LUMAT
International Journal on Math, Science and Technology Education

Research article

Development and impact of histoguide application towards drawing and labelling in microscopic practical

Muhamad Ikhwan Mat Saad[1], Teoh Chern Zhong[1], Mohd Mokhzani Ibrahim[1], Mohd Termizi Borhan[1], Eng Tek Ong[2] and Mohd Afifi Bahurudin Setambah[1]

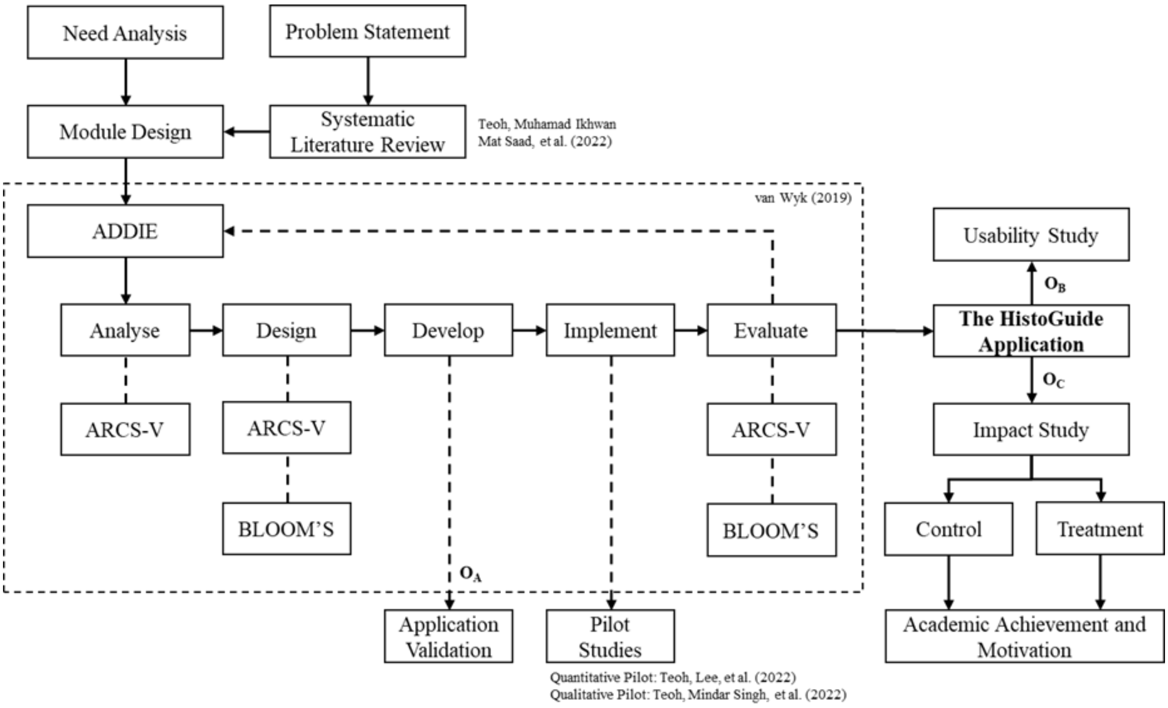
[1] Universiti Pendidikan Sultan Idris, Malaysia
[2] UCSI University, Malaysia

Abstract: The HistoGuide is an Android application used for virtual microscopy and slides to solve the problems of incorrect drawing and labelling in microscopic practicals. It is developed based on modified Analyze, Design, Develop, Implement, and Evaluate (ADDIE), the van Wyk model, as a self-regulated mobile learning, complementary to optical microscopy. However, as a newly developed application, many still do not understand the usability and impact of virtual microscopy. Hence, the HistoGuide was validated using Cohen's kappa agreement coefficient, strengthened with the Content Validity Index (CVI). Data were analysed descriptively using mean, standard deviation and percentages for the usability study and inferentially using independent and paired sample t-tests for the impact study. Findings revealed that Cohen's kappa for content, pedagogy, and technology constructs are 1.00, 1.00, and 0.90, respectively, with an overall of 0.96. The HistoGuide application also achieved high I-CVI and excellent content validity of the overall validation with S-CVI/UA of 0.80 and S-CVI/Ave of 0.96. As for the usability study, the HistoGuide application recorded a high usability level for the overall usability and its four usability constructs: usefulness, ease of use, ease of learning and satisfaction. In the assessment achievement study, there were significant differences between pre- and post-test scores for the treatment group and post-test scores between the treatment and control groups. Thus, the treatment group performed very well compared to the control group in terms of assessment achievement. In the motivation study, the treatment group performed better than the control in motivation and its five motivation constructs: attention, relevance, confidence, satisfaction, and volition. Overall, students from the treatment group outperformed in assessment achievement and motivation compared to the control after using the HistoGuide application. HistoGuide application could enhance the drawing and labelling based on the usability and impact study. This study implies that virtual microscopy could promote innovative learning of microscopic practicals.

Keywords: HistoGuide, virtual microscopy, validity, usability, labelling, drawing, assessment achievement, motivation

Correspondence: mohdafifi@fpm.upsi.edu.my

Received 30 August 2024 / Accepted 4 March 2025 / Published 12 March 2025
Published by the University of Helsinki, Finland and LUMA Centre Finland.



1 Introduction

During the COVID-19 pandemic, schools shifted to online teaching, which is also true in the case of Malaysia. Thus, educators employed various teaching methods, including virtual microscopy, during the lockdown to replace laboratory work. However, during post-pandemic, these practices have continued. Thus, the acquisition of students' manipulative skills (dexterity) has been a concern regarding laboratory work (Mojica & Upmacis, 2022).

Preliminary findings from two biology classes in a sixth-form centre (equivalent to matriculation/pre-university/A-level studies) revealed that 68.8% of students scored below five marks in the results section of the School-Based Assessment (SBA). Drawing, labelling, and applying magnification and scales encompass about 40.0% of the marks allocated in the single assessment, which consisted of manipulative skills (A), results (B), discussion (C), and conclusion (D). Accordingly, students are considered weak in drawing, labelling and applying magnification and scales. Drawing and labelling aspects must be fulfilled together to acquire the allocated marks. Cheung and Winterbottom (2021) supported the findings, which explored students' visualisation competence and discovered that they are weak in perceiving microscopic entities through drawing and labelling. They reported that 60.0% of the students could not label their biological drawings, and a higher proportion tended to give fewer labels (Cheung & Winterbottom, 2021). In addition, students have a moderately low level of sketching skills (Fatimah Mohamed et al., 2011). It might be due to students' inability to draw, label, and observe details, as there is a lack of quality practical images (García et al., 2019). Consequently, these incorrect drawings and labelling or both will thus decrease the SBA and students' motivation to execute practically.

