Croatian Journal of Education
Vol.23; No.3/2021, pages: 903-934
Preliminary communication
Paper submitted: 26th April 2020
Paper accepted: 23rd February 2021
https://doi.org/10.15516/cje.v23i3.4018

A Delphi Study and Verification of Assessment of Chemistry Experiments Videos

Cui Liu[1], Sufen Chen[2, 3], Shuo Liu[4], Jingyuan Liu[4], Jie Wang[4] and Yingying Han[4]
[1]*College of Teacher Education, Jiangsu Normal University*
[2]*Graduate Institute of Digital Learning and Education, National Taiwan University of Science and Technology*
[3]*Optentia Research Focus Area, North-West University*
[4]*School of Chemistry & Materials Science, Jiangsu Normal University*

Abstract

The purpose of this study was to use the Delphi method to combine consensus opinion to develop indicators for pre-service chemistry teachers to assess videos of chemical experiments. This study had two stages. The first was to construct the Video Assessment Questionnaire using a modified Delphi consensus process delivering three rounds of face-to-face or online surveys. The 50 participants included professors, associate professors, undergraduate students, and middle school teachers. The Video Assessment Questionnaire has a high degree of consensus as shown by the coefficients of variation and Kendall's W of the second and third rounds. The final version included 20 indicators in three major aspects: content organization, experiment operation, and video quality. In the second stage, a quasi-experiment was designed to test the effect of using the questionnaire in Chemistry Teaching Methodology Course. The experimental group (n = 22) used the questionnaire to evaluate their own videos and reflect on their performance, whereas the control group (n = 11) carried out the same tasks without the tool. We found that the pre-service teachers who used the Video Assessment Questionnaire had better laboratory instruction ($t(31) = 4.25, p<.001$), especially the operation in experiments ($t(31) = 4.95, p<.001$) than those who only made videos. This study verified the effectiveness of the Video Assessment Questionnaire in terms of helping pre-service chemistry teachers to improve experiment operations skills.

Keywords: *assessment; chemistry education research; laboratory instruction; multimedia-based learning.*

Introduction

The purpose of this study was to develop indicators for pre- and in-service chemistry teachers to assess their own videos of chemical experiments, in order to improve their experiment teaching level. The use of videos in teacher education can help pre-service teachers share experiences, focus attention on specific sequences of behaviors, and identify merits and faults to improve the teaching practices (Fadde & Sullivan, 2013; Marsh et al., 2010; Masats & Dooly, 2011). Researchers have used videos to bridge the gap between scientific theory and practice (e.g., Bencze et al., 2001; Stoll et al., 2006; Seidel et al., 2013). Specifically, videos can demonstrate the implication of a scientific concept and various experiment skills. Moreover, videos can be incorporated into different pedagogical methods. For instance, videos in combination with self-reflection might help an in- or pre-service teacher identify a problem (Christ et al., 2017), and reflecting on a video with peers promotes problem-solving ability (Arya et al., 2015). Video assessment can serve as feedback and feedforward. It is a key element in formative assessment because it involves pre-service teachers in thinking about the quality of their own work, rather than relying on their instructors as the sole source of evaluative judgments (Heidi & Anna, 2009). However, no systematic analysis of the quality of chemistry videos has been done, leading to pre-service teachers' lack of scientific methods to evaluate their own and their peers' videos. In this study, we used the Delphi method to combine consensus opinion to develop indicators of video assessment of chemistry experiments.

Literature Review

Video observation and learning

Video usage has a significant positive effect on learning performance and learning satisfaction (Nagy, 2018), as well as on teaching effectiveness and course satisfaction (Tse et al., 2019). Videos are advantageous to the teaching and learning of practical work (Tano et al., 2017). They can be integrated into teaching at different stages. For instance, instructors may use them as examples (1) to illustrate certain rules, concepts, and principles or (2) to demonstrate phenomena which then lead to a rule. Seidel et al. (2013) found that, in the context of teaching pre-service teachers the pedagogical knowledge, the former facilitates the learning of factual knowledge and the evaluation of classroom teaching, whereas the latter improves lesson planning. Videos play various roles and have different effects.

Chemistry videos are widely used as teaching media in the classroom (Tierney et al., 2014), and recently more have been shared on YouTube (Blonder et al., 2013). They are easy to use. They can be paused, sped up, slowed down, and replayed so that students learn at a comfortable pace and can revisit the materials on demand (Kraft et al., 2012; Richards et al., 2014). Videos could reveal such clear images and focused phenomena or skills that sometimes they are more effective than human instruction.

Jordan et al. (2015) have shown that students who observed short and concise video instructions were able to work more independently throughout the laboratory and perform better than students who had received instruction from a teaching assistant.

Moreover, videos can provide rich visual and dynamic information (Fadde & Sullivan, 2013; Marsh et al., 2010). In a video, learners can clearly observe the performance of oral expression, dynamic experiment operation, and content in multiple representations such as graphs, animations, texts, slow motion, and so on. Videos of chemistry experiments facilitate teaching and reduce concerns of chemical waste, safety, and ambiguousness/hazy manners.

Although chemistry videos have been used for many years and have positive effects on learning (Tierney et al., 2014), rarely have researchers proposed a Video Assessment Scale for chemistry experiments and empirically investigated its effectiveness for pre-service teachers (Christ et al., 2017). An effective assessment scale conveys the standards and expectations of teaching, and promotes self-assessment by pre-service teachers (Beacher et al., 2013).

Self-Assessment

As one of the assessment methods, self-assessmen is a good way for pre-service teachers to reflect on and learn from their experience. Many researchers have studied the impact of self-assessment (Kerr & Bruun, 1983; McDowell, 1995). Self-assessment gives students a chance to evaluate their own accomplishments and enhances self-efficacy, as well as self-regulated learning (Raaijmakers et al., 2019; Yan et al., 2019; Zimmerman, 2000). For example, students who conducted self-assessment expressed more positive attitudes toward and confidence in designing experiments compared to those who had teacher evaluation or no evaluation (Olina & Sullivan, 2004). Moreover, Paris and Paris (2001) propose that self-assessment often involves the internalization and processing of a wide variety of standards, which may help students regulate their learning in a more effective, insightful, and accurate way. Different types of self-assessment have different functions. Predictive self-assessment refers to the ability of predicting future performance. Summative self-assessment is used to rate performance after having completed the learning activities. Concurrent self-assessment is the ability to identify gaps between the actual and desirable performance and the learning needs during the activities. Self-assessment using videos can be summative to the overall performance and concurrent to constantly regulating learning.

Baecher et al. (2013) compared pre-service English teachers' self-assessment of their teaching performance with and without using teaching videos. Although both groups received the same evaluation rubrics, the use of videos increased the pre-service teachers' understanding of quality teaching in the field and reduced overestimation of their own teaching performance. Likewise, Tano et al., (2017) let students monitor their dental examination skills via video recordings. The video monitoring significantly improved the students' practical skills and the consistency between their self-evaluation and the instructors' evaluation. In general, as revealed by LeFebvre et al., (2015) in their basic communication courses using video self-assessment, students tend to overestimate their performance. Nevertheless, rubrics, videos, and self-assessment promote self-regulative learning and reduce the level of overestimation.

Furthermore, Korkmaz et al. (2016) asked students to assess their own performance by recording videos and found that students using videotapes were more likely to spot insufficient/incorrect practices during the self-assessment process. Researchers found that video is intuitive and easy to observe and remember. Video feedback seems to be a promising alternative to traditional written feedback (Mahoney et al., 2019). In the process of self-assessment, pre-service teachers play two roles – those of students and reviewers. Overcoming self-consciousness in order to observe oneself critically is a substantial challenge, especially when self-observation involves their teaching videos (Fadde & Sullivan, 2013). They often struggle to get beyond focusing on their own lesson delivery to reinterpret their oral expression, experiment operations and principle explanations. It is expected that, through self-assessment, pre-service teachers understand the experiment teaching design, experiment teaching organization, experiment teaching innovation and learning content, and make constructive evaluation of their own performance. The use of a self-assessment questionnaire may achieve these objectives and help them make a comprehensive assessment of their own videos.

Research Questions

This study used the Delphi method to construct a Video Assessment Scale for pre-service teachers to evaluate teaching videos in chemistry, investigate the use of the scale, and seek to answer the following questions:

1. What are important indicators of evaluating chemistry experiment videos?
2. Does the Video Assessment Scale promote creating chemistry experiment videos?
3. What are pre-service teachers' perceptions of using the Video Assessment Scale?

Study Design

The use of the Delphi method to create a teaching tool has been well recognized (Ogden et al., 2016). However, it has been criticized for the lack of methodological rigor due to the flexibility of its technique. This study had two stages to enhance the reliability. The first stage was to construct the Video Assessment Questionnaire. The Delphi technique consisted of three consecutive rounds to collect experts' opinions in order to obtain the most reliable consensus. In the second stage, pre-service teachers used the questionnaire to evaluate their own videos and reflect on their performance. Although there is no universal guideline for Delphi, it is strongly recommended that key action steps be taken and reported clearly (Hasson et al., 2000). Figure 1 shows the key actions and the results in each round.

