Schema and emotion in memory retrieval following video-based learning: An artificial intelligence study

Anjana Junius Vidanaralage, Anuja Thimali Dharmaratne
School of Information Technology, Monash University

Shamsul Haque
Department of Psychology, Jeffrey Cheah School of Medicine and Health Sciences, Monash University

Adapting innovative educational technologies to bolster students' academic learning is increasing rapidly. This study explored schema congruent and incongruent participants behaviour when experiencing video-based materials as the medium of learning within the frame of a flipped learning environment. The participants watched an educational learning video on a given topic and completed memory retention tests in different time variations: immediate and delayed. Additionally, an artificial intelligence-based emotion analysis examined the emotional valency of participants during two phases: study phase and test phase. The experiment comprised 16 healthy young adult volunteers (8 schema congruent, 8 schema incongruent; 9 males [56.25%], 7 females [43.75%]; age range 20–34 years, mean age 27.31 years, SD = 2.87 years). A combination of statistics-based and AI-based analysis evaluated the effectiveness of video-based learning in terms of retrieval accuracy, response time and emotional valence. The findings indicate that retrieval accuracy for the schema incongruent group was better than schema congruent. Response time for schema congruent group was quicker than schema incongruent. Both groups exhibited more negative emotions during the study phase but more positive emotions during the test phase.

Implications for practice or policy:

- Acceptance testing of video-based learning in tertiary education for different schema groups of students by assessing their emotional state helps educators to enhance pedagogy.
- Nourishing positive learning experiences from videos and questionnaires should be the goal, considered at the design stage for courses that rely on video-based materials.
- Adaptation of video-based learning strategy is more instructionally efficient and scalable for academic institutions and educators during a pandemic situation.

Keywords: artificial intelligence, education technology, emotion recognition, memory recall, schema theory, video-based learning

Introduction

Technology has heavily influenced education and learning in recent decades (R. Huang et al., 2019). The exponential growth of technology usage in education resulted in a proliferation of pedagogical software learning tools known as digital learning tools (SooHwan et al., 2013). These digital learning tools support the concept known as educational technology, which is also termed as instructional design and technology (Lowenthal & Wilson, 2010). Advances in technology are driving instructors towards the use of education technology applications (Kolekar et al., 2018) and they are contributing to significant changes in students' learning experience (Sloan & Lewis, 2014). Technology-based instruction has evolved from early uses of simple audiovisual aids to complex forms of e-learning approaches (Halawi et al., 2009). The success of e-learning from the 20th century has led to the emergence of new technology-based digital learning strategies that shape learning and instruction (R. Huang et al., 2019). Learning is a natural ongoing process; the initial phase of learning is mastering and memorising basic facts (Riedesel & Charles, 2018). Education is often regarded as synonymous with learning as the acquired experience (Sampath, 1981).

Studies have found that differences in prior knowledge and training can impact learning outcomes (Anderson, 1981; Jonassen & Grabowski, 1993; Rumelhart, 2017; van Kesteren et al., 2012). This phenomenon has been explained in terms of schema or what is known as a pre-existing knowledge base (van Kesteren et al., 2012). Schema is defined as the abstract knowledge structures (Ghosh & Gilboa, 2014) that facilitate what we learn and remember. The knowledge we learn (e.g., information, event memory) could be schema congruent (consistent) and schema incongruent (inconsistent). Research has shown that schema congruent and schema incongruent event memories are processed and recalled better than schema neutral memories (Greve et al., 2019). Moreover, memory can also be modulated by emotion (LaBar & Cabeza, 2006; Righi et al., 2012) and emotions have a substantial influence on the cognitive processes of humans, including learning and memory (Phelps, 2004; Um et al., 2012). Emotions drive attention (Vuilleumier, 2005), influencing learning motivation and facilitating active encoding in memory (Pekrun, 1992; Seli et al., 2016).

Rationale of the study

During the recent COVID-19 pandemic, digital learning reached a whole new level as the default learning strategy (Sá & Serpa, 2020). Despite the benefits digital learning provided when deployed on a smaller scale in the past, the rapid adoption during the pandemic failed to improve teaching practices and even created a negative mindset among students towards video-based learning (Lin & Nguyen, 2021). Moreover, according to Inan et al.'s (2010) study, one primary reason for the failure in digital learning is that instructors used technology as a content delivery platform rather than as an actual learning tool.

Video-based learning is not envisioned to substitute traditional methods of learning or teaching. The goal is to be an enabler supporting the cognitive learning process by creating opportunities in flipped learning environments that help develop skills better (Cervi et al., 2013). Students in a video-based learning environment are naturally more reliant on watching learning videos than in a traditional face-to-face learning setting (Myllymäki et al., 2017). Therefore, it is worthwhile to explore the acceptance of video-based learning and investigate viewing behaviours in terms of prior knowledge (i.e., schema theory) and their affective states (i.e., emotions).