1.1 Literature review

Hence, a comprehensive guideline is necessary for guiding students in executing microscopic practicals. A guide is essential to help students draw and label precisely, bearing in mind the usefulness, ease of use, ease of learning and satisfaction the guide causes to the students. Based on the systematic literature review on leveraging technology in Histology (Teoh, Muhamad Ikhwan Mat Saad, et al., 2022) and the need analysis study, virtual microscopy is recommended to suit the current students' learning styles post-pandemically (Lockee, 2021). The HistoGuide application is employed as virtual microscopy and slides for self-regulated learning and is feasible and complementary to optical microscopy. The HistoGuide application can view the mentioned prepared slides as users would in optical microscopy without losing the correct magnification and size. The application is developed based on van Wyk (2019), who proposed a new framework for developing online courses, the van Wyk model. The model was a course development framework for emerging technologies (van Wyk et al., 2020). It comprises the ADDIE instructional design model, including depictions of where Attention, Relevance, Confidence, and. Satisfaction (ARCS-V) and Bloom's taxonomy are used to enhance learning motivation and evaluate learning objectives. Remarkably, it was reported to be especially valid and useful when integrating motivational aspects at each level of the ADDIE design process (van Wyk et al., 2020). Despite that, as a newly developed application, many still do not understand the usability and impact of virtual microscopy. Thus, developing and assessing the HistoGuide application as a virtual microscopy is necessary.

Surveys are often used to obtain self-reported data about users' interactions with a particular product or system. The usefulness, satisfaction, and ease of use (USE) questionnaire by Lund (2001) measures a product's or service's subjective usability. The items underwent a complete psychometric instrument development process to develop a standardised instrument (Gao et al., 2018). Using the USE questionnaire, Hariyanto et al. (2020) revealed that the adaptive e-learning system's usability for students was initially well-approved in all dimensions of usability. Usability testing refers to testing the usability of the HistoGuide application through four constructs: usefulness, ease of use, ease of learning and satisfaction. In particular, the usability testing uses the survey method by distributing survey instruments to the respondents. The survey has 30 items from the USE questionnaire (Lund, 2001), adapted from Hariyanto et al. (2020). The data is analysed to gauge the usability of the HistoGuide application.

As for the impact study, the impact is defined as a change due to an action or other causes. The positive impact or benefits happens when a desired or positive impact on the outcome makes the learning process efficient and motivating (Giannakas et al., 2018; Lavasani et al., 2011). Therefore, the impact is associated with the competence of the HistoGuide application as a student's learning tool to gauge their SBA achievement and motivation in microscopic practicals. Both impacts are measured by the differences in the mean score of students' pre and post-tests for both control and treatment groups.

Meanwhile, the Instructional Materials Motivation Survey (IMMS) questionnaire developed by Keller (2010) has been widely utilised as a pre and post-test tool for assessing motivational needs before training or measuring people's reactions to instructional materials afterwards. However, it has not been substantially validated (Loorbach et al., 2015). Hence, according to the results of structural equation modelling, the IMMS can be reduced to 12 items, and the Reduced Instructional Materials Motivation Survey (RIMMS) is preferable to the original IMMS (Loorbach et al., 2015). Furthermore, it has been confirmed that the ARCS model's conditional nature is reflected in the RIMMS. It also suggests that the ARCS model's underlying motivating theory holds and is represented in the RIMMS in the self-regulated instructional setting geared at working with technology (Loorbach et al., 2015). Correspondingly, in this study, the RIMMS and the volition questionnaire (Keller et al., 2020) will be adopted to analyse the impact of the HistoGuide application.

In summary, the objectives of the study are as follows.

i. Develop a mobile application as a self-regulated guide in drawing and labelling with a high validity index (O_A).
ii. Determine the usability index of the newly developed HistoGuide application (O_B).
iii. Determine the impact of the HistoGuide application on academic achievement and motivation (O_C).

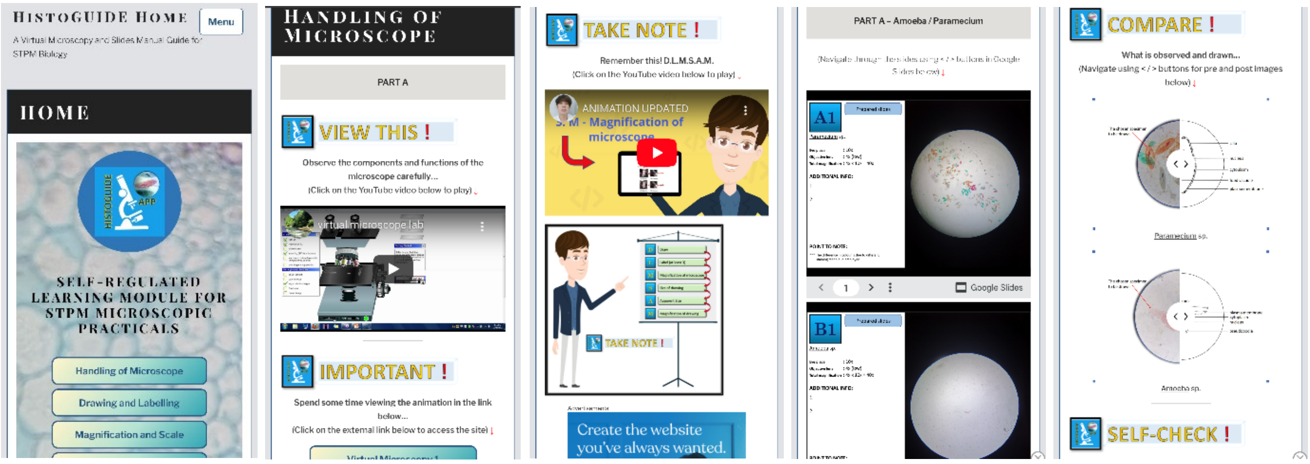
2 Methodology

The research employed a quantitative study. At the same time, the research applied the developmental research design and survey design for the usability study, while the quasi-experimental design for the impact study, as displayed in Figure 1. The developmental research involves designing and developing the HistoGuide application based on the modified ADDIE, the van Wyk model.

After the development, 15 experts validated the HistoGuide application in content, pedagogical, and technology aspects, including qualitative insights, such as suggestions from experts' validation and the revisions done. As a newly developed mobile application, the HistoGuide underwent strong content validation using Cohen's kappa coefficient. In addition, the most widely used index in the quantitative evaluation of the content validity of a scale is the Content Validity Index (CVI) (Polit et al., 2007). Thus, for CVI, I-CVI (Item-level CVI), S-CVI/UA (Scale-level CVI/Universal Agreement approach) and S-CVI/Ave (Scale-level CVI/Average approach) are employed. Note that I-CVI and S-CVI approaches can lead to different values, making it challenging to conclude content validity (Rodrigues et al., 2017). Thus, this study considered both the I-CVI and the S-CVI since the S-CVI is an average score that outliers can skew to support and strengthen the validation through Cohen's kappa coefficient. Items with an I-CVI of 0.78 or higher for three or more experts could be considered evidence of good content validity (Polit et al., 2007).