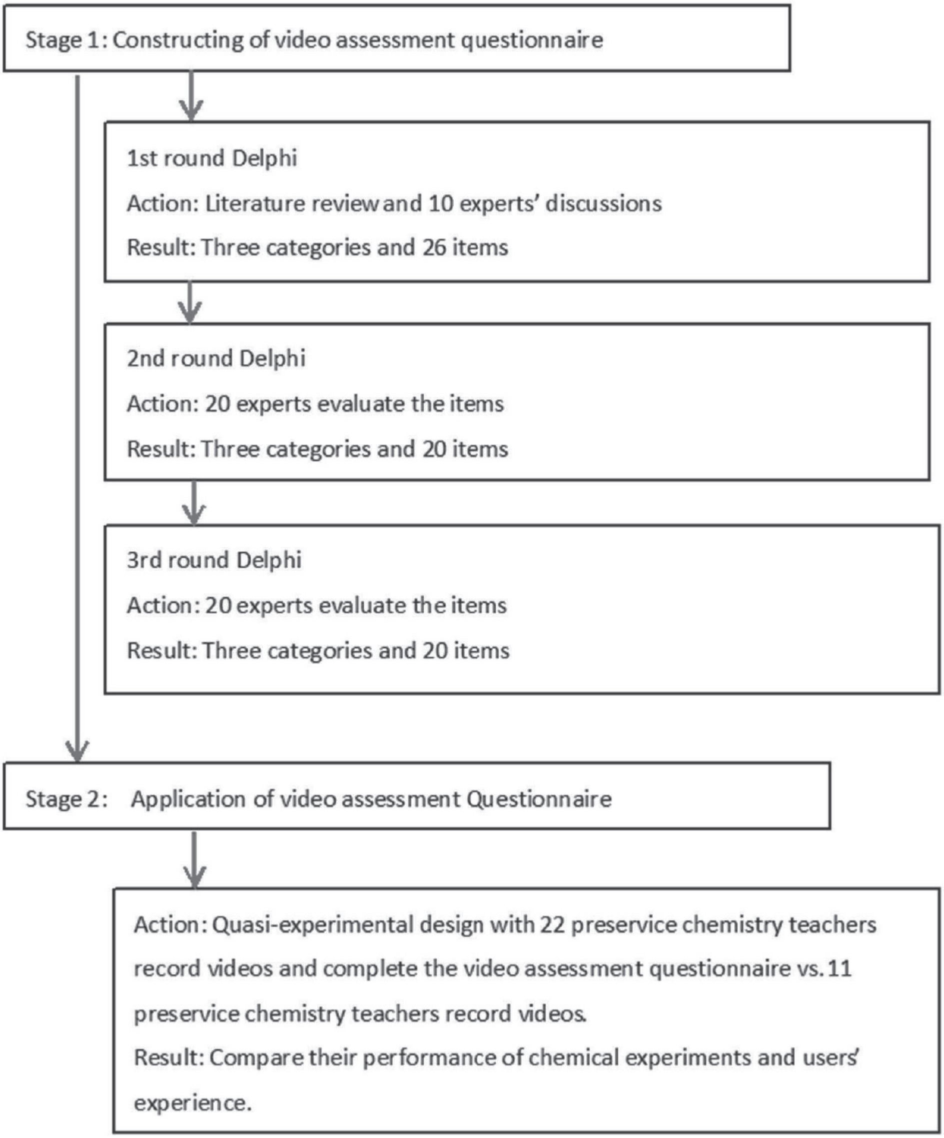

Figure 1. Key actions and results of the Delphi study

Stage 1: Construction of the Video Assessment Questionnaire

The first round of the Delphi research. The Delphi panel of the first round was comprised of ten educational leaders from four provinces of China, including four chemists, two educational technologists, and four chemistry educators. These experts had more than five years of research in chemistry or chemistry education and more than one year of experience in making videos and offered web-based courses. They were invited to list individually the important indicators of chemistry experiment videos. We provided books (Chen et al., 2015; Chen & Tian, 2014), papers (Li et al., 2017; Seidel et al., 2013; Seshadrinathan, Seshadrinathan et al., 2010), and five chemistry experiment videos to the panel. They put forward their own views and explained the basis and reasons for their views. They then discussed and exchanged opinions and agreed on three domains of indicators: content organization, operation in experiments, and video quality, consisting of 26 items in total.

The second and third rounds of the Delphi research. In the second and third rounds, college teachers, undergraduate students and middle school teachers in the field of chemistry were invited to evaluate the appropriateness and importance of the indicators. They all had more than one year of experience making videos. The two rounds each involved 20 different experts.

The second round was delivered via a word spreadsheet. Experts were required to rate the importance of each indicator. Their rating was used to decide whether the index should be excluded or not. A five-point scale was applied in the study, from the most important to the least important. The mean, the percentage of agreement, and standard deviation were used to evaluate the collective judgments of the experts. The coefficient of variation (CV) and the Kendall's coefficient of concordance (Kendall's W) were applied to represent the variability of the experts' opinions. In Kendall's W, a mean score of 4 (indicating relative importance), an 80 % agreement, a SD of 1.2, a CV of 0.2 and $p < .05$ were used as the criteria for consensus (Balasubramanian & Agarwal, 2012; Hu et al., 2017). The indicators fulfilling these criteria were retained. The third round was carried out via an online platform, and the same criteria were used to decide on the final indicators.

The demographics of experts in each round are presented in Table 1.

Table 1
Demographics of Participants in each Round

	Round 1		Round 2		Round 3	
	n	%	*n*	%	*n*	%
Professor	6	60	0	0	1	5
Associate professor	4	40	5	25	6	30
Undergraduate student	0	0	3	15	4	20
Middle school teacher	0	0	12	60	9	45
Working experience[a]						
1-5 years	0	0	10	50	9	45
6-10 years	7	70	2	10	2	10
11-15 years	2	20	4	20	6	30
16-20 years	1	10	4	20	3	15

Note.[a]including the teaching practicum experience

Video Assessment Questionnaire. The indicators fall into three categories. The first category, regarding the content organization, evaluates the general organization and the specific content of the video (Malekian & Aliabadi, 2012), including rationality, attractiveness, reasonable structure, detailed content, guidance, innovation, time control, performance and protective clothing, oral expression, and chemicals. Three indicators such as interaction design, reference, and timely reflection were eliminated in the second round.

Concerning the second category, chemistry experiment operation is a complex system composed of many sub-factors according to certain logic and hierarchical relations. The sub-factors in the system are interrelated and interdependent. They are experiment instrument, chemicals, and operation (Hong, 1994). Therefore, this category includes standardized operation, demonstration, clear and smooth steps, phenomenon, scientific and experiment environment. All indicators passed the three rounds of Delphi.

The third category, video quality, is assessed based on the picture quality, picture content, editing technology, picture effect, sound effect, and synchronization coordination (Zhan, 2010). Indicators include screen quality, video editing technology, and clarity. Three indicators, i.e., video clip, captions, and special audio and video effects were eliminated in the second round.

The questionnaire also set open questions to collect other opinions from experts. The items are defined in Table 2. The Video Assessment Questionnaire scores 100 (5 points per indicator), of which Content Organization accounts for 50 %, Operation in Experiment for 35 %, and Video Quality for 15 %.

Table 2
Items and Descriptions

Indicators and items	Descriptions
Content organization	The general organization and the specific content of the video.
Rationality	The theme of the experiment is related to the science concepts and has the value of exploring.
Attractiveness	The experiment is novel and interesting.
Reasonable structure	The content of the video is complete, hierarchical and logical, which is in accordance with the students' cognitive law.
Detailed content	The experimental phenomenon is explained clearly, and the correct experimental conclusion is obtained.
Guidance	The teacher asks questions and leads students to think.
Innovation	The teacher adopts a variety of learning methods.

Indicators and items	Descriptions
Time control	The teacher completes the experiment operation within the specified time.
Performance and protective clothing	The teacher wears the protective clothing and behaves appropriately.
Oral expression	The teacher has clear articulation, concise language and accurate expression.
Chemicals	The experiment instrument is used correctly, and the amount of chemicals used is reasonable, with no waste.
Operation in experiment	Design, explanation and operation of chemical experiments.
Standardized operation	The experiment operation should follow the standardized and safe procedures.
Demonstration	Experiment operations should be exemplary.
Clear steps	Experiment methods and steps should be clear.
Smooth steps	The operation process should be in line with the best procedures.
Phenomenon	The experiment phenomenon is obvious.
Scientific	Experiment phenomena are consistent with scientific principles.
Environment	The experiment environment is quiet, and the experiment instruments and chemicals are neatly placed on the experiment table.
Video quality	The quality of video recording and editing.
Picture quality	The picture is clear. The play is smooth. The filming quality is stable. The appropriate scene can be used reasonably to vividly display the experiment materials and instruments.
Editing technology	The video is complete, and the editing method is appropriate. The algorithm of the video combination is logical. The transition of the video screen is natural and reasonable.
Clarity and synchronization	Teacher dubbing and background music are consistent with the presented science phenomena.

Stage 2: Application of Video Assessment Questionnaire in preservice Chemistry Teaching Methodology Course

The participants comprised 33 third-year students at Jiangsu Normal University, a major teacher preparation institute in one of the largest provinces in China. They were enrolled in the Chemistry Teaching Methodology Course. The mean age was 19.4 years. The majority were female students (n = 27.82 %). They were informed of the research procedure and signed a consent form.

The duration of the intervention was ten weeks, and students completed one experiment every week, starting from the second week. The weekly lesson lasted 3 hours. Each week the participants conducted a chemistry experiment (topics shown in Table 3), made an instructional video out of it, and turned in an experiment report that included self-assessment of the performance. The first 8 videos were basic chemistry experiments, in which students followed the laboratory manual in the course package. Students had to preview the scheduled experiment and design the related instruction for demonstrating the experiment in a video. In the class, they were introduced to the experiment procedures, they completed the experiment and finalized the instructional design. They then recorded the instruction of the experiment using smart phones and edited the video to 10-20 minutes. The video, along with an experiment report, was submitted to the teacher within one week.

The last video was an innovative experiment designed by each individual student. It was used to measure their ability to design an experiment and make a teaching video. Students were instructed to consult the literature, determine the topic, and design the plan. The instructor reviewed the plan and provided the chemicals and equipment.

Table 3
Topics of Laboratory Instruction

Number	Topics
Experiment 1	Determination of the Avogadro's constant
Experiment 2	Electrolyte solution
Experiment 3	Preparation and properties of colloids
Experiment 4	Effect of concentration temperature on chemical reaction rate
Experiment 5	Solubility of potassium nitrate
Experiment 6	Preparation and properties of oxygen
Experiment 7	Preparation of ethyl acetate
Experiment 8	Preparation and properties of hydrogen
Experiment 9	Designing own experiment

The participants were split into two groups: the control group (n = 11) and the experimental group (n = 22). The control group self-evaluated the performance in their videos, whereas the experimental group carried out self-assessment using the Video Assessment Questionnaire. In the last video, each student had to design a new chemistry experiment. It represented an integrated skill of making chemistry experiment videos. In addition to self-assessment, four researchers used the Video Assessment Questionnaire to rate all videos. The scores given by the four researchers represented students' ability of teaching chemistry experiments. The inter-rater reliability was also assessed based on the four researchers' data.

Finally, we distributed an evaluation questionnaire to the 22 students in the experimental group to understand whether the Video Assessment Questionnaire was easy to use and to determine the impact of the use of the Video Assessment Questionnaire on learning, followed by interviews.

Results

Important indicators of evaluating chemistry experiment videos and consensus of the experts

Both the second and third rounds of Delphi confirmed the importance of 20 out of 26 indicators. The means of all dimensions were higher than 4, ranging from relatively important to the most important (Table 4). The coefficients of variation of the third round of consultation were between 1 % and 11 %, which is lower than 0.2, and the experts had a high level of agreement ranging from 0.675 - 1.0. Kendall's W values of the 2nd and 3rd rounds were 359.98 and 328.98, $p < .001$.

