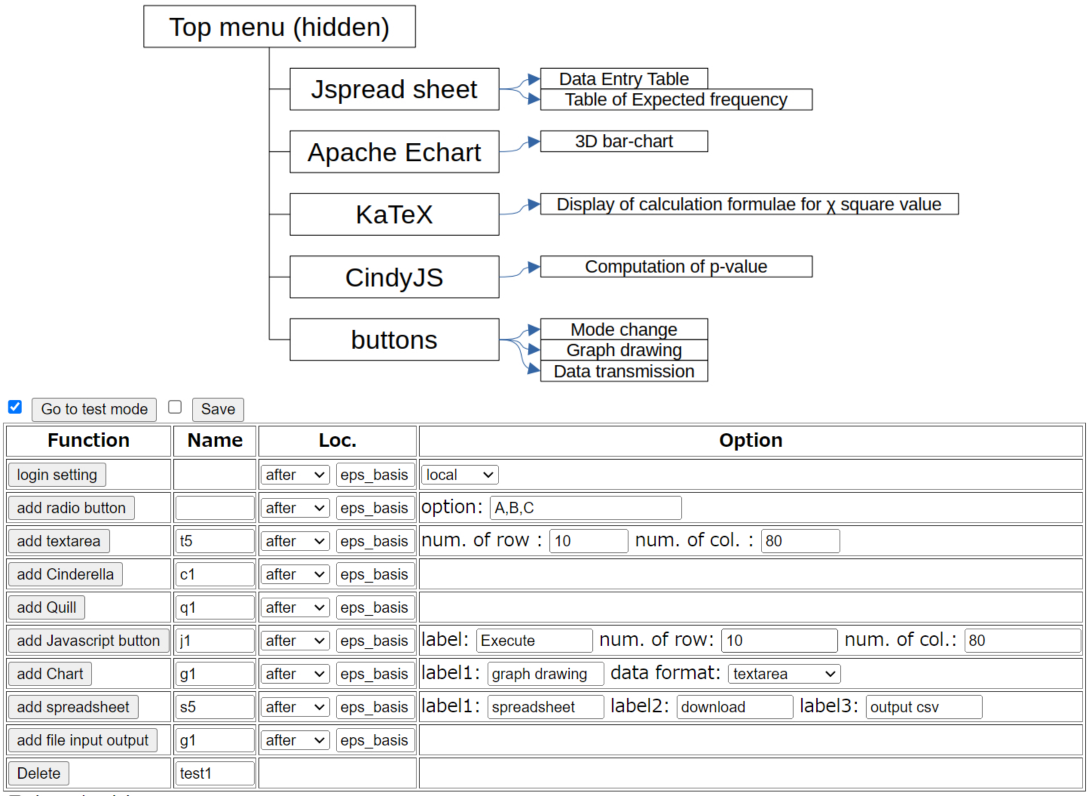


Fig. 2: Summary and the interface of the system to create the HTML content

The features of this interactive content and the technical backgrounds for creating it are as follows.

1. **Data entry in spreadsheet format in a browser**
   As shown in the cross table, learners can input data exactly in the same way as they would input data into a spreadsheet (moving cells with the arrow keys on the keyboard and copying and pasting are also possible). This function is realized by using a Javascript library JspreadSheet CE (https://bossanova.uk/jspreadsheet/v4/) under the MIT License.
2. **High-quality expression of mathematical formulas**
   As shown in the process of calculating the χ^2 value, both numeric and symbolic expressions are rendered quickly with the same quality as Donald Knuth's $\TeX$. This function is realized using KaTeX (https://katex.org/) which is a Javascript library that renders TeX sources in the browser.
3. **3D Graphs**
   3D graphs are rendered using Apache Echats (https://echarts.apache.org/en/index.html), a Javascript library for drawing various types of graphs, including 3D graphs, in a browser under the Apache License 2.0. Since these 3D graphs are movable by clicking and dragging, learners can view them from various perspectives.

4. **2D Graphs**
The χ^2 distribution is plotted by using the interactive geometry system CindyJS (https://cindyjs.org/, [3]), a Javascript library developed for the dynamic geometry software Cinderella ([4]). The relevant items including the rejection region, χ^2 values, and p-values, are calculated and plotted by porting the source code of Numerical Recipes in C (https://s3.amazonaws.com/nrbook.com/book_C210.html) to CindyJS.
5. **Logging function**
The learners' inputs are logged into their browsers with time stamps. The logged data is sent to the server when the user clicks on the "Send" button in Figure 1 after completing their simulations and is subsequently used to trace the learners' thinking.