Consequently, a qualitative pilot study (Teoh, Mindar Singh, et al., 2022) and a quantitative pilot study (Teoh, Lee, et al., 2022) on the HistoGuide prototype were implemented. The validity and reliability of instruments were also sought. The finalised product, the HistoGuide application, illustrated in Figure 1, underwent a usability study with 126 respondents and an impact study with 68 respondents.

Figure 1. Screenshots of the histoguide application



3 Data findings

The HistoGuide application was validated in content, pedagogical, and technological aspects using Cohen's kappa coefficient and CVI, as summarised in Table 1. The kappa coefficient on a 2-point agreement scale for content, pedagogy, technology and overall aspect of the HistoGuide application are 1.00, 1.00, 0.90 and 0.96, respectively. Meanwhile, the overall Cohen's kappa is 0.96, and according to Cohen (1960), Cohen's kappa coefficient, $\kappa > 0.81$, demonstrated a very good agreement between the raters. Since Cohen's kappa coefficient is greater than 0.81, the HistoGuide application is deemed highly valid.

In addition, CVI was also referred to strengthen and determine the content validity. The I-CVIs of all the 30 items in the HistoGuide application validation ranged from 0.80 to 1.00, with only six items having an I-CVI of less than 1.00. Values range from 0 to 1, where I-CVI > 0.79, the item is relevant. Between 0.70 and 0.79, the item needs revisions, and if the value is below 0.70, the item is eliminated (Zamanzadeh et al., 2015). Ultimately, none of the 30 items is eliminated. Only six undergo minor revisions. However, CVI alone is insufficient. Therefore, kappa statistic was employed together. Kappa statistic in CVI is a consensus index of inter-rater agreement that adjusts for chance agreement and is an essential supplement to CVI since kappa provides information about the degree of agreement beyond chance (Polit et al., 2007). Table 2 indicates that all 30 items are still excellent. The S-CVI/UA for content, pedagogy, technology and overall aspect of the HistoGuide application are 1.00, 0.60, 0.80 and 0.80, respectively. The S-CVI/Ave for content, pedagogy, technology and overall aspect of the HistoGuide application are 1.00, 0.92, 0.96 and 0.96, respectively. An S-CVI/UA ≥ 0.80 and an S-CVI/Ave ≥ 0.90 have excellent content validity (Shi et al., 2012). Nevertheless, the overall CVI of the instrument using the universal agreement approach was low in terms of the pedagogy aspect (0.60). Thus, the high number of content experts can be advocated, making consensus difficult (Zamanzadeh et al., 2015) and compensating by the high value of the S-CVI/Ave, which was equal to 0.96.

In summary, the HistoGuide application revealed high content validity of individual items (I-CVI range: 0.80 to 1.00) and excellent content validity of the overall validation (S-CVI/UA = 0.80; S-CVI/Ave = 0.96). In addition, qualitative methods refined the clarity of items for the mentioned six revisioned items (Rodrigues et al., 2017). Using Cohen's kappa on a 2-point agreement scale or S-CVI/Ave, both methods recorded a high validity index of 0.96. Hence, it can be concluded that the developed HistoGuide application has a high validity index of 0.96, which is very good.

Table 1. Summaries of findings for research objective (O_A)

	Content [a]	**Pedagogy** [b]	**Technology** [c]	**Overall** [d]
Cohen's kappa validity index on a 4-Point Agreement Scale	0.90	0.70	0.90	0.83
Description of Cohen's kappa on a 4-Point Agreement Scale *	Very Good	Good	Very Good	Very Good
Cohen's kappa validity index on a 2-Point Agreement Scale	1.00	1.00	0.90	0.96
Description of Cohen's kappa on a 2-Point Agreement Scale *	Very Good	Very Good	Very Good	Very Good
	Content [e]	Pedagogy [f]	Technology [g]	Overall [h]
Average I-CVI	1.00	0.92	0.96	0.96
S-CVI/UA	1.00	0.60	0.80	0.80
S-CVI/Ave	1.00	0.92	0.96	0.96

Note:
a evaluated by Experts 01 and 02
b evaluated by Experts 07 and 08
c evaluated by Experts 13 and 14
d is the average of the three constructs (content, pedagogy and technology)
e evaluated by Experts 01, 02, 03, 04 and 05
f evaluated by Experts 06, 07, 08, 09 and 10
g evaluated by Experts 11, 12, 13, 14, and 15
h is the average of the three constructs (content, pedagogy and technology)
* Cohen's kappa description is interpreted from Cohen (1960)

The characteristics of the respondents in this study are classified in Table 3. There were 126 respondents for the usability test. The demographic data presents respondents from mode two sixth-form centres (74.60%) and national-type secondary schools/SMJK (40.50%) dominated the study. Female respondents in gender categories (65.90%), respondents from the city area (69.00%) and Android users (70.60%) recorded more than two-thirds of the total respondents. In summary, the mean and standard deviation scores for the usability of the HistoGuide application are 4.41 and 0.39, respectively. By having a percentage of 88.20% and from the deciphered mean score level, which is high, the HistoGuide application can be considered to have a high usability level. In addition, the four constructs of usability are also mentioned in Table 4. The mean scores for usefulness, ease of use, ease of learning, and satisfaction were 4.40, 4.36, 4.45 and 4.47, respectively. Note that all four constructs recorded a high level of usability.

Table 2. Ratings on a 30-item scale by 15 experts: Item rated 3 or 4 on a 4-point relevance scale