Table 4
Agreement Score for each Item

	2nd Round				3rd Round			
Indicator	*M*	*SD*	*CV* (%)	Agreement rate (%)	*M*	*SD*	*CV* (%)	Agreement rate (%)
Content organization	4.44	0.32	7	81.92-100	4.44	0.06	1	100
Rationality	4.85	0.36	7		5.00	0.00	0	
Attractiveness	4.90	0.30	6		5.00	0.00	0	
Reasonable structure	4.25	0.43	10		4.10	0.30	7	
Detailed content	4.95	0.22	4		4.90	0.30	6	
Guidance	4.10	0.30	7		4.00	0.00	0	
Innovation	4.95	0.22	4		5.00	0.00	0	
Time control	4.10	0.30	7		4.00	0.00	0	
Performance and dress	4.00	0.00	0		4.00	0.00	0	
Oral expression	4.15	0.57	14		4.00	0.00	0	
Chemicals	4.10	0.54	13		4.00	0.00	0	
Operation in experiment	4.70	0.25	5	98.57-100	4.75	0.13	3	100
Standardized operation	4.95	0.22	4		5.00	0.00	0	
Demonstration	4.20	0.40	10		4.35	0.48	11	
Clear steps	4.85	0.36	7		4.95	0.22	4	
Smooth steps	5.00	0.00	0		5.00	0.00	0	
Phenomenon	5.00	0.00	0		5.00	0.00	0	
Scientific	4.90	0.30	6		4.95	0.22	4	
Environment	4.00	0.45	11		4.00	0.00	0	
Video quality	4.95	0.22	4	67.5-100	4.93	0.13	3	100
Screen quality	4.95	0.22	4		5.00	0.00	0	
Video editing technology	4.95	0.22	4		4.80	0.40	8	
Clarity and synchronization	4.95	0.22	4		5.00	0.00	0	

Effect of Video Assessment Questionnaire

Four researchers first evaluated four videos to establish inter-rater reliability. They reached an acceptable reliability with Kendall's $W = 0.74, 0.83, 0.81$, and 0.93, $p < .001$, respectively for the four categories. They then split to evaluate the videos. Each evaluated 74 or 75 videos.

Table 5 shows the two groups' scores in nine experiments. There was no significant difference in performance at the beginning of the use of the questionnaires, $t(31) = -1.49$ and 0.25, $p > .05$ for Experiments 1 and 2. With the increase in the number of uses, the results of the experimental group were significantly better than the results of the control group, except for Experiment 6 and Experiment 9. The sixth experiment was the preparation and properties of oxygen. The experiment operation was relatively simple, and the phenomenon was obvious. Therefore, there was no difference in the performance of the two groups of students. The ninth was a design experiment, which mainly examined students' innovative design capabilities. Overall, the t-test results show that the use of the scale can improve the experiment teaching performance of students with large effect sizes (Cohen's d = 0.91-1.59), but the effect on simple experiments such as Experiment 6 ($t(31) = -0.93$, $p = .362$) and innovative experiments such as Experiment 9 ($t(31) = 0.24$, $p = .812$) is not significant.

Table 5

Comparison of the Two Groups in the Laboratory Teaching and Video-making Performance

Videos	Groups	*n*	*M(SD)*	*t(31)*	*p*	*Cohen's d*
Experiment 1	Control group	11	85.91(2.64)	-1.49	.146	0.56
	Experimental group	22	87.36(2.51)			
Experiment 2	Control group	11	87.55(2.71)	0.25	.801	0.09
	Experimental group	22	87.32(2.14)			
Experiment 3	Control group	11	86.18(2.69)	-4.31	.001	1.59
	Experimental group	22	90.09(2.21)			
Experiment 4	Control group	11	88.00(1.48)	-3.39	.002	1.28
	Experimental group	22	89.86(1.42)			
Experiment 5	Control group	11	88.64(1.43)	-2.64	.013	0.91
	Experimental group	22	90.55(2.59)			
Experiment 6	Control group	11	90.73(1.91)	-0.93	.362	0.35
	Experimental group	22	91.41(1.95)			
Experiment 7	Control group	11	89.18(1.80)	-2.66	.012	0.96
	Experimental group	22	90.68(1.29)			

Videos	Groups	n	M(SD)	t(31)	p	Cohen's d
Experiment 8	Control group	11	88.73(1.05)	-2.85	.008	1.12
	Experimental group	22	90.09(1.35)			
Experiment 9	Control group	11	90.09(2.11)	0.24	.812	0.09
	Experimental group	22	89.91(1.93)			

Table 6 provides the summary of the nine experiments by the sub-scales. The t-tests showed that the experimental group, who used the Video Assessment Questionnaires, had better skills in laboratory instruction than the control group, who self-assessed videos without the questionnaire (t(31) = -4.25, p = .001). The effect size was large (d = 1.34). In particular, the abilities of experiment operation were significantly higher when using the video assessment questionnaire for self-assessment (t(31) = -4.95, p = .001, d = 1.96). However, content organization (t(31) = -1.39, p = .174) and video quality (t(31) = -1.29, p = .207) were not significantly different between the groups. The effect of the use of the questionnaire was mainly on managing the procedures and operations of experiments.

Table 6
Comparison of the Two Groups in the Sub-scales

Sub-scales	Control group M(SD)	Experimental group M(SD)	t(31)	p	Cohen's d
Content Organization	44.16(0.42)	44.43(0.56)	-1.39	.174	0.55
Operation in experiment	31.05(0.39)	32.04(0.60)	-4.95	.001	1.96
Video quality	13.12(0.31)	13.23(0.48)	-1.29	.207	0.27
Total score	88.33(0.83)	89.70(1.19)	-4.25	.001	1.34

Pre-service teachers' perceptions of using the Video Assessment Questionnaire

We distributed an evaluation questionnaire to the 22 students in the experimental group. The results of the survey are summarized in Table 7.

The results of the questionnaire showed that most students (72.73 %) thought that the questionnaire was easy to use, and a few students (27.27 %) thought that the operation was inconvenient. Most students believed that the use of Video Assessment Questionnaire could help them improve experiment operation skills (90.91 %), focus on various aspects of laboratory instruction (86.37 %), reflect on problems in a laboratory (86.36 %), enhance self-evaluation ability (95.48 %), and promote learning efficiency (77.27 %). Moreover, they all expressed that it could guide them to identify problems in experiment operation and reflect on teaching, and they were willing to use it in the future.

Table 7
Users' Experience of the Video Assessment Questionnaire

Category	Strongly agree *n*(%)	Agree *n*(%)	Uncertain *n*(%)	Disagree *n*(%)	Strongly disagree *n*(%)
Easy to use.	12(54.55)	4(18.18)	4(18.18)	2(9.09)	0(0)
Helps to identify problems in experiment operations.	15(68.18)	7(31.82)	0(0)	0(0)	0(0)
Improves experiment operation skills.	11(50.00)	9(40.91)	2(9.09)	0(0)	0(0)
Helps to pay attention to my performance in all aspects of the experiment.	12(54.55)	7(31.82)	2(9.09)	1(4.55)	0(0)
Helps to reflect on problems in laboratory instruction in a timely manner.	8(36.36)	11(50.00)	3(13.64)	0(0)	0(0)
Improves self-evaluation ability.	14(63.64)	7(31.82)	1(4.55)	0(0)	0(0)
Improves teaching reflection ability.	17(77.27)	5(22.73)	0(0)	0(0)	0(0)
Improves innovation in experiments.	4(18.18)	9(40.91)	8(36.36)	1(4.55)	0(0)
Improves learning efficiency.	6(27.27)	11(50.00)	5(22.73)	0(0)	0(0)
Intention to use the scale in future learning activities.	17(77.27)	5(22.73)	0(0)	0(0)	0(0)

In the interviews, some students with opposing views believed that the use of the Video Assessment Questionnaire took up part of the learning time and did not improve the learning efficiency. Moreover, the indicators in the Video Assessment Questionnaire were numerous and could be streamlined. Some students also suggested that the use of the questionnaire can improve laboratory instruction, but not enhance the ability to innovate.

Discussion

Stage 1 focused on the first research question to develop the scale using the Delphi panels. The scale includes three dimensions: content organization, experiment operation, and video quality. Although previous studies have created video evaluation rubrics, they were mostly for online learning and research purposes. Survey questions may be given to students to focus on certain content while watching videos prior to laboratories. This type of question or scale is normally used as guidance for students on what to learn from a video, and is topic-specific. For example, questions pertaining to safety issues go with a safety video, whereas questions related to the number of glassware and apparatus to set up a microscale distillation experiment match a video on distillation (Box et al., 2017). Rubrics are also applied for analyzing videos for qualitative or mixed method studies. They vary according to the purpose of the study. The scale presented in the current study is uniquely designed for pre- and in-service chemistry teachers to prepare, evaluate, and choose quality videos.

Regarding the second research question, the study results demonstrated that using the scale could improve students' experiment operation skills. However, this improvement required a certain amount of practice. We found that when the scale was first used, students were unfamiliar with the evaluation criteria and the video production. Although we had explained the scale in detail in the class, there were still some problems during its use. For example, during the experiment, students had experiment operation errors, but they were unable to find these errors in time when watching the videos and using the scale for evaluation. The 20 indicators may have been excessive and may thus have created too much cognitive load for the future teachers. One solution is to increase the amount of practice to allow them to familiarize themselves with the indicators. Another solution is to break the chemistry experiments down into multiple simple experiments. When there is less content or fewer experiment operations, the cognitive load due to evaluation may be reduced. Finally, future studies may investigate students' cognitive load while using the scale to refine the indicators or arrange teaching appropriately for students.

The study found that the use of the scale had little effect on the students' content organization. The students lacked teaching experience, and their content organization basically followed the textbook. As a result, the experimental and control groups organized the content similarly. Moreover, the last experiment required students to design their own experiments. However, the students' innovation ability was weak. Most of them adopted experiments from the textbook. Overall, the content organization of the students in both groups was very similar. Instruction other than video making and self-assessment using the scale is needed to specifically enhance their content organization.