This research aimed at a statistics-based and AI-based analysis to explore the behaviour of two focus groups of participants (schema congruent and schema incongruent) when experiencing video-based materials as the medium of learning, within the frame of a flipped learning session which many academic institutions follow during the pandemic. The criteria for evaluating the effectiveness of video-based learning were (a) retrieval accuracy, (b) response time and (c) emotional valence (i.e., degree of positivity or negativity). The above synopsis has been employed to formulate three research questions for this study:

(1) Is there a measurable difference in the response time between schema congruent and schema incongruent participants during a video-based learning session?
(2) Is video-based based learning effective for enhancing adult learning in the short term and the long term?
(3) What is the impact of positive and negative emotions in study and test phases during a video-based learning session?

Background study

Interactions between cognitive psychology, affective science and education technology are the cornerstones underpinning this multidisciplinary research. It evaluates a concept (cognitive psychology: *schema theory*, affective science: *six basic emotions* and education technology: *flipped learning*) within the three cornerstone domains. Figure 1 illustrates each research domain and its relevant concept studied.

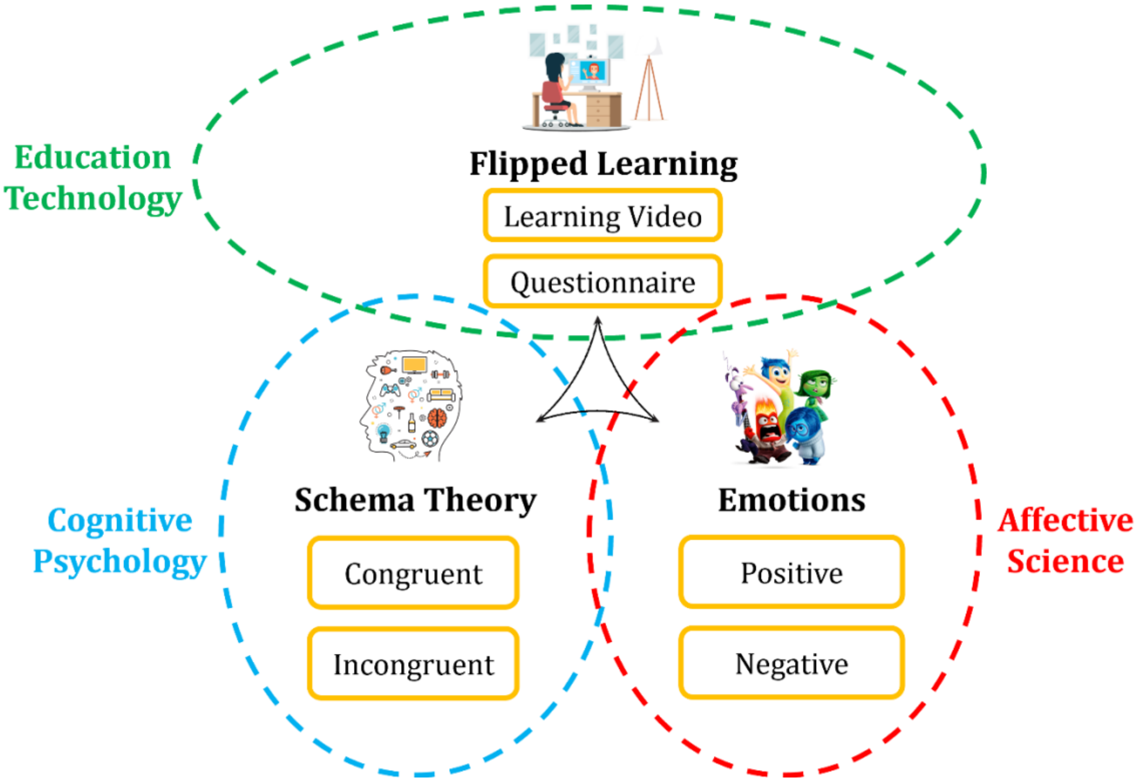


Figure 1. Overview of the research study

Schema theory
The schema concept can be traced back to Plato (428–348 B.C.) and Aristotle (384–322 B.C.) (Kristjánsson, 2016; Laura & Olivier, 2020), but Kant (1899) is generally considered as the first person to talk about schema. Although the term schema has roots in Kant's philosophical work, schema did not gain cognitive research attention until Bartlett (1932) used the term schema and conducted experiments to explore the effects of schema in memory construction. This work was the most widely cited by schema theorists working on the cognitive domain (Saito, 1996). A schema (plural: schemas or schemata) is an organised mental structure of knowledge stored in memory (Fleming, 1987; Gagne, 1986; Richey et al., 2010; Winn, 2004). In other words, schema interprets how we see the external world with internal metal structures (i.e., experience) (Johnson, 1987). Piaget (1952) improved the schema explanation as a continuous interaction in which the individual either absorbs new experience congruent to existing schema or adjust schema to fit in the incongruent experience. The standard information congruency continuum distinguishes two extremes:

- Schema congruent or consistent is an event that conforms to an individual's expectations or the world of knowledge (Greve et al., 2019). Schema consistent information is embodied in long-term memory (Bartlett, 1932; Piaget, 1952) and processed in the brain against an individual's existing prior knowledge. Information retrieval for a congruent event facilitates inferential processing, where the brain finds a similarity from long-term memory and fits it with existing knowledge (Anderson, 1981). This is known as the consistency effect or congruency effect.
- Schema incongruent or inconsistent is an event that conflicts with an individual's expectations for an unexpected novel stimulus encountered from the surrounding environment (Greve et al., 2019). Schema inconsistent information is embodied as discrete propositions in memory (Sentis & Burnstein, 1979) and undergoes distinctive processing for information retrieval, an elaborative form of processing where the brain notices a difference against existing knowledge (Prull, 2015). This asymmetrical mapping of unexpected information over expected information is known as the inconsistency effect or incongruency effect.