Experts	**01**	**02**	**03**	**04**	**05**	**Experts in agreement**	**Item CVI**	***p_c***	***k****	**Evaluation**
Items										
Content										
C1	/	/	/	/	/	5	1.00	0.041	1.00	Excellent
C2	/	/	/	/	/	5	1.00	0.041	1.00	Excellent
C3	/	/	/	/	/	5	1.00	0.041	1.00	Excellent
C4	/	/	/	/	/	5	1.00	0.041	1.00	Excellent
C5	/	/	/	/	/	5	1.00	0.041	1.00	Excellent
C6	/	/	/	/	/	5	1.00	0.041	1.00	Excellent
C7	/	/	/	/	/	5	1.00	0.041	1.00	Excellent
C8	/	/	/	/	/	5	1.00	0.041	1.00	Excellent
C9	/	/	/	/	/	5	1.00	0.041	1.00	Excellent
C10	/	/	/	/	/	5	1.00	0.041	1.00	Excellent
Proportion	1.00	1.00	1.00	1.00	1.00	Average I-CVI	**1.00**			
						S-CVI/UA	**1.00**			
						S-CVI/Ave	**1.00**			
Experts	04	05	06	07	08					
Pedagogy										
P1	/	/	/	/	/	5	1.00	0.041	1.00	Excellent
P2	X	/	/	/	/	4	0.80	0.156	0.76	Excellent
P3	/	/	/	/	/	5	1.00	0.041	1.00	Excellent
P4	X	/	/	/	/	4	0.80	0.156	0.76	Excellent
P5	/	/	/	/	/	5	1.00	0.041	1.00	Excellent
P6	X	/	/	/	/	4	0.80	0.156	0.76	Excellent
P7	/	/	/	/	/	5	1.00	0.041	1.00	Excellent
P8	/	/	/	/	/	5	1.00	0.041	1.00	Excellent
P9	/	/	/	/	/	5	1.00	0.041	1.00	Excellent
P10	X	/	/	/	/	4	0.80	0.156	0.76	Excellent
Proportion	0.60	1.00	1.00	1.00	1.00	Average I-CVI	**0.92**			
						S-CVI/UA	**0.60**			
						S-CVI/Ave	**0.92**			
Experts	09	10	11	12	13					
Technology										
T1	/	/	/	/	X	4	0.80	0.156	0.76	Excellent
T2	/	/	X	/	/	4	0.80	0.156	0.76	Excellent
T3	/	/	/	/	/	5	1.00	0.041	1.00	Excellent
T4	/	/	/	/	/	5	1.00	0.041	1.00	Excellent
T5	/	/	/	/	/	5	1.00	0.041	1.00	Excellent
T6	/	/	/	/	/	5	1.00	0.041	1.00	Excellent
T7	/	/	/	/	/	5	1.00	0.041	1.00	Excellent
T8	/	/	/	/	/	5	1.00	0.041	1.00	Excellent
T9	/	/	/	/	/	5	1.00	0.041	1.00	Excellent
T10	/	/	/	/	/	5	1.00	0.041	1.00	Excellent
Proportion	1.00	1.00	0.90	1.00	0.90	Average I-CVI	**0.96**			
						S-CVI/UA	**0.80**			
						S-CVI/Ave	**0.96**			

Note:

p_c (probability of a chance occurrence) was calculated using the formula for a binomial random variable, with one specific outcome: $p_c = [N!/A!(N – A)!]*.5^N$ where N = number of experts and A = Number agreeing on good relevance.

k* = kappa designating agreement on relevance: $k* = (\text{I-CVI} – p_c)/(1– p_c)$.

Evaluation criteria for kappa, using guidelines described in Cicchetti and Sparrow (1981) and Fleiss (1981): Fair = k* of 0.40 to 0.59; Good = k* of 0.60 to 0.74; and Excellent = k* > 0.74.

The characteristics of the respondents in this study are classified in Table 3. There were 126 respondents for the usability test. The demographic data displays respondents from mode two sixth-form centres (74.60%) and national-type secondary schools/SMJK (40.50%) dominated the study. Female respondents in gender categories (65.90%), respondents from the city area (69.00%) and Android users (70.60%) recorded more than two-thirds of the total respondents. In summary, the mean and standard deviation scores for the usability of the HistoGuide application are 4.41 and 0.39, respectively. Notably, by having a percentage of 88.20% and from the deciphered mean score level, which is high, the HistoGuide application can be considered to have a high usability level. In addition, the four constructs of usability are also mentioned in Table 4. The mean scores for usefulness, ease of use, ease of learning, and satisfaction were 4.40, 4.36, 4.45 and 4.47, respectively. All four constructs recorded a high level of usability.

Table 3. Demographic data for usability study (n = 126)

Categories	Descriptions	Frequency	Percentage (%)
Type	Mode 1 Sixth-form Centre *	25	19.80
	Mode 2 Sixth-form Centre	94	74.60
	Mode 3 Sixth-form Centre	7	5.60
School	National Secondary School/SMK	39	31.00
	National-type Secondary School/SMJK	51	40.50
	Missionary	8	6.30
	Private	3	2.40
	Others	25	19.80
Gender	Male	43	34.10
	Female	83	65.90
Location	City	87	69.00
(Area)	Rural	39	31.00
OS	Android	89	70.60
	iOS	37	29.40

Note: * also known as Sixth-form College

Table 4. Summaries of findings for research objective (O_B)

Constructs (No. of items)	Mean Score	Standard Deviation	Percentage * (%)	Mean score ** Level
Usefulness (8)	4.40	0.3829	88.00	High
Ease of use (11)	4.36	0.4113	87.20	High
Ease of learning (4)	4.45	0.4469	89.00	High
Satisfaction (7)	4.47	0.4452	89.40	High
Usability (30)	4.41	0.3915	88.20	High

Note:
* Percentage is counted using the formula (mean score*100/5).
** Mean score level is interpreted from Hamidah Yusof et al. (2015).

In the impact study, both the pre-test and post-test scores for the assessment achievement and motivation were determined by the normal distribution, referring to skewness and kurtosis value. As for the assessment achievement study, there were no significant differences in the pre-test scores between the treatment and control groups. Furthermore, there were no significant differences between pre and post-test scores for the control group. However, there were significant differences between pre- and post-test scores for the treatment group and post-test scores for the treatment and control groups. Thus, the treatment group performed very well compared to the control group in the assessment achievement of SBA 1 and 2. As for the motivation study, there were no significant differences in the pre-test scores between the treatment and control groups. Moreover, there were significant differences between pre and post-test scores for the treatment and control groups, respectively. There were also significant differences in the post-test scores between the treatment and control groups. Nonetheless, the treatment group performed better than the control group in motivation and its five motivation constructs: attention, relevance, confidence, satisfaction, and volition. Overall, students from the treatment group outperformed those from the control group in assessment achievement and motivation after using the HistoGuide application, as provided in Table 5. HistoGuide application has impacted the assessment achievement and motivation in sixth-form microscopic practicals. Accordingly, it could enhance the student's skill in drawing and labelling based on the usability and impact study.

Table 5. Summaries of findings for research objective (O_C)