Furthermore, the difference in the quality of the videos made by the two groups was not significant. The videos were recorded by the students in pairs. They used smart phones to record each other's experiment teaching. The quality was therefore limited to the smart phone's video function and whatever software was available to them. Before the class, the teacher had explained the video recording specifications and precautions in detail and provided 2-3 teaching videos for the students to observe. Although the teacher emphasized that the students should use their creativity to record their videos, in the teaching process, the students' attention was mostly focused on the experiment operations and explanations. The video quality was limited to ensuring clear and stable pictures or adding opening titles and subtitles.

Concerning the third research question, the students held positive perceptions of the scale. The use of the scale urged them to reflect in time. Timely feedback can support students' learning (Van den Hurk et al., 2016). Finally, assessment of videos can support teachers to notice the teaching of chemistry experiments from multiple angles and identify problems and solutions, which can help them to conduct more in-depth analyses of the teaching practices (Christ et al., 2017).

Limitations

This study used the scale for self-assessment, and as such, the results are limited in the following aspects. Firstly, the validity of self-assessment is subject to students' personal knowledge and experience. Secondly, the research results only represent the video assessment learning results of one school, not other schools. More longitudinal studies with a larger sample size are needed to track how students' experiment teaching abilities change over time. Thirdly, the Delphi method depends on the expert panel. Although we tried to involve experts from chemistry, technology, and education with a variety of teaching experience, they all embraced the same culture. As a result, they emphasized an indicator, ideology and value in the content organization. The content was required not only to be scientific, but also to be morally, politically, and socially correct. This has been a common indicator for all publications and films created in the culture. Taking into account its relevance to chemistry education, it was eventually deleted. Nevertheless, it points out a potential problem with the use of the Delphi method.

Conclusions

The multiple rounds of Delphi created indicators important for assessing chemistry videos. The follow-up empirical study verified the effectiveness of the Video Assessment Questionnaire in terms of helping pre-service chemistry teachers to improve the chemistry experiment operation skills. The survey and interviews revealed that the Video Assessment Questionnaire was easy to use, and students were willing to apply it to evaluate experiment teaching.

However, for relatively simple experiments, that is, experiments with few steps, and innovative experiments, the Video Assessment Questionnaire was not considerably useful. Simple experiments are easy to master. Students could conduct them without extra help. Innovative experiments required creativity, which is beyond the scope of the questionnaire. Nevertheless, for most of the experiments, the questionnaire can help students reduce experiment operation errors and improve experiment skills.

Moreover, the questionnaire did not promote content organization, nor video quality. It could be attributed to the fact that chemistry majors possessed adequate content knowledge and were less likely to make mistakes in fundamental experiments. In the future, these indicators may be tested with non-chemistry majors or using experiments with different levels of complexity to confirm their effectiveness.

Furthermore, previous research found that pre-service teachers benefited by viewing model videos when preparing their own teaching (Beacher et al., 2013). Future studies may provide experiment videos to students to help them think about how to record good videos. Furthermore, the quality of a teacher's contributions to a group's discussion is important (Vrikki et al., 2017). Future research may analyze teachers' participation in student video evaluation activities and how it interacts with the use of the scale. Finally, future research may use the Video Assessment Questionnaire to guide the video making of chemical experiments for MOOCs, for example, or to evaluate appropriate experiment videos for students. Pre-service and in-service chemistry teachers can use it to improve their teaching and in preparing materials for students.

Acknowledgments

This work was supported by the 13th Five-Year Plan of Educational Science in Jiangsu Province of China (YZ-c/2016/18), FinCEAL Plus (51770), and the Academy of Finland (318380).

References

Arya, P., Christ, T., & Chiu, M. M. (2015). Links between characteristics of collaborative peer video analysis events and literacy teachers' outcomes. *Journal of Technology and Teacher Education*, *23*(2), 159-183.

Baecher, L., Kung, S. C., Jewkes, A. M., & Rosalia, C. (2013). The role of video for self-evaluation in early field experiences. *Teaching* & *Teacher Education*, 36, 189-197. https://doi.org/10.1016/j.tate.2013.08.001

Balasubramanian, R., & Agarwal, D. (2012). Delphi technique - A review. *International Journal of Public Health Dentistry*, *3*(2), 16-25.

Bencze, L., Hewitt, J., & Pedretti, E. (2001). Multi-media case methods in pre-service science education: Enabling an apprenticeship for praxis. *Research in Science Education*, *31*(2), 191-209. https://doi.org/10.1023/A:1013121930945

Blonder, R., Jonatan, M., Bar-Dov, Z., Benny, N., Rap, S., & Sakhnini, S. (2013). Can You Tube it? Providing chemistry teachers with technological tools and enhancing their self-efficacy beliefs. *Chemistry Education Research and Practice*, *14*(3), 269-285. https://doi.org/10.1039/C3RP00001J

Box, M. C., Dunnagan, C. L., Hirsh, L. A., Cherry, C. R., Christianson, K. A., Gibson, R. J., Wolfe, M. I., & Gallardo-Williams, M. T. (2017). Qualitative and quantitative evaluation of three types of student-generated videos as instructional support in organic chemistry laboratories. *Journal of Chemical Education*, *94*(2), 164-170. https://doi.org/10.1021/acs.jchemed.6b00451

Chen, Y. J., Guo, Y. H., & Zhou, X. Q. (2015). *Modern educational technology*. China: Higher Education Press.

Chen, Y. K., & Tian, A. L. (2014). *An introduction to MOOC and Flip Classroom*. East China Normal University Press.

Christ, T., Arya, P., & Chiu, M. M. (2017). Video use in teacher education: An international survey of practices. *Teaching and Teacher Education*, *63*, 22-35. https://doi.org/10.1016/j.tate.2016.12.005

Fadde, P., & Sullivan, P. (2013). Using interactive video to develop preservice teachers' classroom awareness. *Contemporary Issues in Technology and Teacher Education*, *13*(2), 156-174.

Hasson, F., Keeney, S., & McKenna, H. (2000). Research guidelines for the Delphi survey technique. *Journal of Advanced Nursing, 32*(4), 1008-1015. https://doi.org/10.1046/j.1365-2648.2000.t01-1-01567.x

Heidi, A., & Anna, V. (2009). Promoting learning and achievement through self-evaluation. *Theory into Practice*, *48*(1), 12-19. https://doi.org/10.1080/00405840802577544

Hong, C. (1994). Theoretical discussion on the training of chemical experiment operation skills. *Chinese Journal of Chemical Education, 15*(9), 4-8.

Hu, J., Fallacaro, M. D., Jiang, L., Wu, J., Jiang, H., & Shi, Z., Ruan, H. (2017). IFNA approved Chinese Anaesthesia Nurse Education Program: A Delphi method. *Nurse Education Today, 56*, 6-12. https://doi.org/10.1016/j.nedt.2017.05.017

Jordan, J. T., Box, M. C., Eguren, K. E., Parker, T. A., Saraldi-Gallardo, V. M., Wolfe, M. I., & Gallardo-Williams, M. T. (2015). Effectiveness of student-generated video as a teaching tool for an instrumental technique in the organic chemistry laboratory. *Journal of Chemical Education, 93*(1), 141-145. https://doi.org/10.1021/acs.jchemed.5b00354

Kerr, N. L., & Bruun, S. E. (1983). Dispensability of member effort and group motivation losses: Free-rider effects. *Journal of Personality and Social Psychology, 44*(1), 78-94. https://doi.org/10.1037/0022-3514.44.1.78

Korkmaz, F. D., Ozsaker, E., Tasdemir, N., & Karacabay, K. (2016). An examination of the effect two different evaluation methods used in CPR training have on the learning process. *Croatian Journal of Education - Hrvatski Časopis za Odgoj i Obrazovanje, 18*(1), 157-175. https://doi.org/10.15516/cje.v18i1.1323

Kraft, A., Rankin, E. S., & Arrighi, V. (2012). Using short videos to supplement lectures on reaction mechanisms, organic spectroscopy, and polymer chemistry. *Advances in Teaching Organic Chemistry*, 209-224. https://doi.org/10.1021/bk-2012-1108.ch013

LeFebvre, L., LeFebvre, L., Blackburn, K., & Boyd, R. (2015). Student estimates of public speaking competency: The meaning extraction helper and video self-evaluation. *Communication Education*, *64*(3), 261-279. https://doi.org/10.1080/03634523.2015.1014384

Li, J. J., XU, Z., Bai, Y. H., & Yang, Y. (2017). Application of microvideo in chemistry experimental practice. *Research & Exploration in Laboratory, 36*(1), 189-191.

Mahoney, P., Macfarlane, S., & Ajjawi, R. (2019). A qualitative synthesis of video feedback in higher education. *Teaching in Higher Education*, *24*(2), 157-179. https://doi.org/10.1080/13562517.2018.1471457

Malekian, F., & Aliabadi, F. M. (2012). Review of methods of organizing the content of the curriculum in the educational system, based on ICT (Information and Communication Technology) from the experts' view. *Procedia - Social and Behavioral Sciences, 51*, 19-23. https://doi.org/10.1016/j.sbspro.2012.08.112

Marsh, B., Mitchell, N., & Adamczyk, P. (2010). Interactive video technology: Enhancing professional learning in initial teacher education. *Computers & Education*, *54*(3), 742-748. https://doi.org/10.1016/j.compedu.2009.09.011

Masats, D., & Dooly, M. (2011). Rethinking the use of video in teacher education: A holistic approach. *Teaching and Teacher Education*, *27*(7), 1151-1162. https://doi.org/10.1016/j.tate.2011.04.004

McDowell, L. (1995). The impact of innovative assessment on student learning. *Innovations in Education and Training International, 32*(4), 302-313. https://doi.org/10.1080/1355800950320402

Nagy, J. T. (2018). Evaluation of online video usage and learning satisfaction: An extension of the technology acceptance model. *International Review of Research in Open & Distance Learning*, *19*(1), 160-185. https://doi.org/10.19173/irrodl.v19i1.2886

Ogden, S. R., Culp Jr, W. C., Villamaria, F. J., & Ball, T. R. (2016). Developing a checklist: Consensus via a modified Delphi technique. *Journal of Cardiothoracic and Vascular Anesthesia, 30*(4), 855-858. https://doi.org/10.1053/j.jvca.2016.02.022

Olina, Z., & Sullivan, H. J. (2004). Student self-evaluation, teacher evaluation, and learner performance. *Educational Technology Research & Development*, *52*(3), 5-22. https://doi.org/10.1007/BF02504672

Paris, S. G., & Paris, A. H. (2001). Classroom applications of research on self-regulated learning. *Educational Psychologist, 36*(2), 89-101. https://doi.org/10.1207/S15326985EP3602_4