Schemas are believed to be the "building blocks of cognition" (Rumelhart, 2017, p. 33); hence schema theory has played a significant role in the history of learning, design and technology domain (i.e., educational technology). Since the 1970s, scholars have applied schema theory to external cognitive strategies (Winn, 2004) to help learners develop schemata when they encounter new information. However, it is challenging to assess improvements on intangible cognitive functions, such as learning and memory recall (Lu, 2014), and explicitly pinpoint how or when technology enhanced the learning experience (Bolton et al., 2008).

Emotions

The study of emotions also has historical roots in Greek philosophy, similar to the schema concept. Plato mentioned emotions as passions or affections (Laura & Olivier, 2020), but it was Aristotle who conducted the first analysis and presented that emotions arise from what we perceive and think and that must be controlled (Kristjánsson, 2016). In fact, both philosophers considered emotions as a disturbance to cognition (Dalgleish & Power, 2000). Darwin (1872) pioneered the first scientific study about emotions and historically dominated as the founder of the science of emotions (Leff, 1973). An emotion (plural: emotions) is defined by Vikan (2017, p. 3) as a "coordination or composite of experience, behavioural expressions and physiological / neurological components with varying duration". Darwin's ideas were related to his notable theory of natural selection, where expression and emotion are equal in both humans and animals and mainly function as a communication medium. Darwin's scientific view between expression and emotion is currently connected to facial expressions in modern research and theories.

Several researchers have tested Darwin's domineering concept of the universality of emotions involving the basic emotions. In this perspective, *basic* does not mean that emotions are discrete or independent (Izard, 1971, 2013). The most delicately formulated study (in support of Darwin) was conducted by Ekman (1973) and considered as the classical study in modern research on emotions. Currently, the following six emotions have the highest level of consensus on the list of basic emotions: (a) happy, (b) sad, (c) anger, (d) fear, (e) surprise and (f) disgust. Izard, pioneering doing emotion research for more than 40 years, himself had proposed several lists of basic emotions, first starting from 10 emotions (Izard, 1991) and then modifying and proposing a revised list of six emotions (Izard, 2009). Izard (2009) emphasised *interest* (the 7th emotion) as a normative emotion because it serves as an entrance to other emotions via attention. For this reason, there will always be controversies regarding the classification of 6–10 most common emotions to conform to the list of basic emotions. The standard emotional valence continuum distinguishes two extremes:

- Positive emotions: These were historically ignored in research but gained an increased interest in recent years (Vikan, 2017). According to Fredrickson's (1998, 2001) theory, positive emotions function to broaden and build adaptation. Broadening refers to the openness to new perspectives and challenges which affects attention, thinking and behaviour. Building implies the stimulation of consistency and duration. According to the list of basic emotions (Izard, 2009), happy and surprise are categorised as positive emotions.
- Negative emotions: Research has focused more frequently on negative than positive emotions (Vikan, 2017). A prime reason for this is that negative emotions often are problems both to individuals and society (Dawkins, 2000). When individuals experience negative emotions, it leads to discomfort (Vikan, 2017). According to the list of basic emotions (Izard, 2009), sad, anger, fear and disgust are categorised as negative emotions.

The impact of emotions on the learning process shows that positive emotions improve and contribute to academic achievement by facilitating attention; this strongly mediates self-motivation and self-satisfaction (Um et al., 2012). Conversely, studies report that negative emotions also improve learning by coordinating confusion. Confusion is a cognitive disequilibrium state induced by contradicting data. A confused student has an increased and better focus on learning, leading to higher performance (D'Mello et al., 2014). Therefore, the effects of emotional experiences are complicated and ubiquitous; positive emotions do not always benefit learning, and negative emotions do not always impede learning (Hascher, 2010; Pekrun, 2014).

Flipped learning

Advancements in educational technology boosted learning capabilities for creating flexible learning environments at home instead of schools and universities (Betihavas et al., 2016). By exploiting this advantage, recent trends show a surge in advocating flipped learning approaches (Gilboy et al., 2015). In flipped learning pedagogical framework, traditional lectures and assignments are flipped (reversed or inverted) (Arnau et al., 2013). In flipped learning, the learner (i.e., student) gains the necessary knowledge by watching a learning video at home before attending the class. While in class, the activities are committed to discussions centred around explored knowledge (Borah, 2021). Consequently, learners learn more through direct interactions with peers than passively watching an instructor during class (Borah, 2021). However, a study by Findlay-Thompson and Mombourquette (2014) compared traditional face-to-face learning and flipped learning and showed a neutral result with no significant differences in outcomes. Furthermore, Willey and Gardner (2014) stated that learners who poorly perform in flipped learning environments typically lack the agency and self-efficacy to take responsibility for their own learning. Moreover, Willey (2016) suggested learners who struggle in flipped learning possess behaviours such as not asking questions in sessions, being unable to apply learning concepts to problem context and heavily dependent on the traditional procedural learning style.