Research questions for research objective (O_C)	Null hypothesis	Result	Remark
3a: Is there any significant difference in students' assessment achievement between the treatment group and control group scores in the pre-test?	H_{01}: There is no significant difference in students' assessment achievement between the treatment group and control group scores in the pre-test.	SBA1: ($t(35) = -0.292$, $p = 0.772$) Null hypothesis fails to be rejected as $p \geq 0.05$ SBA2: ($t(29) = 0.524$, $p = 0.604$) Null hypothesis fails to be rejected as $p \geq 0.05$	There is no significant difference in students' assessment achievement between the treatment group and control group scores in the SBA1 and SBA2 pre-test
3b: Is there any significant difference in students' assessment achievement between pre-test and post-test scores for the treatment group?	H_{02}: There is no significant difference in students' assessment achievement between pre-test and post-test scores for the treatment group.	SBA1: ($t(18) = -4.253$, $p = 0.000$) Null hypothesis rejected as $p < 0.05$ SBA2: ($t(14) = -2.955$, $p = 0.010$) Null hypothesis rejected as $p < 0.05$	There is a significant difference in students' assessment achievement between pre-test and post-test scores for the treatment group
3c: Is there any significant difference in students' assessment achievement between pre-test and post-test scores for the control group?	H_{03}: There is no significant difference in students' assessment achievement between pre-test and post-test scores for the control group.	SBA1: ($t(17) = -1.567$, $p = 0.135$) Null hypothesis fails to be rejected as $p \geq 0.05$ SBA2: ($t(15) = -1.576$, $p = 0.136$) Null hypothesis fails to be rejected as $p \geq 0.05$	There is no significant difference in students' assessment achievement between pre-test and post-test scores for the control group
3d: Is there any significant difference in students' assessment achievement between the treatment group and control group scores in the post-test?	H_{04}: There is no significant difference in students' assessment achievement between the treatment group and control group scores in the post-test.	SBA1: $t(35) = 2.396$, $p = 0.022$) Null hypothesis rejected as $p < 0.05$ SBA2: $t(29) = 2.123$, $p = 0.042$) Null hypothesis rejected as $p < 0.05$	There is a significant difference in students' assessment achievement between the treatment group and control group scores in the SBA1 and SBA2 post-test
4a: Is there any significant difference in students' motivation between the treatment group and control group scores in the pre-test?	H_{05}: There is no significant difference in students' motivation between the treatment group and control group scores in the pre-test.	Motivation: ($t(35) = 0.183$, $p = 0.856$) The null hypothesis fails to be rejected as $p \geq 0.05$	There is no significant difference in students' motivation between the treatment group and control group scores in the pre-test

		*attention: (t(35) = 0.439, p = 0.664) *relevance: (t(35) = 1.540, p = 0.133) *confidence: (t(35) = 0.032, p = 0.639) *satisfaction: (t(35) = 0.925, p = 0.361) *volition: (t(35) = -0.473, p = 0.639) The null hypothesis for each construct fails to be rejected as $p \geq 0.05$	There is no significant difference in students' attention, relevance, confidence, satisfaction and volition between the treatment group and control group scores in the pre-test
4b: Is there any significant difference in students' motivation between pre-test and post-test scores for the treatment group?	H_{06}: There is no significant difference in students' motivation between pre-test and post-test scores for the treatment group.	Motivation: (t(18) = -7.681, p = 0.000) The null hypothesis was rejected as $p < 0.05$	There is a significant difference in students' motivation between pre-test and post-test scores for the treatment group
		*attention: (t(18) = -4.296, p = 0.000) *relevance: (t(18) = -3.637, p = 0.002) *confidence: (t(18) = -6.296, p = 0.000) *satisfaction: (t(18) = -5.115, p = 0.000) *volition: (t(18) = -7.053, p = 0.000) The null hypothesis for each construct was rejected as $p < 0.05$	There is a significant difference in students' attention, relevance, confidence, satisfaction and volition pre-test and post-test scores for the treatment group
4c: Is there any significant difference in students' motivation between pre-test and post-test scores for the control group?	H_{07}: There is no significant difference in students' motivation between pre-test and post-test scores for the control group.	Motivation: (t(17) = -2.962, p = 0.009) The null hypothesis was rejected as $p < 0.05$	There is a significant difference in students' motivation between pre-test and post-test scores for the control group
		*attention: (t(17) = -2.404, p = 0.028) *volition: (t(17) = -3.855, p = 0.001) The null hypothesis for each construct was rejected as $p < 0.05$	There is a significant difference in students' attention and volition pre-test and post-test scores for the control group
		*relevance: (t(17) = -1.844, p = 0.083) *confidence: (t(17) = -0.622, p = 0.542) *satisfaction: (t(17) = -1.844, p = 0.083) The null hypothesis for each construct fails to be rejected as $p \geq 0.05$	There is no significant difference in students' relevance, confidence and satisfaction pre-test and post-test scores for the control group
4d: Is there any significant difference in students' motivation between the treatment group and control group scores in the post-test?	H_{08}: There is no significant difference in students' motivation between the treatment group and control group scores in the post-test.	Motivation: (t(35) = 3.444, p = 0.002) The null hypothesis was rejected as $p < 0.05$	There is a significant difference in students' motivation between the treatment group and control group scores in the post-test
		*confidence: (t(35) = 5.458, p = 0.000) *satisfaction: (t(35) = 2.842, p = 0.007) *volition: (t(35) = 2.506, p = 0.017) The null hypothesis for each construct was rejected as $p < 0.05$	There is a significant difference in students' confidence, satisfaction and volition between the treatment group and control group scores in the post-test

		*attention: (t(35) = 1.735, p = 0.092), *relevance: (t(35) = 1.854, p = 0.072) The null hypothesis for each construct fails to be rejected as p ≥ 0.05	There is no significant difference in students' attention and relevance between the treatment group and control group scores in the post-test

Note: * Constructs of Motivation in ARCS-V

4 Discussion and implication of research

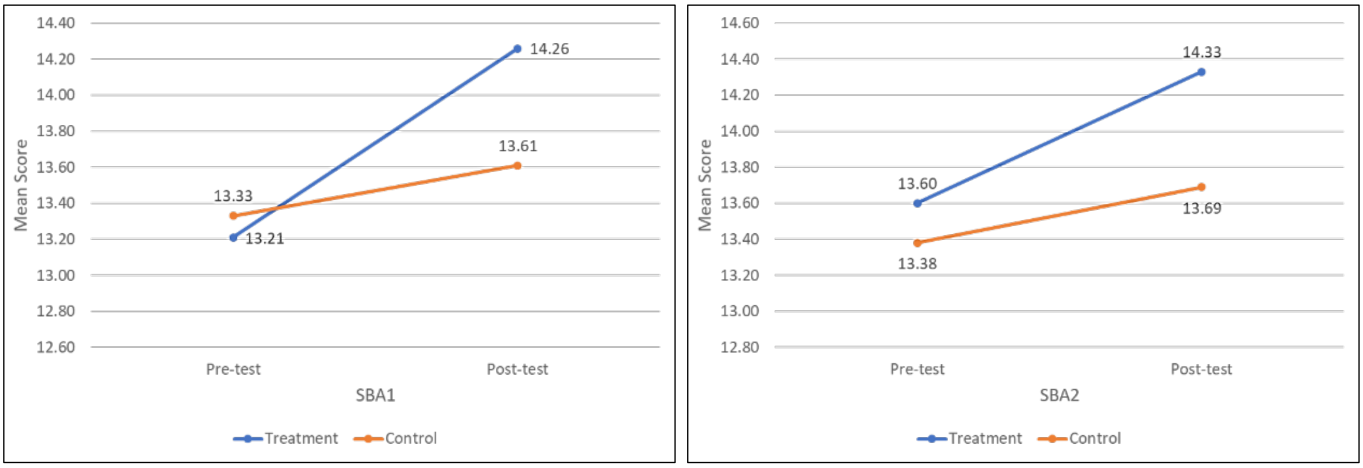
The study focused on measuring the validity and usability of virtual microscopy and slides. Based on the collected data, the overall views of the experts and respondents were reasonably positive towards using the HistoGuide application, the flexibility provided in the mobile applications, and the user-friendly interfaces.