Raaijmakers, S. F., Baars, M., Paas, F., van Merriënboer, J. J., & van Gog, T. (2019). Effects of self-assessment feedback on self-assessment and task-selection accuracy. *Metacognition and Learning*, *14*(1), 21-42. https://doi.org/10.1007/s11409-019-09189-5

Richards-Babb, M., Curtis, R., Smith, V. J., & Xu, M. (2014). Problem solving videos for general chemistry review: Students' perceptions and use patterns. *Journal of Chemical Education, 91*(11), 1796-1803. https://doi.org/10.1021/ed500280b

Seidel, T., Blomberg, G., & Renkl, A. (2013). Instructional strategies for using video in teacher education. *Teaching and Teacher Education, 34*, 56-65. https://doi.org/10.1016/j.tate.2013.03.004

Seshadrinathan, K. Soundararajan, R. Bovik, A. C., & Cormack, L. K. (2010). Study of subjective and objective quality assessment of video. *IEEE Transactions on Image Processing, 19*(6), 1427-1441. https://doi.org/10.1109/TIP.2010.2042111

Stoll, L., Bolam, R., McMahon, A., Wallace, M., & Thomas, S. (2006). Professional learning communities: A review of the literature. *Journal of Educational Change*, 7, 221-258. https://doi.org/10.1007/s10833-006-0001-8

Tano, R., Takaku, S., & Ozaki, T. (2017). The efficacy of video monitoring-supported student self-evaluation of dental explorer skills in dental hygiene education. *International Journal of Dental Hygiene, 15*(4), 113-118. https://doi.org/10.1111/idh.12243

Tierney, J., Bodek, M., Fredricks, S., Dudkin, E., & Kistler, K. (2014). Using web-based video as an assessment tool for student performance in organic chemistry. *Journal of Chemical Education, 91*(7), 982-986. https://doi.org/10.1021/ed400195c

Tse, W. S., Choi, L. Y., & Tang, W. S. (2019). Effects of video-based flipped class instruction on subject reading motivation. *British Journal of Educational Technology*, *50* (1), 385-398. https://doi.org/10.1111/bjet.12569

Van den Hurk, H. T. G., Houtveen, A. A. M., & Van de Grift, W. J. C. M. (2016). Fostering effective teaching behavior through the use of data-feedback. *Teaching & Teacher Education*, *60*, 444-451. https://doi.org/10.1016/j.tate.2016.07.003

Yan, Z., Brown, G. T., Lee, J. C. K., & Qiu, X. L. (2019). Student self-assessment: Why do they do it? *Educational Psychology*, 1-24. https://doi.org/10.1080/01443410.2019.1672038

Zhan, R., & Zhang, X. B. (2010). The study on evaluation system of instructional video in the extractive course. *Modern Educational Technology, 20*(6), 35-39.

Zimmerman, B. J. (2000). Self-efficacy: An essential motive to learn. *Contemporary Educational Psychology*, *25*(1), 82-91. https://doi.org/10.1006/ceps.1999.1016

Cui Liu
College of Teacher Education
Jiangsu Normal University
101 Shanghai Rd., Xuzhou, Jiangsu Province, China
6020040038@jsnu.edu.cn

Sufen Chen
Graduate Institute of Digital Learning and Education, National Taiwan University of Science and Technology
43, Sec.4, Keelung Rd., Taipei, Taiwan 106, R.O.C.
sufchen@mail.ntust.edu.tw

Shuo Liu
School of Chemistry & Materials Science
Jiangsu Normal University
101 Shanghai Rd., Xuzhou, Jiangsu Province, China
372950082@qq.com

Jingyuan Liu
School of Chemistry & Materials Science
Jiangsu Normal University
101 Shanghai Rd., Xuzhou, Jiangsu Province, China
1924615447@qq.com

Jie Wang
School of Chemistry & Materials Science
Jiangsu Normal University
101 Shanghai Rd., Xuzhou, Jiangsu Province, China
1905464998@qq.com

Yingying Han
School of Chemistry & Materials Science
Jiangsu Normal University
101 Shanghai Rd., Xuzhou, Jiangsu Province, China
2893576648@qq.com

Delfi-metoda istraživanja i verifikacija procjene videozapisa o kemijskim eksperimentima

Sažetak

Cilj ovoga istraživanja bio je primijeniti Delfi-metodu za dobivanje i kombiniranje podudarnih mišljenja, kako bi se izradili pokazatelji pomoću kojih budući nastavnici kemije mogu procijeniti videozapise o kemijskim eksperimentima. Ovo istraživanje provedeno je u dvije faze. U prvoj je izrađen Upitnik za procjenu videozapisa o kemijskim eksperimentima pomoću modificirane Delfi metode. Na taj su način utvrđena podudarna mišljenja i provedena tri kruga anketiranja uživo ili u online obliku. Uzorak se sastojao od 50 ispitanika – profesora, docenata, studenata dodiplomskih studija te srednjoškolskih nastavnika. Upitnik za procjenu videozapisa pokazao je visoku razinu konsenzusa, što potvrđuju koeficijent varijacije i Kendallov W koeficijent podudarnosti u drugom i trećem krugu anketiranja. Finalna verzija obuhvaća 20 pokazatelja za tri glavna aspekta: organizaciju sadržaja, izvođenje eksperimenta i kvalitetu videozapisa. U drugoj fazi provedenen je kvazieksperiment kako bi se testirao učinak primjene Upitnika u kolegiju Metodika nastave kemije. Eksperimentalna skupina ($n = 22$) koristila je upitnik kako bi procijenila vlastite videozapise i provela refleksiju o izvedbi eksperimenta, dok je kontrolna skupina ($n = 11$) odradila iste zadatke, no bez upitnika. Utvrđeno je da su budući nastavnici koji su koristili Upitnik za procjenu videozapisa o kemijskim eksperimentima izvodili bolje nastavu u laboratoriju ($t(31) = 4{,}25$, $p < .001$) i kemijske eksperimente ($t(31) = 4{,}95$, $p < .001$) od onih koji su samo izrađivali videozapise. Istraživanje je potvrdilo učinkovitost Upitnika za procjenu videozapisa o kemijskim eksperimentima kao pomoć budućim nastavnicima Kemije u poboljšanju vještina izvođenja kemijskih eksperimenata.

Ključne riječi: *procjena; istraživanje u nastavi kemije; nastava u laboratoriju; učenje temeljeno na multimediji*

Uvod

Svrha ovoga istraživanja bila je izraditi pokazatelje pomoću kojih će nastavnici i budući nastavnici Kemije vrednovati svoje videozapise o kemijskim eksperimentima, kako bi unaprijedili vlastito poučavanje o kemijskim eksperimentima. Korištenje videozapisa u izobrazbi nastavnika može pomoći budućim nastavnicima u razmjeni iskustava, fokusiranju na određene obrasce ponašanja te prepoznavanju prednosti i nedostataka u nastavnoj praksi (Fadde i Sullivan, 2013; Marsh, Mitchell i Adamczyk, 2010; Masats i Dooly, 2011). Voditelji istraživanja koristili su videozapise kako bi premostili jaz između znanstvene teorije i prakse (npr. Bencze, Hewitt i Pedretti, 2001; Stoll, Bolam, McMahon, Wallace i Thomas, 2006; Seidel, Blomberg i Renkl, 2013). Da budemo precizniji, videozapisi mogu pobliže pokazati implikacije određenoga znanstvenog pojma i različite vještine provođenja eksperimenata. Štoviše, videozapise može se integrirati u različite pedagoške metode. Na primjer, videozapisi u kombinaciji sa samorefleksijom mogu pomoću nastavnicima i budućim nastavnicima prepoznati problem (Christ, Arya i Ming, 2017), dok refleksija o videozapisima s kolegama razvija vještine rješavanja problema (Arya, Christ i Chiu, 2015). Vrednovanje videozapisa može poslužiti kao povratna informacija ili kao informacija o onome što bi trebalo poboljšati u sljedećim koracima. Ono je ključan element formativnoga vrednovanja jer potiče buduće nastavnike na promišljanje o kvaliteti vlastita rada, umjesto da se oslanjaju na nastavnike kao na jedine izvore vrednovanja (Heidi i Anna, 2009). Međutim, do sada nije provedena sustavna analiza kvalitete videozapisa o kemijskim eksperimentima zbog čega budući nastavnici ne poznaju dovoljno znanstvene metode pomoću kojih bi vrednovali vlastite videozapise i videozapise svojih kolega. U ovome smo istraživanju koristili Delfi-metodu za kombiniranje podudarnih mišljenja kako bi izradili pokazatelje za vrednovanje videozapisa o kemijskim eksperimentima.

Pregled literature

Gledanje videozapisa i učenje

Upotreba videozapisa ima značajan pozitivan učinak na učenje i zadovoljstvo procesom učenja (Nagy, 2018), kao i na učinkovitost poučavanja i zadovoljstvo kolegijem (Tse, Choi i Tang). Videozapisi su jako korisni u poučavanju i učenju o praktičnom radu (Tano, Takaku i Ozaki, 2017). Mogu se lako integrirati u nastavu u različitim fazama. Na primjer, nastavnici ih mogu koristiti kao primjere pomoću kojih (1) ilustriraju određena pravila, pojmove i principe ili (2) demonstriraju fenomene pomoću kojih se kasnije izvode pravila. Seidel, Blomberg i Renkl (2013) utvrdili su da, u kontekstu poučavanja budućih nastavnika o pedagoškim principima, ilustracija pravila olakšava učenje o činjenicama i evaluaciju nastave, dok demonstracija fenomena vodi k boljem planiranju nastave. Videozapisi imaju raznolike uloge i učinke.

Videozapisi o kemijskim eksperimentima uvelike se koriste u nastavi Kemije (Tierney, Bodek, Fredricks, Dudkin i Kistler, 2014), a u posljednje se vrijeme sve više dijele na YouTubeu (Blonder, Jonatan, Bar-Dov, Benny, Rap i Sakhnini, 2013). Lako ih je koristiti. Mogu se zaustaviti, ubrzati, usporiti i odgledati nekoliko puta, tako da studenti mogu učiti umjerenim tempom i vratiti se na materijale kada god im to zatreba (Kraft, Rankin i Arrighi, 2012; Richards-Babb, Curtis, Smith i Xu, 2014). Videozapisi mogu pokazati vrlo jasne slike i fokusirati se na fenomene i vještine te su ponekad učinkovitiji od nastave koju izvode nastavnici. Jordan i sur. (2015) pokazali su da su studenti koji su odgledali kratke i sažete videoupute u laboratorijskim vježbama radili puno samostalnije i bolje od studenata kojima je nastavu izvodio nastavnik.