Experiment

The videoconferencing tool Zoom (https://www.zoom.us/) was used as the primary audio and video acquisition tool. Zoom conveniently conducts confidential live research interviews seamlessly among participants in a remote location (Gray et al., 2020). Additionally, the Quizizz web application (https://quizizz.com/) was used as the multiple-choice questions (MCQ) questionnaire tool for participant response acquisition. This online assessment tool provides a gamified experience (Chaiyo & Nokham, 2017; Göksün & Gürsoy, 2019) to keep participants engaged and motivated while responding to memory retention tests questions of the experiment. Figure 2 illustrates the flow of the experiment. Participants were sent a Zoom email invitation link for the scheduled meeting. This invitation could be synchronised with their electronic calendar (e.g., Gmail, iCal or Outlook) and comprised the date, time and password to the meeting along with the explanatory statement of the study and consent form to be submitted online. Specifically, the email informed that participation was voluntary and that the scheduled Zoom session would be recorded. However, participants were not informed or given any clue about the learning topic (i.e., cryptocurrency).

Ethical clearance was obtained via an online application from the Monash University Human Research Ethics Committee for this research. The committee has categorised this study as a "low-risk" project that meets the requirements of the national statement on ethical conduct in human research and Australian code for the responsible conduct of research and has granted approval – MUHREC Project ID: 22084.

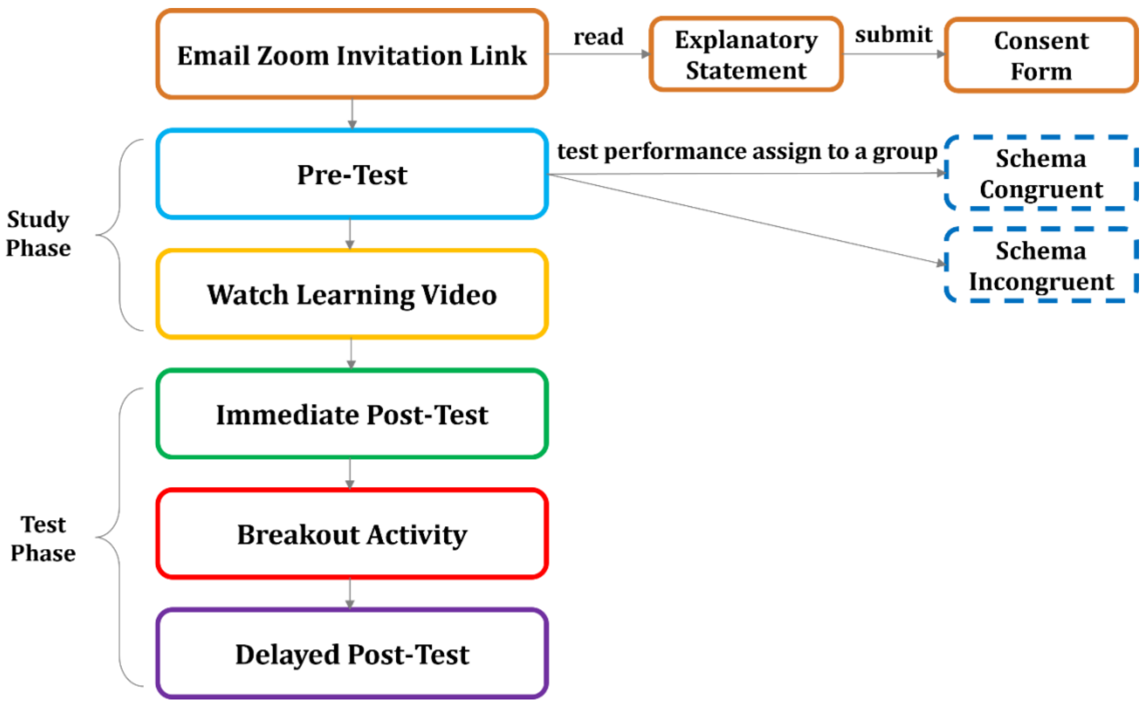


Figure 2. Experiment flow

Participants

The study comprised 16 healthy young adult volunteers (8 schema congruent, 8 schema incongruent; 9 males [56.25%], 7 females [43.75%]; age range 20–34 years, mean age 27.31 years, *SD* = 2.87 years). The sample size was calculated using the *G*Power* statistical power analysis software (Faul, Erdfelder, Buchner et al., 2009; Faul, Erdfelder, Lang et al., 2007) for an effect size of 0.20 (d = 0.20) at a 0.05 level of significance (α = 0.05) with a power of 0.80 (β = 0.80) and 2 groups (congruent, incongruent) and 2 measurements (immediate, delayed).