In addition to the knowledge attainment measure and solving the microscopic drawing and labelling problems, this usability study revealed that students had a positive perception of the usability of virtual microscopy as a learning tool, evidenced by a significantly higher satisfaction score of 89.40%. Although several past studies have recorded students' opinions, insufficient data is drawn from directly measuring students' level of satisfaction with virtual microscopy or optical microscopy. Notably, this usability study aligns with the study by Hande et al., which measured student satisfaction using optical, virtual, and optical and virtual microscopes (Hande et al., 2017). Their study suggested 87.6% satisfaction with virtual microscope usage (Hande et al., 2017).

Overall, this usability study enhances and is aligned with other previous studies. This study supported other research that reported several advantages of using a virtual microscope as a learning tool, namely easy navigation with optimum contrast, clear images, presence of interactive features that allow collaborative learning and easy access to virtual microscopy for self-regulated study (Nauhria & Hangfu, 2019). In addition, previous studies suggest that a crucial factor of virtual microscopy is the facilitation of collaboration (Triola & Holloway, 2011), a prominent feature of the HistoGuide application. However, it should be noted that some students and educators also strongly preferred the continued use of traditional microscopy, supplemented with virtual microscopy, as both tools in adjunct optimised students' learning (Xu, 2013). Additionally, Raja (2010) presented similar findings, where students accepted optical microscopy as a supplementary learning tool. Hence, the HistoGuide application is suggested as complementary to optical microscopy due to its high usability in terms of usefulness, ease of use, ease of learning, and satisfaction.

The impact study on assessment achievement exhibited a significant difference in pre-test and post-test scores for the treatment and control groups, respectively. It might be closely related to the exercises given to the students during treatment, which could train them to become more skilful in drawing and labelling. The treatment effect occurs in SBA1, as depicted in Figure 2. It demonstrates that the treatment effectively increases the SBA1 achievement in the treatment group. In summary, both SBA1 and SBA2 recorded significant assessment achievement for the treatment groups in the post-test. Building on this, the HistoGuide application has impacted the students' SBA achievement in microscopic practicals. However, other effects are probably present. Regression towards the direction of the population mean may occur (Chua, 2020). In addition, there may be a maturity, historical, or testing effect (Chua, 2020), enabling both respondent groups to obtain higher scores in the post-test. Nevertheless, all those threats were adequately attended to and addressed when employing quasi-experimental research.

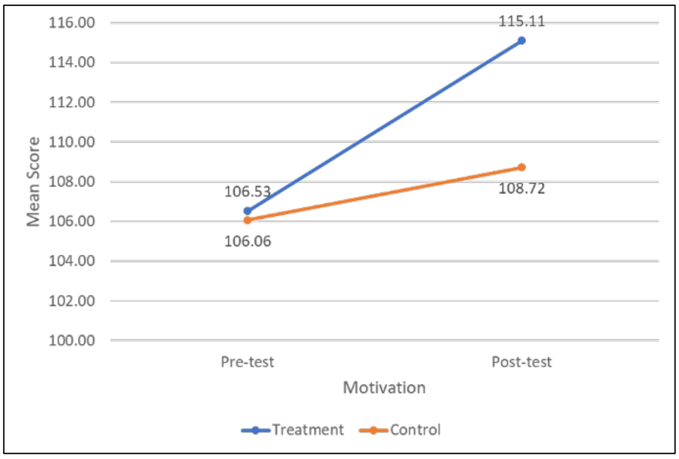
Figure 2. Mean score between pre-test and post-test for SBA1 and SBA2



The present study result is in line with the research conducted by Cheng et al. (2017) and Tauber et al. (2019), whereby students' test results are more successful with virtual microscopy with slides due to the improvement of the quality of preparation for microscopic practicals (Cheng et al., 2017; Tauber et al., 2019). In addition, the students obtained better results in performance with virtual microscopy and slides (Bacha et al., 2020; Manganello et al., 2019). Considering this perspective, this study provides clear evidence for the impact of the HistoGuide application, a self-regulated web-enhanced learning, on the students' assessment achievement.

The impact study on motivation significantly differed in pre-test and post-test scores for the treatment and control groups, respectively. It might be closely related to the exercises given to the students during treatment, which could train them to become more motivated in drawing and labelling. The motivation graph, as shown in Figure 3, indicates that both groups' motivational levels increased during the post-test.

Figure 3. Mean Score between Pre-test and Post-test for Motivation



The present study recorded significant motivation achievement for the treatment groups in the post-test. The HistoGuide application has impacted students' motivation in microscopic practicals. The present study result aligns with the research conducted by Cheng et al. (2017) and Tauber et al. (2019). The researchers also noted higher student motivation (Cheng et al., 2017; Tauber et al., 2019). This study contradicted Simok et al. (2019), who mentioned that virtual microscopy failed to stimulate students' motivation compared to optical microscopy. The researchers suspected it might be due to lacking a motivational model in their module. Hence, the HistoGuide application is designed according to Cromley et al. (2020) suggestion. Meanwhile, Cromley et al. (2020) concluded that combinations of cognitive and motivational interventions were offered to increase students' performance with minimal additional work for the instructors. At the same time, van Wyk (2019), who mentioned that the ARCS-V motivational design model is appropriate for designing motivational features in a course, also adhered to. The motivational theories included self-determination, flow, and the ARCS model formed the basis of the HistoGuide application. These motivational features stimulated students' motivation in the treatment group.

However, there was a significant difference between the pre-test and post-test scores for the motivation control group. It can be concluded that there is a significant difference in students' motivation between pre-test and post-test scores for the control group. Accordingly, both the treatment and control groups recorded significantly increased motivation. The significantly increased motivation for the control group might be due to the use of the ministry's SBA manual. Moreover, attention and volition constructs recorded significant differences as well. However, by not using the HistoGuide application to complement the assessment manual, the relevance, confidence and satisfaction constructs were not significantly increased. In conclusion, there was a significant difference in students' motivation between the treatment group and control group scores in the post-test. The constructs of confidence, satisfaction and volition also recorded significant differences. Conversely, attention and relevance constructs were not significantly increased since both the treatment and control groups utilised the assessment manual as the basis of the study. Concurrently, the HistoGuide application as virtual microscopy is recommended to complement optical microscopy as it enables users to draw and label better for their microscopic practicals, influencing their assessment achievement and motivation to use it.

Indeed, many challenges are faced when implementing virtual microscopy in schools. One notable challenge faced is the internet availability and connectivity for the HistoGuide application to function, as it is also a platform for communication.

