Intelligent Human Systems Integration (IHSI 2024), Vol. 119, 2024, 437–445
https://doi.org/10.54941/ahfe1004508

The Effect of Visual and Auditory Channel Interaction on Cognitive Performance in Online Education

Yingliang Sun, Wenyu Wu, Ke Zhu, Hao Huang, Jiayu Guo, and Yilian li

Southeast University, Nanjing, Jiangsu 210000, China

ABSTRACT

With the arrival of the multimedia era, online education has become one of the main teaching modes. At the same time, the presentation of learning materials has become more diversified, and the interaction mode has gradually developed in the direction of multimodality. According to multichannel cognitive theory and multimedia cognitive theory, human can process information from different channels at the same time, but due to limited cognitive resources, inappropriate content presentation and interaction methods can cause extra load, which will affect learning efficiency. In this paper, we design an experiment to investigate the effects of the visual and auditory channels, as well as the audio-visual channel on people's learning efficiency by presenting and delivering learning information in different ways. The result shows that learning through audio-visual channel is the most efficient interaction method, and learning through a single auditory channel is the least efficient method.

Keywords: Online education, Human-computer interaction, Cognitive performance, Multichannel interaction

BACKGROUND

With the development of digitalisation and intelligence, the online education model has gradually emerged and has been widely popularised since the epidemic era. Online education makes use of the Internet, combining massive educational resources and multimodal interaction, breaking the time and space barriers of traditional education while greatly improving the interest in education and learning efficiency. Multimedia teaching has gradually become the basic form of teaching with the advantages of intuition, vividness and efficiency, and online education is more dependent on this. Slides and their accompanying text, pictures, animations, as well as video and audio are the main information transfer media for online teaching, and the visual and auditory channels are the main channels for information transfer. However, due to the limited cognitive resources of learners and the difference in cognitive efficiency through different channels, the choice of information transfer channels and the design of multimedia learning methods affect the learning efficiency of learners to a certain extent.

Multichannel Cognitive Theory in Online Education

People acquire external information through vision, hearing, touch, taste, and smell. Multi-channel interaction means using multiple sensory channels for information input and output at the same time, so that human-computer interaction can be carried out in a more natural, efficient, and collaborative way, and user experience can be enhanced (Cohen and McGee, 2004). Among them, vision and hearing are the main sources of receiving information, and are also the main channels of information transfer in remote education. Therefore, studying the effects of visual and auditory channels as well as audio-visual composite channel on people's cognitive performance in the process of online education is one of the ways to effectively improve learning efficiency.

Vision is the main way for people to acquire information, and more than 80 % of information comes from it (Matsubara and Nagamachi, 1997). Hearing is the second largest information acquisition channel after vision and is the medium through which sound information is received. However, visual and auditory channels have their own disadvantages. The human eye has a limited visual range, and it will be further reduced in the case of concentration. When the density of information is too large or there are dynamic changes, the human cognitive overload and the rate and accuracy of information reception will also be reduced. The auditory channel can break the spatial limitation, but relative to the visual channel, the reception efficiency is lower and is susceptible to environmental noise (Li, 2022). The combination of visual and auditory channels can make up for the disadvantages of a single channel and improve the accuracy and comprehensiveness of information reception. In addition, the audio-visual mode of information is closer to the human information reception mode in the real state. It can improve the vividness and intuition of the learning content, as well as the naturalness and immersion of the learning process. On the other hand, a multichannel approach increases the amount of information transmission. From the viewpoint of cognitive load theory, human cognitive capacity is limited, and can only process quantitative information per unit of time, unable to deal with excess information, which may cause cognitive overload, interfering with one's understanding and processing of other information (Mayer, 1999). Therefore, when making the choice of interaction channels for online teaching, it is necessary to consider the advantages and disadvantages of different channels as well as the information processing capacity of human beings.

Cognitive Theory of Multimedia Learning

Mayer has conducted research on multimedia learning since the 1980s in an attempt to find a more scientific and efficient learning method (Xiao, 2023). His dual-channel assumption, limited capacity assumption and active processing assumption are the cornerstones of the cognitive theory of multimedia learning, and based on the three assumptions, Mayer constructed seven basic principles of multimedia instructional design (Mayer, 2009).

The dual-channel assumption refers to the fact that when people are engaged in cognitive processing, there are corresponding information processing channels for both visual and auditory information. The limited capacity assumption refers to the fact that people consume cognitive resources when performing cognitive processing, and cognitive resources are limited, so the amount of information that can be processed at one time in each information processing channel is also limited. The active processing assumption refers to the fact that people actively select learning materials, conduct cognitive processing on them with the internal mental model, and finally integrate new knowledge with old knowledge to establish a consistent mental representation. In multimedia learning, firstly, the multimedia materials in the form of text and images go into the human sensory memory centre through the auditory channel and the visual channel respectively, and then the learners select, organise and integrate the information from the sensory memory, which is the cognitive processing process of multimedia learning.