Inclusion criteria consisted of (a) age between 20 and 34 years; (b) minimum high school college graduate; (c) literacy in English reading, writing and listening skills; (d) information technology literacy in typing, web navigation and videoconferencing skills (e.g., Zoom). Exclusion criteria consisted of (a) failed to complete consent form; (b) self-reported history of learning disability; (c) self-reported history of psychotic disorders (i.e., mental disorders); (d) self-reported alcohol consumption in last 24 hours. Table 1 summarises an overview of the participants' demographics.

Table 1
Participants' demographics

Demographic	***n***	**%**
Age		
20–24 years	2	12.50%
25–29 years	11	68.75%
30–34 years	3	18.75%
Gender		
Male	9	56.25%
Female	7	43.75%

Materials and measures

Traditional learning modes and recalling learned content follow rote memorisation techniques (Thong et al., 2016), which drastically differs in how knowledge is delivered and measured for memory retention in a flipped learning environment.

Stimuli – learning video
Educational videos are extensively used to deliver varied educational content (Moussiades et al., 2019). Since watching a learning video is an integral part of the flipped learning framework, educational design research in recent times tends to focus more on video-based learning experiences (i.e., emotions) exhibited among students (Kolås, 2015). This experiment selected the topic cryptocurrency for the educational video content material. Researchers and domain experts were consulted to produce this material at Monash University. The entire video was in English and covered the theoretical fundamentals of bitcoin, the blockchain platform and cryptocurrency mining. The cryptocurrency learning video did not exceed 9 minutes for maintaining proper participant engagement (Guo et al., 2014).

Measures – questionnaires
Tests are viewed as tools for measuring learners (i.e., students) mastery of skills and knowledge and testing nature is ubiquitous in educational research (Marsh et al., 2007). The two most common ways to measure memory retention are recall test and recognition test (Davis & Moore, 1935). The medium of the memory retention tests is in English and involves two MCQ tests: pre-test and post-test. The pre-test questionnaire comprised 20 questions – 10 on demographic information and 10 on cryptocurrency concepts (refer to Appendix A). Demographic questions were related to age, gender and current level of education; the cryptocurrency questions were designed following a recognition test format that examines participants' pre-existing knowledge about cryptocurrency. Similarly, the post-test also consisted of 20 questions – 10 on recognition and 10 on recall (refer to Appendix B), based on the cryptocurrency learning video. Here, the recall test questions were designed according to Little and Bjork's (2015) suggested format. Each question on both questionnaires had only one correct answer.

Design

The study includes two factors (i.e., independent variables: schema, recall point), each with two levels; hence, it can be denoted as $s^k = 2^2$. The *s* represents the number of factors and *k* represents the number of levels. This experiment with two factors, each with two levels, is designated as 2 (schema: congruent, incongruent) x 2 (recall point: immediate, delayed) mixed factorial design (i.e., spilt-plot design) and tabularised in Table 2.

Table 2
Two-by-two (2x2) mixed factorial design

2 x 2 mixed factorial design (Split-plot)		Schema	
		Congruent	Incongruent
Recall point	Immediate	Immediate + Congruent	Immediate + Incongruent
	Delayed	Delayed + Congruent	Delayed + Incongruent

Independent variables (IV) are factors manipulated in an experiment. Furthermore, the IV schema can be classified as a participant variable that the researcher cannot manipulate (e.g., schema congruent or incongruent characteristic of the participant). Based on pre-test results, when participants reach the criterion (test score ≥ 50% correct), they are assigned to the schema congruent group and if they could not reach the threshold, they are designated to the schema incongruent group (refer to Figure 2). The IV recall point is classified as a stimulus variable which the researcher can manipulate (e.g., the time duration of immediate and delayed recall) (Mandler, 1959). Additionally, the IV schema was a between-subjects factor and the IV recall point was a within-subjects factor; refer to Figure 3 for the experiment design paradigm.