6 Conclusion

Students faced problems in drawing and labelling for their microscopic practicals, which would affect their understanding of subsequent subtopics or even the theoretical class. If the students' problems are not solved, it might impact or influence their assessment achievement and motivation in the practical examination. In particular, current students are of the Gen-Z generation and are attracted to mobile devices daily. Thus, mobile smartphone applications could be students' assistants in self-regulated learning. In this study, the HistoGuide application is an Android-based application developed based on the modified ADDIE, the van Wyk model, to study its usability and impact on students' assessment achievement and motivation. The HistoGuide application is hoped to assist students in addressing their weaknesses in drawing and labelling during microscopic practicals. Notably, the validation of the HistoGuide application was estimated by six experts using Cohen's kappa coefficient on a 2-point agreement scale for content, pedagogy, technology, and overall aspects of the HistoGuide application, resulting in 1.00, 1.00, 0.90 and 0.96, respectively. Further validation with 15 experts using CVI produced high I-CVI and excellent content validity of the overall validation with S-CVI/UA of 0.80 and S-CVI/Ave of 0.96. In addition, the HistoGuide application was reported to be of high usability level for the overall usability and its four usability constructs: usefulness, ease of use, ease of learning and satisfaction. The usability study involved 126 students. Furthermore, the first impact study involved 37 students and the second involved 31 students. The assessment achievement and motivation were measured using pre- and post-test scores. Overall, students from the treatment group outperformed in assessment achievement and motivation compared to students from the control group after using the HistoGuide application. HistoGuide application has impacted assessment achievement and motivation in microscopic practicals. In essence, it could enhance the student's skill in drawing and labelling based on this usability and impact study.

Thus, it is recommended that the HistoGuide application as virtual microscopy is utilised to complement optical microscopy as it enables users to be able to draw and label better for their microscopic practicals, possibly influencing the assessment achievement and their motivation to use it.

The assumptions presented in this study were: (a) virtual microscopy will have an influence on the way that students perceive laboratory activities; (b) the sample selected is representative and adequate; (c) the researcher maintains neutrality on teaching with digital or virtual technology. The study's limitations were: (a) students will find it inconvenient to use virtual microscopy and slides when interpretation is obtained from cloud-based secure systems, and (b) the internet availability. The recommended future development is (a) an iOS-based HistoGuide application.

Acknowledgement

This research is supported by Universiti Pendidikan Sultan Idris through Geran Galakan Penyelidikan Universiti 2021 (Grant number: 2021-0109-107-01). This study was carried out with the cooperation of Sektor Pembelajaran Jabatan Pendidikan Negeri Pulau Pinang (JPNPP) and Jabatan Pendidikan Wilayah Persekutuan Kuala Lumpur (JPWPKL).

References

Bacha, D., Ferjaoui, W., Charfi, L., Rejaibi, S., Slama, S. Ben, Njim, L., & Lahmar, A. (2020). The interest of virtual microscopy as a means of simulation learning in pathological anatomy and cytology. *Onkologia i Radioterapia*, *14*(5), 23–29.

Cheng, X., Lee, K. K. ho, Chang, E. Y., & Yang, X. (2017). The "flipped classroom" approach: Stimulating positive learning attitudes and improving mastery of histology among medical students. *Anatomical Sciences Education*, *10*(4), 317–327. https://doi.org/10.1002/ase.1664

Cheung, K. K. C., & Winterbottom, M. (2021). Exploring students' visualisation competence with photomicrographs of villi. *International Journal of Science Education*, *43*(14), 2290–2315. https://doi.org/10.1080/09500693.2021.1959958

Chua, Y. P. (2020). *Mastering research statistics* (2nd ed.). McGraw-Hill Education (Malaysia).

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*(1), 37–46.

Cromley, J. G., Perez, T., Kaplan, A., Dai, T., Mara, K., & Balsai, M. J. (2020). Combined cognitive-motivational modules delivered via an LMS increase undergraduate biology grades. *Technology, Mind, and Behavior*, *1*(2), 1–19. https://doi.org/10.1037/tmb0000020.supp

Dickerson, J., & Kubasko, D. (2007). Digital microscopes: Enhancing collaboration and engagement in science classrooms with information technologies. *Contemporary Issues in Technology and Teacher Education*, *7*(4), 279–292.

Donnelly, A. D., Mukherjee, M. S., Lyden, E. R., & Radio, S. J. (2012). Virtual microscopy in cytotechnology education: Application of knowledge from virtual to glass. *CytoJournal*, *9*(1). https://doi.org/10.4103/1742-6413.95827

Fatimah Mohamed, Tan, S. W., & Noor, N. N. M. (2011). Observing and sketching skills in plant anatomy practical class. *Jurnal Sains Dan Matematik*, *3*(2), 66–73.

García, M., Victory, N., Navarro-Sempere, A., & Segovia, Y. (2019). Students' views on difficulties in learning histology. *Anatomical Sciences Education*, *12*(5), 541–549. https://doi.org/10.1002/ase.1838

Giannakas, F., Kambourakis, G., Papasalouros, A., & Gritzalis, S. (2018). A critical review of 13 years of mobile game-based learning. *Educational Technology Research and Development*, *66*(2), 341–384. https://doi.org/10.1007/s11423-017-9552-z

Hamidah Yusof, Jamal Yunus, & Khalip Musa. (2015). *Kaedah penyelidikan : Pengurusan pendidikan*. Penerbit Universiti Pendidikan Sultan Idris.

Hande, A. H., Lohe, V. K., Chaudhary, M. S., Gawande, M. N., Patil, S. K., & Zade, P. R. (2017). Impact of virtual microscopy with conventional microscopy on student learning in dental histology. *Dental Research Journal*, *14*(2), 111–116. https://doi.org/10.4103/1735-3327.205788

Hariyanto, D., Triyono, M. B., & Köhler, T. (2020). Usability evaluation of personalised adaptive e-learning system using USE questionnaire. *Knowledge Management and E-Learning*, *12*(1), 85–105. https://doi.org/10.34105/j.kmel.2020.12.005

Helle, L., Nivala, M., Kronqvist, P., Gegenfurtner, A., Björk, P., & Säljö, R. (2011). Traditional microscopy instruction versus process-oriented virtual microscopy instruction: A naturalistic experiment with control group. *Diagnostic Pathology*, *6*(8), 1–9. https://doi.org/10.1186/1746-1596-6-S1-S8

Keller, J. M. (2010). *Motivational design for learning and performance: The ARCS model approach*. Springer.