According to the cognitive load theory, cognitive load is divided into intrinsic cognitive load, extraneous cognitive load and germane cognitive load (Low, 2009). Among them, intrinsic cognitive load refers to the amount of cognitive resources consumed in the process of receiving learning knowledge, which is determined by the learning content and the learners' own knowledge and experience; extraneous cognitive load refers to the additional cognitive load, which is determined by the organisation of the information and the way of presenting it, and the inappropriate design of the teaching method will increase the cognitive difficulty as well as the non-essential cognitive load; germane cognitive load refers to the additional cognitive load which is beneficial for the learners to understand and remember, which is related to the construction of mental schema and the degree of subjective engagement, and can affect the selection, organisation and integration in learners' cognitive process. Multimedia cognitive theory suggests that instructional design should minimise the extraneous cognitive load and increase the germane cognitive load, and the presentation of information and the choice of interaction channels will directly affect the production of both. An excellent teaching approach will reduce the extra cognitive load and improve students' cognitive performance (Xu, 2019).

EXPERIMENTAL DESIGN AND IMPLEMENTATION

The purpose of this experiment is to investigate the effect of information transfer channels on learners' cognitive performance in the online teaching process. In this experiment, slides are selected for presenting information, three different interaction modes of visual, auditory, and audio-visual channels are selected as experimental variables, and the correctness rate of the subjects' retelling of the teaching content is used as an assessment index.

Experimental Purpose

1. Explore the effects of visual and auditory channels as well as audio-visual channels on learners' cognitive performance and analyse the differences in cognitive performance between different channel approaches.

2. To find out the most efficient channel for information transfer and provide guidance for instructional design.

Participants

Ninety volunteers aged between 14 and 35 were randomly selected as experimental subjects, 45 of each gender. The volunteers had normal cognitive and expressive abilities, and the volunteers were not aware of the learning content of the task beforehand, so the possibility that the results of the experiment would be affected by the fact that they were informed in advance was ruled out.

Experimental Materials

Three teaching slides. The slides show the same points in different ways, i.e., the characteristics of the Rococo style: "Rococo art is an art form or style of art that emerged in France in the 18th century and spread throughout Europe. The general characteristics of Rococo are ornate, elaborate, delicate, often asymmetrical, with a preference for curved and S-shaped lines, naturalistic decorations, and bright colours". The key words are: 18th century, emerged in France, throughout Europe, ornate, elaborate, delicate, asymmetrical, curved, S-shaped lines, naturalistic, brightly coloured. The first slide contains a picture with text (shown in Figure 1); the second slide does not contain text and conveys information to the participants only by means of audio (shown in Figure 2); the third slide conveys visual information by means of text and audio information by means of a player (shown in Figure 3). All three slides are presented for the same time, which is the time required for the audio to be played once, i.e., 18 seconds.

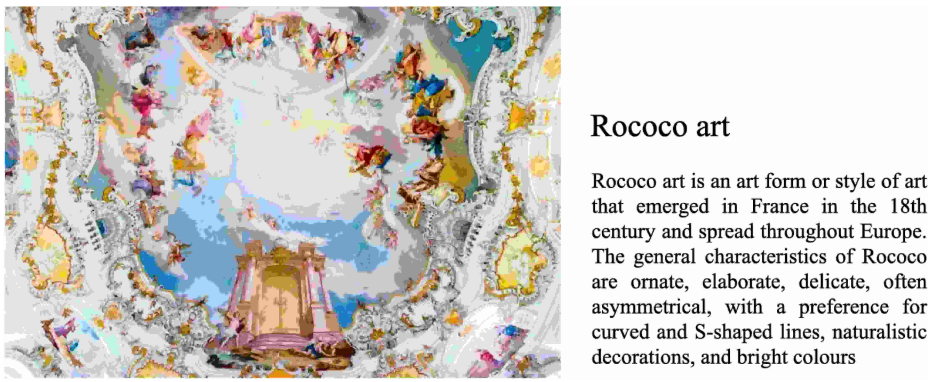

Figure 1: Slide with only image and text.

Experimental Procedure

1. The subjects were randomly divided into three groups, a, b, and c, according to the proportion of men and women, assigned different tasks, and numbered 1-30.