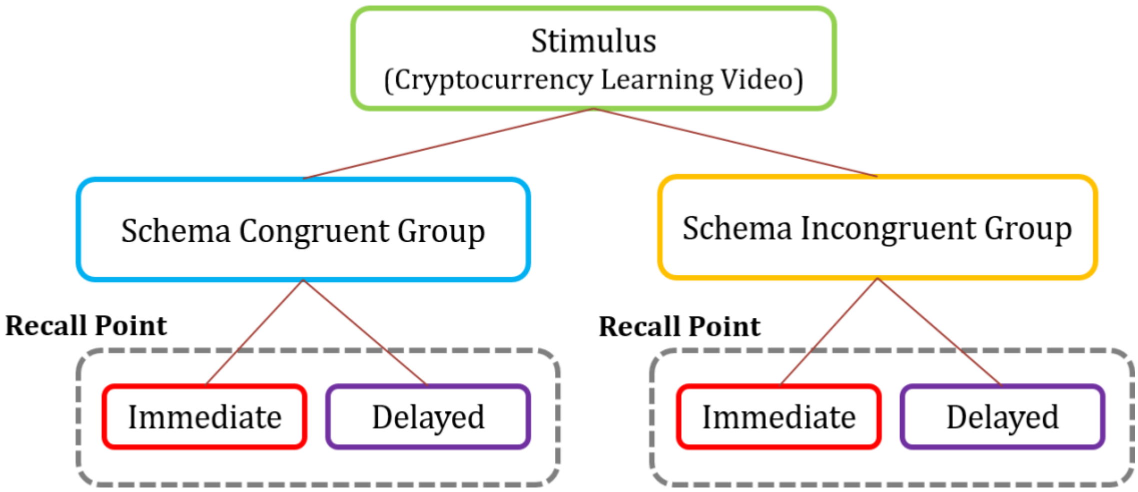


Figure 3. Experiment design paradigm

Each participant was exposed to both conditions of the within-subjects variable (i.e., recall point). Dependent variables (DV) are the outcomes measured in an experiment to examine its relationship with IV. The three DVs in this experiment are (a) retrieval accuracy, (b) response time and (c) emotional valence. Table 3 tabularise the experiment variables.

Table 3
Overview of experiment variables

Variable type	Variable name	Value
Independent	Schema congruent	–
Independent	Schema incongruent	–
Dependent	Retrieval accuracy	Percentage
Dependant	Response time	Seconds
Dependant	Emotional valence	Percentage

Procedure

The experiment consists of two phases: study phase and test phase. As illustrated in Figure 2, the study phase includes a pre-test and watching the cryptocurrency learning video; likewise, the test phase comprises immediate and delayed post-test and a breakout session activity.

Study phase
All participants joined the session individually. The web app link to the experiment was shared with the self-contained instructions and stated that they are later questioned based on the learning video. After reading the instructions, the software starts recording the participant's face and facial expressions until the experiment ends. The AI-based emotion analysis was conducted later based on recorded video data, not in real-time during the session. Before playing the cryptocurrency video, the participants need to complete the pre-test questionnaire, which took on average 165 seconds (i.e., 2 minutes and 45 seconds) to complete. Participants made their responses by clicking the corresponding answer box for each question. The system randomly shuffles both questions and answers to ensure that the sequence of questions and multiple-choice answers are different for each participant. Meanwhile, in the background, the web app maintains a detailed time-log of the total duration spent in the learning video by each participant, duration taken to respond to each question and time snapshots when participants enter and exist each stage (i.e., from pre-test till delayed post-test) as illustrated in Figure 2. All participants watched the entire video and some even replayed the learning video since they were allowed to do so for a better understanding. The entire study phase took 838 seconds (i.e., 13 minutes and 58 seconds) on average per participant.

Test phase
The post-tests consist of two types (immediate and delayed) with the same set of questions. A breakout session activity was placed after the immediate post-test, before assessing the participant's long-term memory retention via delayed post-test. After watching the learning video, participants were presented with an immediate post-test where they responded by clicking the corresponding answer box and a grace time period of 60 seconds was given for each question with a visual countdown progress bar displayed on top of the screen. After each question, nor at the end of the questionnaire, the test score or feedback (i.e., correct or incorrect answer) was not conveyed to the participant. The breakout session activity encompasses a series of basic psychological distractor tasks (e.g., vowel count, word list, pattern recognition and Stroop tests) to minimise participants rehearsing the currently studied material and ensure nothing is stored in their working memory before moving to delayed post-test. Each participant took 244 seconds (i.e., 4 minutes and 4 seconds) on average in the breakout session and the entire test phase took 752 seconds (i.e., 12 minutes and 32 seconds) on average.

Methodology

The recorded video data was analysed using the open-source Face-API.JS framework library (Mühler, 2019). This application programming language is a JavaScript module built on top of the Tensorflow.JS core kernel (Smilkov et al., 2019) and runs on the Node.JS platform (Tilkov & Vinoski, 2010). The framework consists of three main modules.

Face detection module

The face detection module consists of two types of convolutional neural networks named (a) single shot-multibox detector (SSD) and (b) tiny face detector (TFD). The SSD module is based on the MobileNetV1 object-detection machine learning algorithm (Howard et al., 2017). The SSD MobileNet model is pre-trained using the WIDERFACE public dataset (Yang et al., 2016), consisting of 32,203+ images and 393,703+ labelled faces. The machine learning algorithm can compute the location of a human face on each frame image on the video and return the bounding box together with a probability score of confidence (value > 0 and < 1) for each face (refer to Figure 4). The TFD module is an optimised version of the YOLO (You Only Look Once) object-detection machine learning algorithm (Redmon & Farhadi, 2017), which utilises depth-wise separable convolution instead of regular convolutions. This makes TFD a very performant real-time face detector, which is much faster, smaller and less resource-consuming than the SSD MobileNet model. However, in return, it performs slightly less accurate in detecting small faces. The TFD YOLO model has been pre-trained on a custom dataset of approximately 14000+ label images (Redmon & Farhadi, 2017). For this study, the TFD was selected as the default face detector module due to its minor trade-off of accuracy for speed, moderate resource consumption, being web-friendly (able to run the complete analysis on a browser) and the ability to easily adapt for different video resolutions (since the participants webcams had different resolutions).