Keller, J. M., Ucar, H., & Kumtepe, A. T. (2020). Development and validation of a scale to measure volition for

learning. *Open Praxis*, *12*(2), 161–174. https://doi.org/10.5944/openpraxis.12.2.1082
Krippendorf, B. B., & Lough, J. (2005). Complete and rapid switch from light microscopy to virtual microscopy for teaching medical histology. *Anatomical Record - Part B New Anatomist*, *285*(1), 19–25. https://doi.org/10.1002/ar.b.20066
Kumar, R. K., Velan, G. M., Korell, S. O., Kandara, M., Dee, F. R., & Wakefield, D. (2004). Virtual microscopy for learning and assessment in pathology. *Journal of Pathology*, *204*(5), 613–618. https://doi.org/10.1002/path.1658
Lavasani, M. G., Mirhosseini, F. S., Hejazi, E., & Davoodi, M. (2011). The effect of self-regulation learning strategies training on the academic motivation and self-efficacy. *Procedia - Social and Behavioral Sciences*, *29*, 627–632. https://doi.org/10.1016/j.sbspro.2011.11.285
Lockee, B. B. (2021). Online education in the post-COVID era. *Nature Electronics*, *4*, 5–6. https://doi.org/10.1038/s41928-020-00534-0
Loorbach, N., Peters, O., Karreman, J., & Steehouder, M. (2015). Validation of the Instructional Materials Motivation Survey (IMMS) in a self-directed instructional setting aimed at working with technology. *British Journal of Educational Technology*, *46*(1), 204–218. https://doi.org/10.1111/bjet.12138
Lund, A. M. (2001). Measuring usability with the USE questionnaire. *Usability Interface*, *8*(2), 3–6.
Manganello, F., Falsetti, C., & Leo, T. (2019). Self-regulated learning for web-enhanced control engineering education. *Educational Technology and Society*, *22*(1), 44–58.
Mojica, E. R. E., & Upmacis, R. K. (2022). Challenges encountered and students' reactions to practices utilised in a general chemistry laboratory course during the COVID-19 pandemic. *Journal of Chemical Education*, *99*(2), 1053–1059. https://doi.org/10.1021/acs.jchemed.1c00838
Nauhria, S., & Hangfu, L. (2019). Virtual microscopy enhances the reliability and validity in histopathology curriculum: Practical guidelines. *MedEdPublish*, *8*, 28. https://doi.org/10.15694/mep.2019.000028.2
Neel, J. A., Grindem, C. B., & Bristol, D. G. (2007). Introduction and evaluation of virtual microscopy in teaching veterinary cytopathology. *Journal of Veterinary Medical Education*, *34*(4), 437–444. https://doi.org/10.3138/jvme.34.4.437
Ordi, O., Bombí, J. A., Martínez, A., Ramírez, J., Alòs, L., Saco, A., Ribalta, T., Fernández, P. L., Campo, E., & Ordi, J. (2015). Virtual microscopy in the undergraduate teaching of pathology. *Journal of Pathology Informatics*, *6*(1), 1–6. https://doi.org/10.4103/2153-3539.150246
Polit, D. F., Beck, C. T., & Owen, S. V. (2007). Focus on research methods: Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. *Research in Nursing & Health*, *30*, 459–467. https://doi.org/10.1002/nur.20199
Raja, S. (2010). Virtual microscopy as a teaching tool adjuvant to traditional microscopy. *Medical Education*, *44*(11), 1126. https://doi.org/10.1111/j.1365-2923.2010.03841.x
Rodrigues, I. B., Adachi, J. D., Beattie, K. A., & MacDermid, J. C. (2017). Development and validation of a new tool to measure the facilitators, barriers and preferences to exercise in people with osteoporosis. *BMC Musculoskeletal Disorders*, *18*(1), 1–9. https://doi.org/10.1186/s12891-017-1914-5
Saco, A., Bombi, J. A., Garcia, A., Ramírez, J., & Ordi, J. (2016). Current status of whole-slide imaging in education. *Pathobiology*, *83*(2–3), 79–88. https://doi.org/10.1159/000442391
Shi, J., Mo, X., & Sun, Z. (2012). Content validity index in scale development. *Journal of Central South University (Medical Sciences)*, *37*(2), 152–155. https://doi.org/10.3969/J.ISSN.1672-7347.2012.02.007
Simok, A. A., Hadie@Haji, S. N. H., Abdul Manan@Sulong, H., Yusoff, M. S. B., Mohd Noor, N. F., Asari, M. A., & Kasim, F. (2019). The impact of virtual microscopy on medical students' intrinsic motivation. *Education in Medicine Journal*, *14*(4), 47–59. https://doi.org/10.21315/eimj2019.11.4.5
Tauber, Z., Cizkova, K., Lichnovska, R., Lacey, H., Erdosova, B., Zizka, R., & Kamarad, V. (2019). Evaluation of the effectiveness of the presentation of virtual histology slides by students during classes. Are there any differences in approach between dentistry and general medicine students? *European Journal of Dental Education*, *23*(2), 119–126. https://doi.org/10.1111/eje.12410
Teoh, C. Z., Lee, T. T., Muhamad Ikhwan Mat Saad, Sagat, M. P., Nur Aqilah Jamaludin Khir, & Rajagopal, T. (2022). HistoGuide mobile application (virtual microscopy and slides): A quantitative usability pilot study. *Journal of ICT in Education*, *9*(2), 162–174. https://doi.org/10.37134/jictie.vol9.2.12.2022
Teoh, C. Z., Mindar Singh, N. K., Siti Nor Badariah, & Marzuki Mohd Uzi Dollah. (2022). Using HistoGuide mobile application (virtual microscopy): A qualitative pilot study on usability and sixth form students' learning experience. *EDUCATUM Journal of Science, Mathematics and Technology*, *9*(2), 136–154. https://doi.org/10.37134/ejsmt.vol9.2.15.2022
Teoh, C. Z., Muhamad Ikhwan Mat Saad, & Che Nidzam Che Ahmad. (2022). Integrating technology-mediated learning in biology education (histology): A systematic literature review. *Journal of ICT in Education*, *9*(1), 86–99. https://doi.org/10.37134/jictie.vol9.1.8.2022
Triola, M. M., & Holloway, W. J. (2011). Enhanced virtual microscopy for collaborative education. *BMC Medical Education*, *11*(1), 9–12. https://doi.org/10.1186/1472-6920-11-4
van Wyk, N. (2019). *A proposed framework for developing online courses: The case for a VR course*. University of Cape Town.

van Wyk, N., Johnston, K., Moeller, K., & Haas, F. (2020). Developing an IT course for emerging technologies using a framework – An example of an IoT Course V1.0. *Proceedings of the 2020 InSITE Conference*, 015–045. https://doi.org/10.28945/4521

Xu, C. J. (2013). Is virtual microscopy really better for histology teaching? *Anatomical Sciences Education*, *6*(2), 138. https://doi.org/10.1002/ase.1337

Zamanzadeh, V., Ghahramanian, A., Rassouli, M., Abbaszadeh, A., Alavi-Majd, H., & Nikanfar, A.-R. (2015). Design and implementation content validity study: Development of an instrument for measuring patient-centered communication. *Journal of Caring Sciences*, *4*(2), 165–178. https://doi.org/10.15171/jcs.2015.017