Nadalje, videozapisi mogu dati bogate vizualne i dinamične informacije (Fadde i Sullivan, 2013; Marsh, Mitchell i Adamczyk, 2010). U njima učenici mogu jasno promatrati govorni izričaj, izvođenje dinamičnoga eksperimenta te sadržaj predstavljen na raznovrsne načine, poput grafova, animacija, teksta, usporenoga snimka itd. Videozapisi o kemijskim eksperimentima olakšavaju poučavanje i reduciraju probleme povezane s kemijskim otpadom, sigurnošću te objašnjavaju nejasnoće.

Iako se videozapisi o kemijskim eksperimentima koriste dugi niz godina i pozitivno utječu na učenje (Tierney, Bodek, Fredricks, Dudkin i Kistler, 2014), stručnjaci su rijetko predlagali Skalu za vrednovanje videozapisa o kemijskim eksperimentima i rijetko su empirijski ispitivali njihovu dobrobit za buduće nastavnike (Christ, Arya i Chiu, 2017). Učinkovita skala za vrednovanje reflektira nastavne standarde i očekivanja te promiče samovrednovanje kod budućih nastavnika (Baecher, Kung, Jewkes i Rosalia, 2013).

Samovrednovanje

Kao jedna od metoda vrednovanja, samovrednovanje je dobar način da se budućim nastavnicima pokaže kako mogu provesti refleksiju o provedenim aktivnostima i učiti iz vlastita iskustva. U mnogim se istraživanjima proučavao utjecaj samovrednovanja (Kerr i Bruun, 1983; McDowell, 1995). Ono studentima pruža priliku da procijene vlastita postignuća i podignu razinu samoučinkovitosti i samoreguliranoga učenja (Raaijmakers, Baars, Paas, van Merriënboer i van Gog, 2019; Yan, Brown, Lee i Qiu, 2019; Zimmerman, 2000). Na primjer, studenti koji su proveli samovrednovanje imali su pozitivniji stav prema istraživanju i veće samopouzdanje u osmišljavanju eksperimenata u usporedbi s onima koji su primili povratnu informaciju od nastavnika ili evaluaciju uopće nisu dobili (Olina i Sullivan, 2004). Nadalje, Paris i Paris (2001) smatraju da samovrednovanje često podrazumijeva i internalizaciju i procesiranje raznovrsnih standarda, što može studentima pomoći pri reguliranju učenja na učinkovitiji, smisleniji i točniji način. Različiti oblici samovrednovanja imaju različite funkcije. Prediktivno samovrednovanje odnosi se na sposobnost predviđanja uspjeha u realizaciji budućih zadataka. Sumativno samovrednovanje koristi se za određivanje uspjeha u radu nakon provedenih aktivnosti učenja. Trenutačno samovrednovanje je sposobnost prepoznavanja razlika između stvarne i željene uspješnosti u radu te potreba u učenju tijekom aktivnosti. Samovrednovanje pomoću videozapisa može biti sumativno za cjelokupni uspjeh u radu i trenutačno za stalnu regulaciju učenja.

Baecher i sur. (2013) usporedili su načine na koje su budući nastavnici Kemije sami procijenili svoj proces poučavanja uz primjenu videozapisa i bez njih. Iako su obje skupine imale iste rubrike za evaluaciju, videozapisi su budućim nastavnicima pomogli bolje shvatiti što je to kvalitetno poučavanje u području kemije te su im pomogli realnije procijeniti vlastito poučavanje. Isto tako, Tano, Takaku i Ozaki (2017) dopustili su studentima da pogledaju snimke vlastitih vještina izvođenja dentalnih pregleda. Promatranje videozapisa uvelike je unaprijedilo praktične vještine studenata te je dovelo do ujednačenih rezultata samoevaluacije i evaluacije koju su proveli nastavnici. Kako su naveli LeFebvre, LeFebvre, Blackburn i Boyd (2015), u kolegijima u kojima se raspravlja o samoprocjeni rada uz videozapise, studenti obično precjenjuju vlastiti rad. Za razliku od toga, rubrike, videozapisi i samoprocjena promiču samoregulacijsko učenje i pomažu studentima objektivnije samoprocijeniti vlastiti rad.

Nadalje, Korkmaz, Ozsaker, Tasdemir i Karacabay (2016) zamolili su studente da procijene svoj rad pomoću snimanja videozapisa. Došli su do saznanja da su studenti koji su ih koristili u većoj mjeri tijekom procesa samoprocjene uspijevali uočiti nedovoljno dobre ili netočne postupke. Istraživači smatraju da su videozapisi intuitivni i da ih je lako analizirati i zapamtiti. Povratna informacija koju pruža videozapis dobra je alternativa tradicionalnoj pisanoj povratnoj informaciji (Mahoney, Macfarlane i Ajjawi, 2019). Tijekom samoprocjene budući nastavnici imaju dvije uloge – ulogu nastavnika i ulogu ocjenjivača. Nadići samosvijest kako bi na kritički način promatrali svoj rad je veliki izazov, osobito kada samoopažanje uključuje i videozapise u kojima oni izvode eksperimente u nastavi (Fadde i Sullivan, 2013). Često im je teško nadići fazu promatranja nastavnoga procesa u videozapisu i analizirati vlastiti govor, izvedbu eksperimenata i objašnjavanje kemijskih principa. Očekuje se da će kroz proces samoprocjene budući nastavnici bolje razumjeti principe poučavanja učenika o eksperimentima, organizaciju takvog poučavanja, inovacije u njemu i sam nastavni sadržaj te da će provesti konstruktivnu evaluacije vlastita rada. Primjena Upitnika za samoprocjenu može im pomoći ostvariti te ciljeve i pomoći da detaljno procijene svoje videozapise.

Istraživačka pitanja

U ovom istraživanju korištena je Delfi-metoda za izradu Skale za procjenu videozapisa pomoću koje su budući nastavnici kemije procjenjivali videozapise za nastavu Kemije, ispitivali upotrebu skale te tražili odgovore na sljedeća pitanja:

1. Koji su pokazatelji važni za procjenu videozapisa o kemijskim eksperimentima?
2. Promiče li Skala za procjenu videozapisa izradu videozapisa o kemijskim eksperimentima?
3. Kakve su percepcije budućih nastavnika o Skali za procjenu videozapisa?

Dizajn istraživanja

Upotreba Delfi-metode u izradi nastavnoga alata odavno je prihvaćena (Ogden, Culp, Villamaria i Ball, 2016). Međutim, metoda je naišla i na kritike zbog toga što joj nedostaje metodološka strogoća jer koristi vrlo fleksibilne tehnike. Ovo istraživanje odvijalo se u dvije faze, kako bi se ostvarila što veća razina pouzdanosti. U prvoj se fazi izradio Upitnik za procjenu videozapisa. Delfi-tehnika sastojala se od tri uzastopna kruga anketiranja kako bi se prikupila mišljenja stručnjaka i dobila maksimalno podudarna mišljenja. U drugoj su fazi budući nastavnici koristili upitnik za procjenu vlastitih videozapisa i za refleksiju o svojemu radu. Iako za Delfi ne postoje jedinstvene smjernice, preporuča se provedba ključnih koraka i vođenje jasnih bilješki o njima (Hasson, Keeney i McKenna, 2000). Slika 1 pokazuje ključne korake i rezultate svakoga kruga.

Slika 1.

1. faza: Izrada Upitnika za procjenu videozapisa

Prvi krug Delfi-istraživanja. Delfijev panel u prvom krugu sastojao se od deset vodećih stručnjaka u području obrazovanja iz četiri provincije u Kini. Sačinjavali su ga četiri kemičara, dva obrazovna tehnologa i četiri nastavnika Kemije. Svi su stručnjaci imali više od pet godina iskustva u istraživanjima u području kemije ili nastavi Kemije te više od jedne godine iskustva u izradi videozapisa i *online* tečajeva i kolegija. Pozvani su da pojedinačno navedu važne pokazatelje u videozapisima o kemijskim eksperimentima. Panelu smo stavili na raspolaganje knjige (Chen, Guo i Zhou, 2015; Chen i Tian, 2014), znanstvene radove (Li, Xu, Bai i Yang, 2017; Seidel, Blomberg i Renkl, 2013; Seshadrinathan, Soundararajan, Bovik i Cormack, 2010) i pet videozapisa o kemijskim eksperimentima. Dali su svoja mišljenja i objasnili ih. Zatim su raspravljali o različitim mišljenjima i složili se oko tri glavne domene pokazatelja: organizaciji sadržaja, izvođenju eksperimenata i kvaliteti videozapisa, s ukupno 26 tvrdnji.

Drugi i treći krug Delfi-istraživanja. U drugome i trećem krugu, sveučilišni profesori kemije, studenti na dodiplomskim studijima kemije i nastavnici Kemije zaposleni na srednjim školama procijenili su prikladnost i važnost pokazatelja. Svi su imali više od jedne godine iskustva u izradi videozapisa. U svakome od ova dva kruga sudjelovalo je 20 različitih stručnjaka.

Drugi krug proveden je uz primjenu proračunske tablice. Stručnjaci su trebali ocijeniti važnost svakoga pokazatelja. Njihove ocjene koristile su se pri odluci o tome treba li se indeks ostaviti ili isključiti. U istraživanju je korištena i skala od pet stupnjeva, u kojoj su pokazatelji prikazani u nizu od najvažnijega do najmanje važnoga. Srednja vrijednost, postotak podudarnosti i standardna devijacija koristili su se za procjenu kolektivne prosudbe stručnjaka. Za prikazivanje varijabilnosti mišljenja stručnjaka korišteni su koeficijent varijacije (CV) i Kendallov koeficijent podudarnosti (Kendallov W). U Kendallovu koeficijentu podudarnosti, srednja vrijednost od 4 (koja označava relativnu važnost), slaganje od 80 %, standardna devijacija od 1,2, koeficijent varijacije od 0,2 i $p < ,05$ koristili su se kao kriteriji za postizanje konsenzusa (Balasubramanian i Agarwal, 2012; Hu i sur., 2017). Pokazatelji koji su ispunjavali ove kriterije bili su zadržani. Treći krug proveden je putem *online* platforme, a korišteni su isti kriteriji za donošenje konačne odluke o pokazateljima.