2. Slides with text only, slides with audio only, and slides with both text and audio were numbered 1, 2, and 3 respectively, and assigned to subjects in groups a, b, and c for testing.
3. The experimental subjects watched the slides and tried their best to remember what was presented in the slides.
4. The subjects retold the teaching content, and then we evaluated their mastery of the learning content by comparing the overlap between the keywords in the retold content and the slides.
5. The experimental results were analysed based on data, and conclusions were drawn.

Figure 2: Slide with only image and audio content.

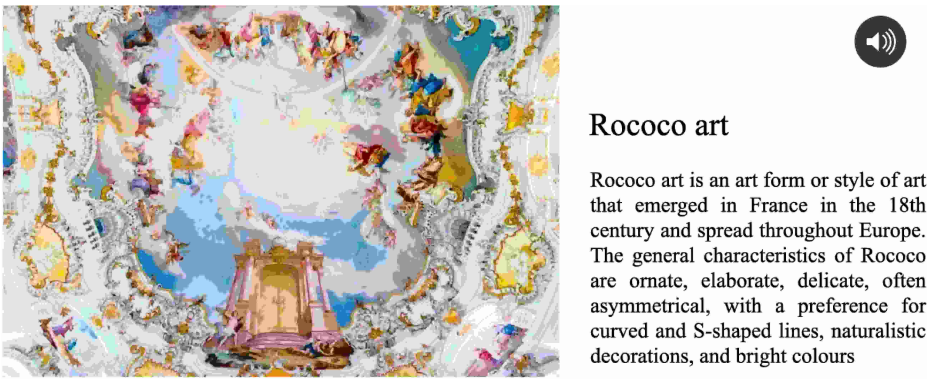

Figure 3: Slide with an image, text and audio content.

Experimental Results

Based on the results of the experiment, we obtained the retellings of 90 individuals, calculated the percentage of keyword matches between these retellings and the materials, and obtained the following table.

Table 1. Matching rate between keywords in the retelling and the source material.

Group Number	a	b	c
1	63.60%	45.50%	63.60%
2	45.50%	27.30%	56.50%
3	36.40%	36.40%	72.70%
4	54.50%	63.60%	45.50%
5	63.60%	45.50%	63.60%
6	45.50%	63.60%	45.50%
7	54.50%	27.30%	36.40%
8	36.40%	36.40%	54.50%
9	63.60%	54.50%	27.30%
10	36.40%	27.30%	63.60%
11	72.70%	54.50%	90.10%
12	63.60%	45.50%	72.70%
13	27.30%	18.20%	27.30%
14	36.40%	45.50%	54.50%
15	45.50%	72.70%	45.50%
16	45.50%	18.20%	36.40%
17	72.70%	63.60%	63.60%
18	63.60%	54.50%	45.50%
19	36.40%	27.30%	63.60%
20	27.30%	45.50%	54.50%
21	54.50%	36.40%	63.60%
22	81.80%	54.50%	72.70%
23	54.50%	27.30%	54.50%
24	72.70%	36.40%	81.80%
25	27.30%	36.40%	45.50%
26	72.70%	81.80%	36.40%
27	63.60%	45.50%	91.00%
28	54.50%	27.30%	63.60%
29	27.30%	45.50%	54.50%
30	45.50%	18.20%	72.70%

DATA ANALYSIS

SPSS Analysis

Firstly, the data were tested for normality and the results were as follows:

Table 2. Tests of normality.

		Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Sensory Channel	Statistic	df	Sig.	Statistic	df	Sig.
Matching rate	Visual	.144	30	.117	.939	30	.088
	Auditory	.133	30	.185	.951	30	.179
	Audio-visual	.118	30	.200*	.965	30	.417

*. This is a lower bound of the true significance.
a. Lilliefors Significance Correction

According to the data in the above table, cognitive performance is significant (Sig>0.05) under different channels of information transfer, which is consistent with normal distribution and can be tested parametrically.

Descriptive statistics were performed to obtain the mean and standard deviation, and the results were as follows:

Table 3. Descriptive statistics.

Dependent Variable: Matching Rate			
Sensory Channel	**Mean**	**Std. Deviation**	**N**
Visual	0.515133	0.1586790	30
Auditory	0.427400	0.1637196	30
Audio-visual	0.573067	0.1648899	30
Total	0.505200	0.1715332	90

From the results, it can be seen that the average matching rate is 51.5133% (SD = 15.86790%) when only visual information is provided, 42.7400% (SD = 16.37196%) when only auditory information is provided, and 57.3067% (SD = 16.48899%) when both visual and auditory information are provided. The highest average matching rate is found under the audio-visual channel, followed by the visual channel, and the lowest average matching rate is found under the auditory channel.