Figure 4. Bounding box with probability score (Mühler, 2019)

Facial expression detection module

The facial expression feature extraction module consists of a 68-point face landmark detection model that employs ideas of both depth-wise separable convolutions and densely connected blocks (Wu et al., 2017). The 68-point model is structured as follows: 10 points for eyebrows (5 points each), 12 points for eyes (6 points each), 9 points for nose, 20 points for mouth and 17 points for face boundary. Figure 5 shows the 68 face landmark points. The model has been pre-trained on a dataset of approximately 35000+ labelled face images (Wu et al., 2017).

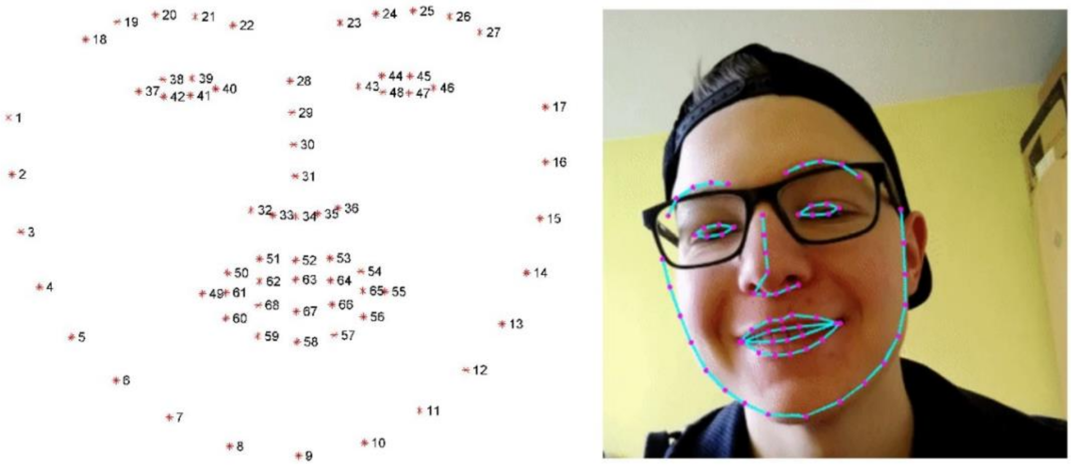


Figure 5. The 68 face landmark points (Mühler, 2019)

Emotion classification module

The emotion classification module also employs depth-wise separable convolutions and densely connected blocks to classify the six basic emotions. It is implemented based on ResNet-34 architecture (He et al., 2016) to compute facial expressions with a feature vector of 128 values for each face image describing a person's facial expression characteristics. The neural network utilises the Dlib library – a toolkit containing machine learning algorithms in the computer vision domain (King, 2009) and achieves a prediction accuracy of 99.38% (Mühler, 2019) based on the Labeled-Faces-in-the-Wild public dataset (G. Huang et al., 2008) that contained 13000+ images. However, it is noted that wearing glasses decreases the emotion prediction accuracy. Figure 6 shows the classification of six basic emotions and the probability of each face being classified as the type of emotion specified.

Figure 6. Classification of six basic emotions with probability scores (Mühler, 2019)

Results

Quantitative data were statistically analysed using IBM SPSS Statistics version 24 (https://www.ibm.com/products/spss-statistics/) software. In a study, the rationale for selecting the statistical tests depends on research questions and independent and dependent variables. For this 2 x 2 mixed factorial design experiment with categorical information (i.e., groups) on the IV and continuous information (e.g., test score retrieval accuracy and response time) on the DV, the analysis of variance (ANOVA) was selected to examine the inferential statistics. The retrieval accuracy is the percentage of correctly answered test scores and calculated for each participant's immediate and delayed sessions. Moreover, the response time is the total reaction time in seconds that each participant took to complete the immediate and delayed tests. Finally, emotional valence is the percentage of positive and negative emotions exhibited during study phase and test phase.

Data from the practice trials were excluded from the analysis. The p values set with an alpha (α) significance level of $p = .05$ and point estimates of effect sizes were presented by reporting the Eta-Squared (η_p^2). At the preliminary level, the Box's test assumptions were met ($p > .05$) for the results of all the ANOVAs to be valid and Levene's test requirements were also achieved ($p > .05$), which checks homogeneity variance for the main effects of the experiment. The interpretation of statistical tests results is discussed in terms of main effects for each independent variable and interaction effects between independent variables along with the level of significance by reporting effect sizes.

Statistical analysis of schema data

Repeated-measures mixed ANOVA (i.e., spilt-plot ANOVA) was applied to examine the retrieval accuracy of schema (congruent and incongruent) and recall point (immediate and delayed). Moreover, the schema was between-subjects and the recall point was within-subjects.