3. The Flow of the Use of the Content

Concerning the innovative educational approach including ICT use, there have been many debates about the extent to which teachers' guidance shold be given during the process of students' inquiry-based learning ([5][6]). One important factor of that guidance is the context in which students use ICT tools. To investigate the interrelation between the context in which learners use this content and their thinking, we conducted a comparative study of the learning outcomes of the two groups of students who used the content in different contexts. The subjects were 221 sophomores (20 years old) from a Japanese university whose majors are not mathematical science but pharmaceutical science. They were randomly divided into Group A (80 females and 29 males) and Group B (76 females and 36 males). While both group of students took the statistics class with similar lesson plans, only the contexts in which they used the content were slightly different between A and B.

Our pilot study using this content indicated that many students find it difficult to unify the increase and decrease of χ^2 statistic and the literal meaning of independence and correlation. Regarding this result, we asked students to input values into the cross table of the content and search for the most probable integer inputs so that the resulting p-values are nearest to $0.9,\ 0.1,\ 0.05$, and 0.01, respectively. Through this simulation process, students were expected to understand that the increase of the entry in B2 cell leads to the stronger correlation of the two attributes and unify that knowledge with the mathematical background of χ^2 test. While no recommendation was given to the students in Group A concerning the range of input, the students in Group B were strongly encouraged to search for the integers between 60 and 100. Therefore, the simulations made by the students in Group A was "goal-free" in a sense, and this flexible context was supposed to impose a larger cognitive load on their working memory compared to Group B ([1]). Though the students in Group B were permitted to input the numbers outside the recommended range, this rigid context was supposed to draw more attention to the procedure for calculating the χ^2 statistic compared to Group A. Typical examples of students' inputs are summarized in Figure 3 (left) from which it can be seen that the favorable range of inputs are the integers between 60 and 80. Figure 3 (right) is a sample visualization of how two students simulated. Here the horizontal axis represents the passage of time, and the vertical axis represents the value of the input into B2 entry. In this figure, the above-mentioned favorable range is specified as the strip bounded by the light blue lines.

cured	not cured	p-value
60	40	1
61	39	0.885
62	38	0.7719
63	37	0.6629
64	36	0.5601
65	35	0.4652
66	34	0.3795
67	33	0.3039
68	32	0.2386
69	31	0.1835
70	30	0.1382
71	29	0.1018
72	28	0.0733
73	27	0.0515
74	26	0.0353
75	25	0.0235
76	24	0.0153
77	23	0.0097
78	22	0.0059

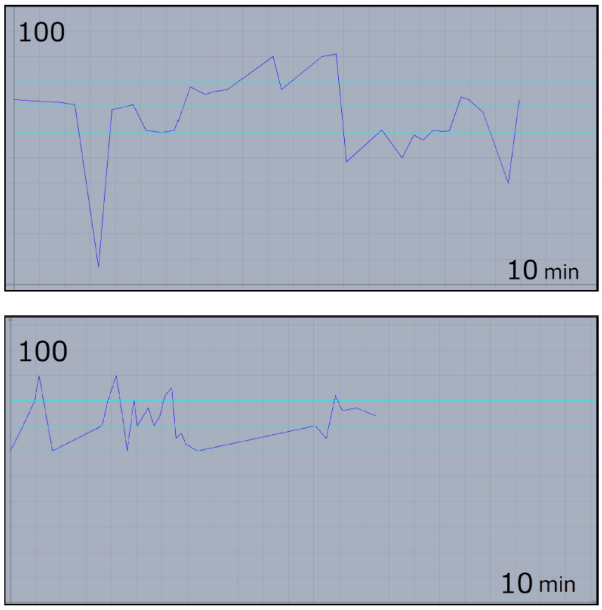


Fig. 3: Summary table and visualization of students' simulation processes

After completing the simulation, students were asked to take a posttest which was designed to evaluate their learning outcomes. While the above summary table was being presented, students were asked to answer the inputs for the same p-values when the entries in B3 and C3 were changed to 62 and 38. The definition of the χ^2 statistic implies that, if 2 is added to the B2 entry and is subtracted from the C2 entry then, all numerators $(a_{ij} - e_{ij})^2$ of each term in the χ^2 statistic will be maintained. Considering that $\frac{1}{e_{ij} \pm 2}$ is nearly equal to $\frac{1}{e_{ij}}$ in this case, the above modification of B2 and C2 gives almost the same χ^2 statistic. The students who paid attention to the interrelation between the calculation process of the χ^2 statistic and the literal meaning of correlation during their simulation could be assumed to come to this conclusion.