Demografske karakteristike stručnjaka koji su sudjelovali u svakome krugu prikazane su u Tablici 1.

Tablica 1.

Upitnik za procjenu videozapisa. Pokazatelji su svrstani u tri kategorije. Prva kategorija obuhvaća organizaciju sadržaja, procjenjuje opću organizaciju i specifični sadržaj videozapisa (Malekian i Aliabadi, 2012), uključujući racionalnost, atraktivnost, smislenu strukturu, detaljan sadržaj, smjernice, inovacije, učinkovito upravljanje vremenom, izvedbu, nošenje zaštitne odjeće, usmeno izražavanje te kemijske supstance. Tri indikatora, interakcija, reference i refleksija, eliminirani su u drugom krugu.

Što se tiče druge kategorije, izvođenje kemijskoga eksperimenta kompleksan je sustav mnogih podfaktora sastavljen prema određenoj logici i hijerarhijskim odnosima. Podfaktori u tom sustavu međusobno su povezani i ovise jedan o drugome, a uključuju: instrument za provođenje eksperimenta, kemijske supstance te izvođenje eksperimenta (Hong, 1994). Stoga ova kategorija obuhvaća standardno izvođenje eksperimenta, demonstraciju, jasne korake koji logički slijede jedan drugoga, fenomene te okruženje u kojemu se izvodi eksperiment i znanstveni postupak. Svi pokazatelji prošli su kroz tri kruga Delfi-istraživanja.

Treća kategorija, kvaliteta videozapisa, procjenjuje se na temelju kvalitete slike, sadržaja slike, tehnologije korištenje za uređivanje videozapisa, efekata na slikama, zvučnih efekata i sinkronizacije (Zhan, 2010). Pokazatelji uključuju kvalitetu ekrana, tehnologiju korištenu za uređivanje videozapisa te jasnoću. Tri pokazatelja - videoisječak, naslovi te specijalni audio i videoefekti, eliminirani su u prvome krugu.

Upitnik je također sadržavao i pitanja otvorenoga tipa kako bi se od stručnjaka prikupila i ostala mišljenja. Tvrdnje su prikazane u Tablici 2. U Upitniku za procjenu videozapisa maksimalan broj bodova je 100 (5 bodova po pokazatelju), a od toga Organizacija sadržaja ima udio od 50 %, Izvođenje kemijskoga eksperimenta 35 %, a Kvaliteta videozapisa 15 %.

Tablica 2.

2. faza: Primjena Upitnika za procjenu videozapisa u kolegiju Metodika nastave kemije za buduće nastavnike

Uzorak ispitanika u istraživanju sastojao se od 33 studenta treće godine na Sveučilištu Jiangsu Normal, koje je glavni institut za izobrazbu nastavnika. Smješteno je u jednoj od najvećih kineskih provincija. Svi su studenti pohađali kolegij Metodika nastave kemije. Prosječna dob ispitanika bila je 19,4 godine. Većina ispitanika bila je ženskoga spola (27,82 %). Bili su obaviješteni o načinu provedbe istraživanja te su dali svoj pisani pristanak.

Intervencija je trajala deset tjedana, a studenti su svakoga tjedna, počevši u drugome, završili po jedan eksperiment. Predavanje je trajalo 3 sata, a održavalo se tjedno. Svakoga tjedna sudionici su proveli po jedan kemijski eksperiment (teme su prikazane u Tablici 3), o njemu izradili videozapis za primjenu u nastavi te predali izvješće o eksperimentu koje je uključivalo samoprocjenu njihova rada. Prvih 8 videozapisa prikazivalo je osnovne kemijske eksperimente koje su studenti izvodili prema laboratorijskim uputama u nastavnim materijalima propisanima za kolegij. Trebali su prvo pročitati informacije o nadolazećem eksperimentu i na temelju toga izraditi nastavni plan za demonstraciju eksperimenta u videozapisu. Na nastavi su im objašnjeni postupci za provedbu eksperimenta, a zatim su dovršili eksperiment i izradili plan nastave. Nakon toga su za potrebe nastave snimili izvedbu eksperimenta koristeći pametne telefone te uredili video u trajanju 10-20 minuta. Videozapis i izvješće o provedenome eksperimentu predali su nastavniku u roku od jednoga tjedna.

Posljednji videozapis bio je inovativni eksperiment koji je osmislio svaki student sam za sebe. Korišten je kako bi se odredila njihova sposobnost osmišljavanja eksperimenata i izrade videozapisa za nastavu. Studenti su dobili upute da pročitaju odgovarajuću literaturu, odrede teme i izrade plan. Nastavnik je njihov plan pregledao te nabavio potrebne kemijske supstance i opremu.

Tablica 3.

Ispitanici su bili podijeljeni u dvije skupine: kontrolnu skupinu (n = 11) i eksperimentalnu skupinu (n = 22). Kontrolna skupina samoprocjenjivala je svoj rad prikazan u videozapisima, dok je eksperimentalna skupina provela samoprocjenu pomoću Upitnika za procjenu videozapisa. U posljednjem je videozapisu svaki student morao izraditi novi kemijski eksperiment, koji je demonstrirao integriranu vještinu izrade videozapisa o kemijskim eksperimentima. Uz samoprocjenu, četvero istraživača koristilo je Upitnik za procjenu videozapisa za ocjenjivanje svih videozapisa. Ocjene istraživača odnosile su se na sposobnost studenata da izvode nastavu koja uključuje kemijske eksperimente. Pouzdanost među ocjenjivačima također se mjerila na temelju podataka četvero istraživača.

Na kraju, 22 studenta u eksperimentalnoj skupini popunili su evaluacijski upitnik kako bi se došlo do saznanja je li Upitnik za procjenu videozapisa lagano koristiti te kako bi se odredio njegov učinak na učenje. Nakon toga su uslijedili intervjui.

Rezultati

Važni pokazatelji evaluacije videozapisa o kemijskim eksperimentima i podudarnost mišljenja stručnjaka

I drugi i treći krug Delfi-metode potvrdili su važnost 20 od 26 pokazatelja. Srednje vrijednosti svih dimenzija bile su veće od 4, u rasponu od relativno važnih do najvažnijih (Tablica 4). Koeficijenti varijacije trećega kruga konzultacija bili su između 1 % i 11 %, što je manje od 0,2, a stručnjaci su postigli visoku razinu podudarnosti u mišljenjima, u rasponu od 0,675 do 1,00. Vrijednosti Kendallova W koeficijenta u drugom i trećem krugu iznosile su 359,98 i 328,98, za svaki pojedinačno, uz $p < 0,001$.

Tablica 4.

Utjecaj Upitnika za procjenu videozapisa

Na početku je četvero istraživača procijenilo četiri videozapisa kako bi uspostavili pouzdanost među ocjenjivačima. Dosegli su prihvatljivu pouzdanost s vrijednostima Kendallova W koeficijenta podudarnosti $W = 0,74, 0,83, 0,81$ i $0,93$, uz $p < 0,001$, za svaku od četiri kategorije posebno. Zatim su se podijelili i ocijenili sve videozapise – ukupno 74 ili 75 videozapisa.

Tablica 5 pokazuje ocjene dviju skupina za devet eksperimenata. Nije uočena značajna razlika u radu na početku korištenja upitnika - t(31) = -1,49 i 0,25, p > 0,05 u 1. eksperimentu i 2. eksperimentu. S povećanjem broja upotrebe upitnika, rezultati eksperimentalne skupine bili su značajno bolji od rezultata kontrolne skupine, osim u 6. i 9. eksperimentu. Šesti eksperiment demonstrirao je pripremu i svojstva kisika. Izvođenje eksperimenta bilo je relativno jednostavno, a fenomen je bio očigledan. Stoga nije bilo razlike u izvođenju eksperimenta između dvije skupine studenata. Deveti eksperiment bio je onaj koji su studenti morali sami osmisliti, čime su se uglavnom analizirale njihove sposobnosti inovativnoga osmišljavanja eksperimenata. Općenito gledajući, rezultati t-testa pokazuju da se upotrebom skale kod studenata može poboljšati sposobnost izvođenja eksperimenata u nastavi i to sa značajnom veličinom učinka (Cohenov d = 0,91 – 1,59). No, njezin utjecaj na jednostavne eksperimente, poput 6. eksperimenta ($t(31) = -0,93$, $p = 0,362$) i na inovativne eksperimente, poput 9. eksperimenta ($t(31) = 024$, $p = 0,812$) nije značajan.

Tablica 5.

Tablica 6 prikazuje sažetak devet eksperimenata prema podskalama. T-testovi pokazuju da je eksperimentalna skupina, koja je koristila Upitnik za procjenu videozapisa, imala bolje vještine za izvođenje nastave u laboratoriju od kontrolne skupine koja je samoprocjenjivala svoje videozapise bez upitnika (t(31) = -4,25, p = 0,001). Veličina učinka bila je velika (d = 1,34). Preciznije, sposobnosti izvođenja eksperimenta bile su značajno veće kada se pri samoprocjeni koristio Upitnik za procjenu videozapisa (t(31) = -4,95, p = 0,001, d = 1,96). Međutim, organizacija sadržaja (t(31) = -1,39, p = 0,174) i kvaliteta videozapisa (t(31) = -1,29, p = 0,207) nisu se značajno razlikovale u ove dvije skupine. Utjecaj korištenja upitnika uočen je uglavnom u postupcima i samom izvođenju eksperimenata.

Tablica 6.

Percepcije budućih nastavnika o korištenju Upitnika za procjenu videozapisa

Evaluacijski upitnik poslan je studentima u eksperimentalnoj skupini (n = 22), a njegovi su rezultati prikazani u Tablici 7.

Tablica 7.

Rezultati evaluacijskoga upitnika pokazali su da većina studenata (72,73 %) smatra da je Upitnik lako koristiti, dok je nekoliko studenata (27,27 %) mišljenja da je njegovo korištenje nepraktično. Većina studenata smatra da im je primjena Upitnika za procjenu videozapisa pomogla unaprijediti vještine izvođenja eksperimenata (90,91 %), fokusirati se na različite aspekte nastave u laboratoriju (86,37 %), razmisliti o problemima s kojima su se susreli u laboratoriju (86,36 %), poboljšati sposobnost samoprocjene (95,48 %) te poboljšati učinkovitost učenja (77,27 %). Štoviše, svi su izjavili da bi im upitnik mogao pomoći prepoznati probleme u izvođenju eksperimenata, provesti refleksiju o izvođenju nastave te da su je voljni nastaviti koristiti u budućnosti.