Afterwards, the results of the above experiments were tested by one-way ANOVA, from which the significance was determined. Before that, the data needs to be verified for homogeneity of variance, and the results are as follows:

Table 4. Tests of homogeneity of variances.

		Levene Statistic	**df1**	**df2**	**Sig.**
Matching rate	Based on Mean	.015	2	87	.985
	Based on Median	.008	2	87	.992
	Based on Median and with adjusted df	.008	2	85.933	.992
	Based on trimmed mean	0.016	2	87	0.984

As can be seen from the above table, Sig>0.05, which is significant, i.e. the variances are homogeneous and suitable for ANOVA, after which the data were subjected to one-way ANOVA and the results were as follows:

Based on the data, it was found that Sig<0.05, i.e. different information transfer channels have a significant effect on the matching rate.

Multiple comparisons showed that there were significant differences in matching rates between the visual and auditory channels, and between the auditory and audio-visual channels, while there were no significant differences in matching rates between the visual and audio-visual channels.

Table 5. ANOVA.

Matching rate					
	Sum of Squares	**df**	**Mean Square**	**F**	**Sig.**
Between Groups	.323	2	.161	6.114	.003
Within Groups	2.296	87	.026		
Total	2.619	89			

Table 6. Multiple Comparisons.

Dependent Variable: Matching Rate							
	(I) Sensory channel	**(J) Sensory channel**	**Mean Difference (I-J)**	**Std. Error**	**Sig.**	**95% Confidence Interval**	
						Lower Bound	**Upper Bound**
LSD	1	2	.0877333*	.0419449	.039	.004363	.171103
		3	−.0579333	.0419449	.171	−.141303	.025437
	2	1	−.0877333*	.0419449	.039	−.171103	−.004363
		3	−.1456667*	.0419449	<.001	−.229037	−.062297
	3	1	.0579333	.0419449	.171	−.025437	.141303
		2	.1456667*	.0419449	<.001	.062297	.229037

CONCLUSION

Based on the multichannel cognitive theory and the multimedia cognitive theory, this paper explores the effects of different forms of information delivery on learners' cognitive performance during online teaching. Three different slides were used as the experimental materials; three groups of subjects were randomly selected to participate in the experiment; the matching rate between the content that the learners retold after the information was presented and the original content was used as the evaluation index; and finally the experimental results were analysed by using the SPSS software.

Based on the results of the data analysis, it can be inferred that the information presentation and delivery channels have a significant impact on the quality of teaching and learning. When both visual and auditory information are provided, learners have the best learning outcomes and cognitive performance; when only visual information is provided, it is the second best; and when only auditory information is provided, learners have the worst learning outcomes and cognitive performance. Of these, the difference between the effects of the visual and audio-visual channels on cognitive performance is small, and the difference between the effects of the visual and auditory channels, the audio-visual and auditory channels on cognitive performance is large. Therefore, the involvement of the visual channel greatly favours cognitive performance, while the audio-visual channel is the best choice. When designing educational materials in the future, designers should consider using visual or audio-visual channel in the design of educational materials, and try to avoid using only the auditory channel for content presentation.

REFERENCES

Cohen, P. R. and McGee, D. R. (2004). Tangible multimodal interfaces for safety-critical applications. Communications of the ACM, 47(1), p. 41. doi: https://doi.org/10.1145/962081.962103.

Li, Y. (2022). Research on multichannel interaction of virtual reality online education based on user cognitive load theory. doi: https://doi.org/10.27014/d.cnki.gdnau.2022.000927.

Low, R. (2009). 'Cognitive Architecture and instructional design in a multimedia context', Cognitive Effects of Multimedia Learning, pp. 1–16. doi:10.4018/978-1-60566-158-2.ch001.

Matsubara, Y. and Nagamachi, M. (1997). Hybrid Kansei engineering system and design support. International Journal of Industrial Ergonomics, 19(2), pp. 81–92. doi: https://doi.org/10.1016/s0169-8141(96)00005-4.

Mayer, R. E. (1999). The Promise of Educational Psychology: Learning in the content areas. Upper Saddle River, NJ: Merrill, Prentice-Hall.

Mayer, R. E. (2009). Multimedia Learning. Cambridge University Press.

Xiao, L. (2023). A study of Chinese teaching strategies in middle schools based on Meyer's multimedia learning theory. doi: https://doi.org/10.27280/d.cnki.gsdsu.2023.000380.

Xu, Y. (2019). Research on the Situation Teaching of Junior High School Chinese Class Based on VR Resources. doi: https://doi.org/0.27356/d.cnki.gtjdu.2019.000902.