Also, we used the log data of students' simulations to evaluate their thinking. It is natural to think that students whose inputs are far from the favorable range most likely have difficulty understanding the mechanism. Moreover, it can be assumed that the inputs at later stages of the whole simulation process may have more significance than those at early stages. Therefore, we computed the following signal from each student's simulation log. As shown in Figure 4, when the i-th input is x_i $(i = 1,\ 2,\ \cdots,\ n)$, its deviation d_i from the favorable range is defined as:

$$d_i = \begin{cases} x_i - 80 & (80 < x_i) \\ 0 & (60 \leqq x_i \leqq 80) \\ 60 - x_i & (x_i < 60) \end{cases}$$

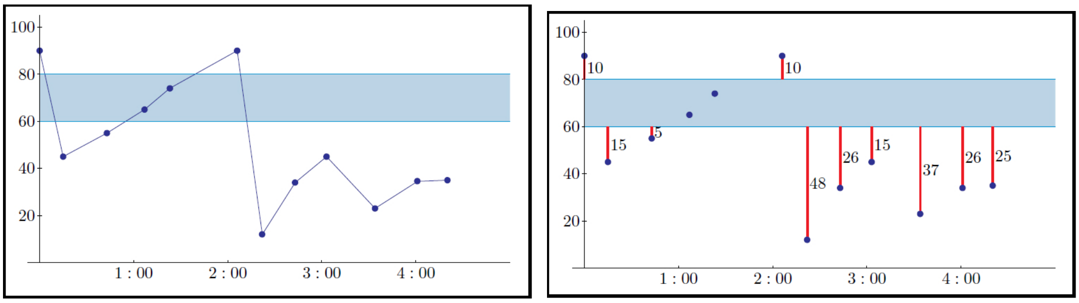


Fig. 4: Definition of the deviation of students' inputs from the favorable region

We defined the signal as follows:

$$\frac{\sum_{i=1}^{n}(i \times d_i)}{\sum_{i=1}^{n} i} \tag{1}$$

For instance, in case of Figure 4, the signal is calculated as

$$\frac{1 \times 10 + 2 \times 15 + 3 \times 5 + \cdots + 11 \times 26 + 12 \times 25}{1 + 2 + 3 + \cdots + 12}$$

We compared the distribution of students' signals with respect to groups A and B together with their performance at the posttest.

4. Results

If a student answered more than three questions correctly with any errors occurring only in cases of $p = 0.9$ or $p = 0.01$, their posttest performance was considered superior. Any results that did not meet these criteria were evaluated as inferior. The results of the posttest are summarized in the following table.

Table 1: Result of the posttest

	Group A	Group B	total
superior	21	36	57
inferior	88	76	164
total	109	112	221

The p-value, derived from Fisher's exact test, is 0.03201, which indicates that the rigid context set up for the simulation of the students in Group B had a positive impact on the posttest performance.

Though the duration of the whole simulation process was slightly shorter in the case of Group B compared to Group A, there was no significant difference in the time-averaged fluctuation of inputs between A and B. However, students in Group B with superior performance on the posttest made inputs concentrated in the neighborhood of the favorable range. Figure 5 shows sample processes of simulation carried out by those students in groups A (left) and B (right).