U intervjuima se pokazalo da neki studenti imaju oprečna mišljenja. Smatraju da im je primjena upitnika za procjenu videozapisa oduzela vrijeme za učenje i nisu uspjeli unaprijediti svoju uspješnost u učenju. Nadalje, smatraju da Upitnik za procjenu videozapisa sadrži previše pokazatelja i da bi mogao biti jednostavniji i učinkovitiji. Neki studenti su bili mišljenja da upitnik može poboljšati nastavu koja se provodi u laboratoriju, ali da ne može poboljšati njihovu sposobnost da inovativno osmišljavaju eksperimente.

Rasprava

U prvoj fazi istraživanja pozornost je usmjerena na prvo istraživačko pitanje – izrada skale kroz panele koje podrazumijeva Delfi-metoda. Skala uključuje tri dimenzije: organizaciju sadržaja, izvođenje kemijskih eksperimenata te kvalitetu videozapisa. Iako su u ranijim istraživanjima izrađene rubrike za procjenu videozapisa, bile su namijenjene uglavnom *online* učenju i za istraživačke svrhe. Studentima se mogu dati pitanja pomoću kojih se lakše mogu fokusirati na određeni sadržaj tijekom gledanja videozapisa, a prije rada u laboratoriju. Ovakav tip pitanja ili skale obično se koristi kao pomoć studentima u tome što trebaju iz videozapisa naučiti i usko je vezan uz temu. Na primjer, pitanja koja se odnose na sigurnosne mjere postavljaju se uz videozapise o istima, dok se pitanja koja se odnose na broj staklenog pribora i instrumente potrebne za eksperimente o destilaciji postavljaju uz videozapis o destilaciji (Box i sur., 2017). Rubrike se također koriste za analizu videozapisa u kvalitativnim istraživanjima ili istraživanjima koja koriste mješovite metode. Rubrike variraju ovisno o svrsi istraživanja. Skala prikazana u ovom istraživanju izrađena je isključivo za buduće nastavnike Kemije i one koji to već jesu, kako bi mogli pripremiti, evaluirati i odabrati kvalitetne videozapise.

Što se tiče drugoga istraživačkog pitanja, rezultati pokazuju da korištenje skale može unaprijediti vještine studenata za provođenje eksperimenata. Međutim, to zahtijeva određenu vježbu. Uočili smo da, kada su prvi put koristili skalu, studentima su kriteriji evaluacije i izrada videozapisa bili nepoznanice. Iako smo im skalu detaljno objasnili na nastavi, još uvijek su se tijekom njezine primjene javljali problemi. Na primjer, studenti su imali pogreške tijekom izvođenja eksperimenata, no nisu te pogreške mogli uočiti pravovremeno kada su gledali videozapise i koristili skalu za procjenu. Možda je broj od 20 pokazatelja bio prevelik i zbog toga je budućim nastavnicima predstavljao preveliko kognitivno opterećenje. Jedno moguće rješenje je povećati broj vježbi kako bi se bolje upoznali s pokazateljima. Drugo je rješenje rastaviti kemijske eksperimente na više jednostavnih eksperimenata. Kada ima manje sadržaja ili manje operacija koje u jednom eksperimentu treba izvesti, kognitivno opterećenje nametnuto evaluacijom može se smanjiti. Na kraju, u budućim bi se istraživanjima moglo ispitati kognitivno opterećenje studenata dok koriste skalu kako bi se bolje prilagodili pokazatelji ili kako bi se nastavni proces prilagodio studentima.

Istraživanje je pokazalo da upotreba skale ima mali utjecaj na način na koji studenti organiziraju sadržaj. Ono što im je nedostajalo bilo je nastavno iskustvo, a način na koji su organizirali sadržaj pratio je udžbenik. Kao rezultat toga, studenti i u eksperimentalnoj i u kontrolnoj skupini organizirali su sadržaj na sličan način. Štoviše, zadnji eksperiment podrazumijevao je da će studenti sami osmisliti vlastite eksperimente. Međutim, njihova sposobnost inovacije bila je slaba. Većina ih se odlučila na eksperimente iz udžbenika. Općenito gledajući, studenti iz obje skupine na sličan su način organizirali sadržaj. Osim izrade videozapisa i samoprocjene pomoću skale, u nastavi bi se trebalo više govoriti o organizaciji sadržaja, s ciljem poboljšanja te vještine kod budućih nastavnika.

Nadalje, nije uočena značajna razlika u kvaliteti videozapisa koje su izradili studenti iz dviju skupina. Videozapise su snimali studenti u paru. Koristili su pametne telefone za snimanje svojih kolega tijekom izvođenja eksperimenta. Kvaliteta je stoga bila ograničena funkcijama mobitela i dostupnim aplikacijama za snimanje. Nastavnik je prije nastave objasnio sastavnice bitne za snimanje videozapisa i sve na što treba obratiti pažnju prilikom snimanja. Također je studentima pokazao 2 ili 3 slična videozapisa za analizu. Iako je nastavnik naglasio da bi studenti trebali koristiti vlastitu kreativnost za snimanje videozapisa, u nastavnom procesu studenti su se uglavnom ograničili na izvođenje eksperimenata i na objašnjenja. Kvaliteta videozapisa ograničena je na to da snimka bude čista i mirna ili na to da se dodaju naslovi ili titlovi.

Što se tiče trećega istraživačkog pitanja, studenti su imali pozitivne percepcije o skali. Njezina ih je primjena prisilila da na vrijeme provedu refleksiju. Povratna informacija koja dolazi na vrijeme može uvelike pomoći studentima u procesu učenja (Van den Hurk, Houtveen i Van de Grift, 2016). Na kraju, evaluacija videozapisa može pomoći nastavnicima da iz više perspektiva sagledaju način na koji se poučava o kemijskim eksperimentima te da uoče probleme i moguća rješenja. To im može pomoći u provedbi dublje analize različitih nastavnih praksi (Christ, Arya i Chiu, 2017).

Ograničenja

Ovo istraživanje koristilo je Skalu za samoprocjenu te su zbog toga rezultati ograničeni u nekoliko aspekata. Kao prvo, valjanost samoprocjene podložna je osobnom znanju i iskustvu studenata. Kao drugo, rezultati istraživanja predstavljaju rezultate učenja kroz procjenu videozapisa samo u jednoj obrazovnoj instituciji. Potrebno je provesti više longitudinalnih istraživanja na većem uzorku ispitanika kako bi se utvrdilo na koji se način vremenom mijenjaju sposobnosti studenata za poučavanje o kemijskim eksperimentima. Kao treće, Delfi-metoda ovisi o panelu stručnjaka. Iako smo pokušali angažirati stručnjake iz kemije, tehnologije i obrazovanja s raznolikim nastavnim iskustvom, svi su došli iz iste kulture. Zbog toga su svi u procjeni organizacije sadržaja naglasili pokazatelje, ideologiju i vrijednosti. Oni smatraju da sadržaj ne bi trebao biti samo znanstvene prirode, nego i moralno, politički i društveno korektan. Ovo je već duže vrijeme uobičajen pokazatelj za sve publikacije i filmove koji se u ovoj kulturi stvaraju. Uzimajući u obzir činjenicu da ti aspekti nisu bitni u nastavi Kemije, pokazatelj je na kraju ipak eliminiran. Ipak, on upućuje na mogući problem u primjeni Delfi-metode.

Zaključci

U višestrukim krugovima Delfi-metode izrađeni su pokazatelji bitni za procjenu videozapisa u području kemije. Empirijska studija koja je uslijedila provjerila je korisnost Upitnika za procjenu videozapisa za buduće učitelje jer im on pomaže unaprijediti vještine izvođenja kemijskih eksperimenata. Istraživanje i intervjui pokazali su da je Upitnik za procjenu videozapisa lako koristiti, a studenti su ga voljni primjenjivati u procjeni vlastita poučavanja o kemijskim eksperimentima.

Međutim, zbog relativno jednostavnih eksperimenata, tj. onih s vrlo malo koraka, kao i zbog malog broja inovativnih eksperimenata, Upitnik za procjenu videozapisa nije bio značajno koristan. Lako je ovladati jednostavnim eksperimentima te ih studenti mogu provesti bez ikakve dodatne pomoći. Inovativni eksperimenti zahtijevaju kreativnost, što je izvan opsega upitnika. Ipak, u većini slučajeva upitnik je studentima pomogao smanjiti broj pogrešaka u izvođenju eksperimenata i unaprijediti vještine izvođenja istih.

Nadalje, upitnik nije naglašavao važnost organizacije sadržaja ni kvalitetu videozapisa. To se može pripisati činjenici da studenti kemije imaju dovoljno znanje o sadržajima iz područja kemije te je bilo malo vjerojatno da će napraviti pogreške u osnovnim eksperimentima. U budućnosti bi se ove pokazatelje moglo testirati na studentima koji ne studiraju kemiju ili odabrati eksperimente različitih razina složenosti kako bi se potvrdila njihova učinkovitost.

Na kraju, ranija su istraživanja pokazala da su budući nastavnici imali koristi od toga što su mogli gledati primjere videozapisa dok su pripremali vlastitu nastavu (Baecher, Kung, Jewkes i Rosalia, 2013). U budućim bi se istraživanjima moglo studentima pokazati videozapise o kemijskim eksperimentima i potaknuti ih na razmišljanje o tome kako snimiti dobar videozapis. Osim toga, kvaliteta nastavnikova doprinosa grupnim diskusijama jako je bitna (Vrikki, Warwick, Vermunt, Mercer i Van Halem, 2017). U budućim bi se istraživanjima moglo analizirati sudjelovanje nastavnika u aktivnostima u kojima studenti procjenjuju videozapise i pokazati kako ono utječe na upotrebu skale. Na kraju, u budućim bi se istraživanjima mogao koristiti Upitnik za procjenu videozapisa kako bi se unaprijedila vještina snimanja videozapisa o kemijskim eksperimentima za potrebe, na primjer, različitih MOOC-ova ili za evaluaciju videozapisa o kemijskim elementima koji su prikladni za studente. Budući nastavnici Kemije, kao i oni koji već rade kao nastavnici Kemije, mogu Upitnik koristiti za unapređenje vlastita poučavanja i u pripremi materijala za učenike.

Zahvale

Ovaj rad podupiru 13. petogodišnji plan obrazovnih znanosti u provinciji Jiangsu u Kini (YZ-c/2016/18), FinCEAL Plus (51770) te Finska akademija (318380).