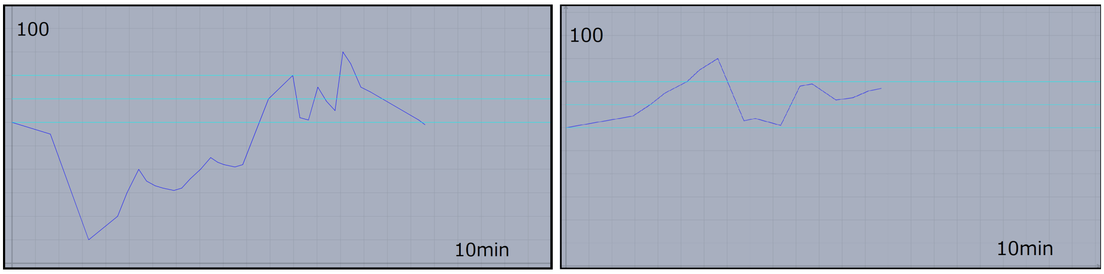


Fig. 5: Simulation process of some students with superior posttest performance

The distribution of the signals defined by the formula (1) in section 3 is displayed in the histogram of Figure 6 in which the left is of Group A and the right is of Group B. For convenience, cases exceeding 40 are accumulated at the class with highest value of the histogram. The distribution of the signal corresponding to superior performance on the posttest is colored blue and that corresponding to inferior performance is colored dark orange. In both histgrams, blue and dark orange distributions are overlapped.

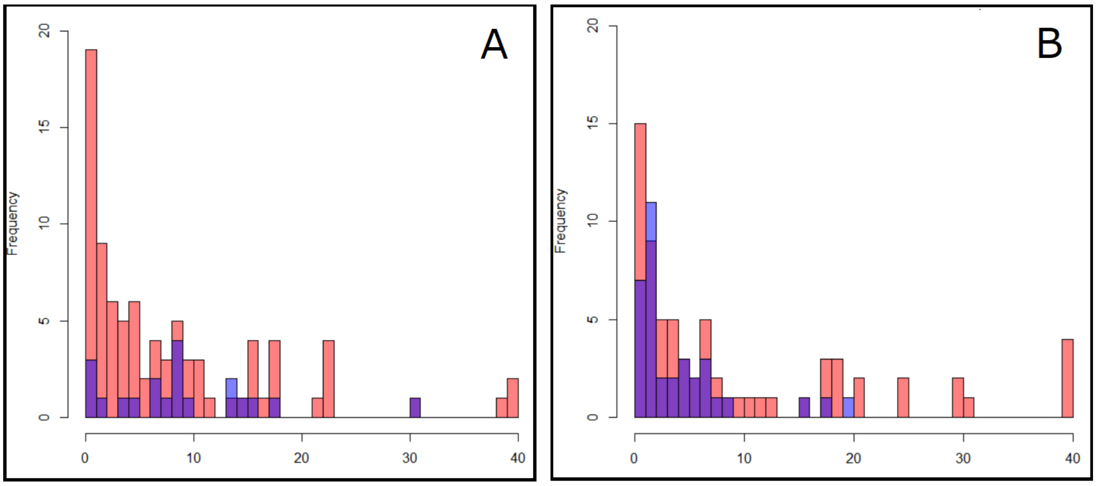


Fig. 6: Distributions of the signals derived from the simulation log

Figure 6 indicates that, while the purple data and the dark orange data are evenly distributed in the case of Group A, the purple data are densely distributed in a smaller region compared to the dark orange data in the case of Group B. The p-values derived by applying Wilcoxon rank-sum test to the blue data and the dark orange data were 0.8864 and 0.0351 in the cases of Group A and Group B respectively. This implies that, while the whole distribution of signals is not different between the two groups, the students in Group B whose signals were low improved their posttest performance remarkably.

5. Discussion

In this study, we developed HTML-based content to help students understand a statistical test and investigated how students use it. Various functions based on Javascript libraries were integrated into a single HTML content to lessen the cognitive load imposed on learners' working memory.

Although the results of the posttest analysis suggest that contextualizing and controlling the simulation process could have a positive impact on students' learning outcomes, further data is required to fully understand how the contextualization affects students' thinking. The results of this study strongly indicate that the log data of learners' simulations can serve as a useful tool for that purpose. For instance, the process from the contextual change to the improvement of learning outcomes is assumed to be as follows:

1. Guidance about the range of input made students recognize the meaning of the entries in the contingency table.
2. Recognition of that meaning guides some students input the integers near the favorable range. This made it easy for students to see how the increase and decrease of input are related to those of the χ^2 statistic.
3. Some students who observed the above-mentioned relation between input and the χ^2 statistic could consider the reason for it while referring to the calculation process of the χ^2 statistic.
4. For the students who were conscious of the form of deviation embedded in the definition of the χ^2 statistic, it would not be so difficult to give correct answers at the posttest.

Conclusively, when the signal derived from some student's simulation process is relatively large, it can be assumed that the student is at risk and should be encouraged to turn her attention to the process for calculating the χ^2 statistic.

6. Concluding Remarks and Future Work

In many topics of the statistical test, it can be expected that similar HTML-based interactive content is effective for the learning those concepts. The authors are planning to use our system to create many HTML-based contents corresponding to those topics. The result of this study strongly indicates that only creating the content is not sufficient. The information about when and how to use it together with its cognitive load theoretic background should be needed. For instance, in case of this study, it should be warned that students' simulation without any guidance might impose heavy cognitive load to them because a random movement of their imput can prevent them from observing some interrelation between verbal aspect and numerical aspect of χ^2 test.

There are some limitations in this study. One major limitation is that we did not analyze students' thinking through their verbal statements. To carry out further research in this direction, we will set up a computer-supported collaborative learning (CSCL) environment and analyze learners' verbal communications. In that case, learners' nonverbal communication including their gestures or body languages should also be analyzed. We are now planning to use recent sensing technologies to monitor those activities. Figure 7 shows our trial of using a depth sensor camera to detect the movements of participants' hands and heads in a CSCL environment ([7][8]).

Fig. 7: Use of depth sensor camera to detect learners' movements

Also, there are some functions to be improved in our content creation system. For instance, there has not been implemented the function for flexible page layout. In fact, though we have been able to control the order of rendering several components on a single HTML content, we are not able to arrange the precise positions where they are located. We are now developing the function to interactively set up the position of each component by clicking and dragging while monitoring the final output. Figure 8 shows the screenshots of the editing process (left) and the final output (right) where the position and size of answer columns and supplementary descriptions can be changed interactively.

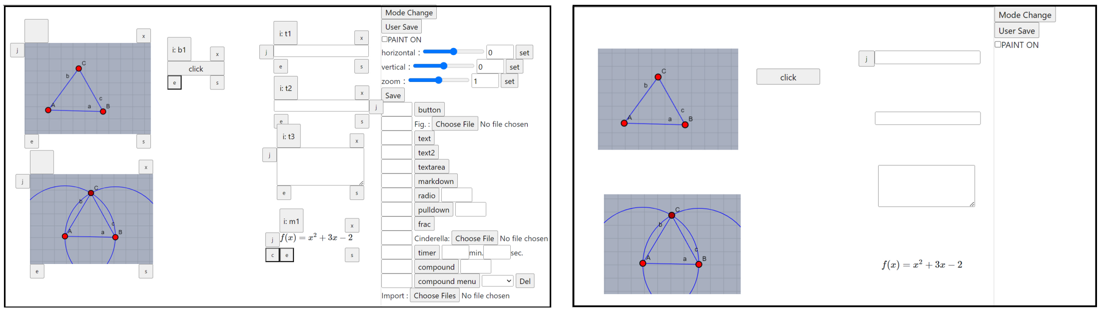


Fig. 8: Function for interactive arrangement of various web components

Acknowledgments

This work was supported by JSPS KAKENHI (21K02752 and 22K02951).

References

[1] J. Sweller, P. Ayres, S. Kalyuga: Cognitive Load Theory, Springer (2011)

[2] T. Kitamoto, M. Kaneko, S. Takato: E-learning system with Computer Algebra based on JavaScript programming language, Proceedings of ATCM2018, 123-132 (2018)

[3] M. von Gagern, U. Kortenkamp, J. Richter-Gebert, M. Strobel: CindyJS – mathematical visualization on modern devices, Lecture Notes in Computer Science 9725, 319–326 (2016)
[4] J. Richter-Gebert, U. Kortenkamp: Interactive Geometry with Cinderella, Springer (2012)
[5] P. Kirschner, J. Sweller, R. Clark: Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experimental, and inquiry-based teaching, Educational Psychologist, 41(2), 75-86 (2006)
[6] C. Hmelo-Silver, R. G. Duncan, C. Chinn: Scaffolding and achievement in problem-based and inquirylearning: A response to Kirschner, Sweller, and Clark (2006), Educational Psychologist, 42(2), 99-107 (2007)
[7] M. Kaneko, H. Egi, T. Kitamoto, T. Noda: The multimodal analysis of learners' mathematical proof activity in a CSCL environment, Proceedings of CSCL2024, 383-384 (2024)
[8] S. Yamamoto, R. Funabashi, T. Noda, M. Kaneko, H. Egi: The multimodal analysis of learners' mathematical proof activity in a CSCL environment, Companion Proceedings of LAK2024, (2024)