Retrieval accuracy

The 2 x 2 mixed ANOVA investigated the impact of schema and recall point on retrieval accuracy of post-test questions. There was no statistically significant main effect of schema on retrieval accuracy, $F(1,14) = .51$, $p = .486$, $\eta_p^2 = .04$. Also, there was no statistically significant main effect of recall point on retrieval accuracy, $F(1,14) = 1.20$, $p = .293$, $\eta_p^2 = .08$. However, there was a statistically significant interaction effect between schema and recall point on retrieval accuracy, $F(1,14) = 5.49$, $p = .034$, $\eta_p^2 = .28$ (refer to Figure 7). A post-hoc paired sample test revealed that the retrieval accuracy in the schema incongruent group is significantly higher than the schema congruent group in delayed recall, $t(7) = -2.76$, $p = .028$.

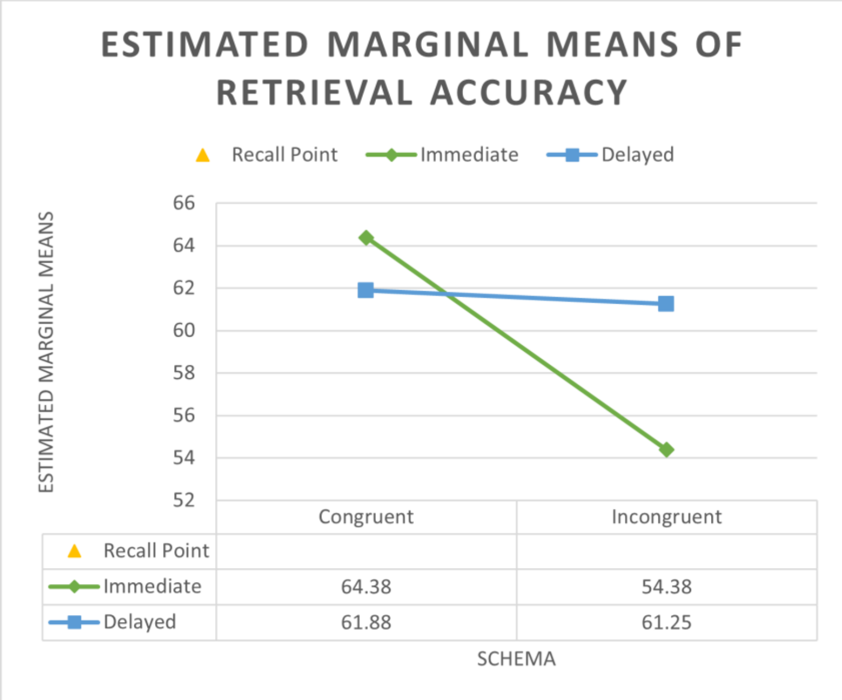


Figure 7. Schema and recall point on retrieval accuracy

Response time

The 2 x 2 mixed ANOVA investigated the impact of schema and recall point on response time in answering the post-test questions. There was no statistically significant main effect of schema on response time, $F(1,14) = .94$, $p = .348$, $\eta_p^2 = .06$. However, there was a statistically significant main effect in recall point on response time, $F(1,14) = 42.58$, $p < .001$, $\eta_p^2 = .75$, with the schema incongruent group ($M = 273.1$) significantly taking more time to respond overall than the schema congruent group ($M = 235.3$). Moreover, there was no statistically significant interaction effect between schema and recall *point* on response time, $F(1,14) = .01$, $p = .938$, $\eta_p^2 < .01$ (refer to Figure 8).

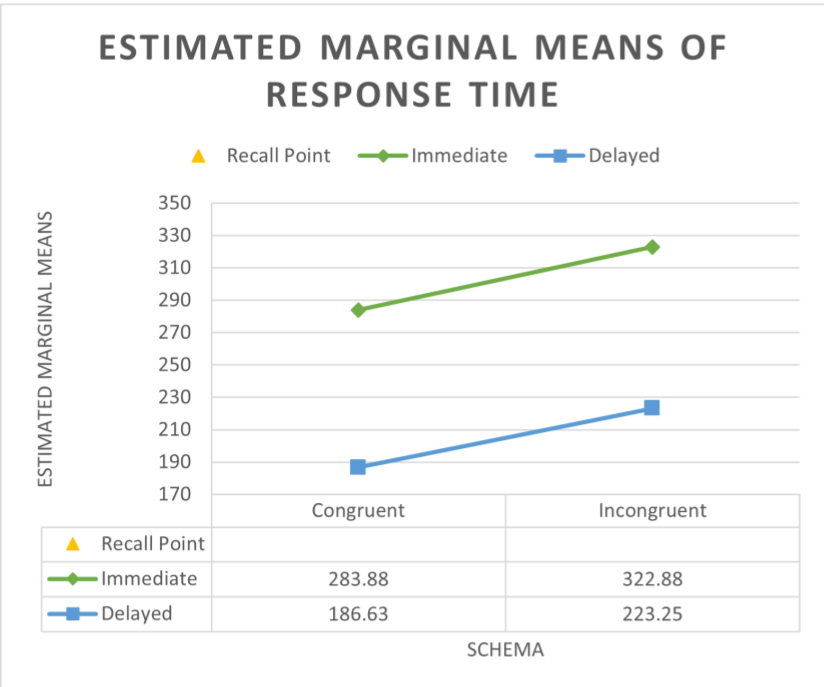
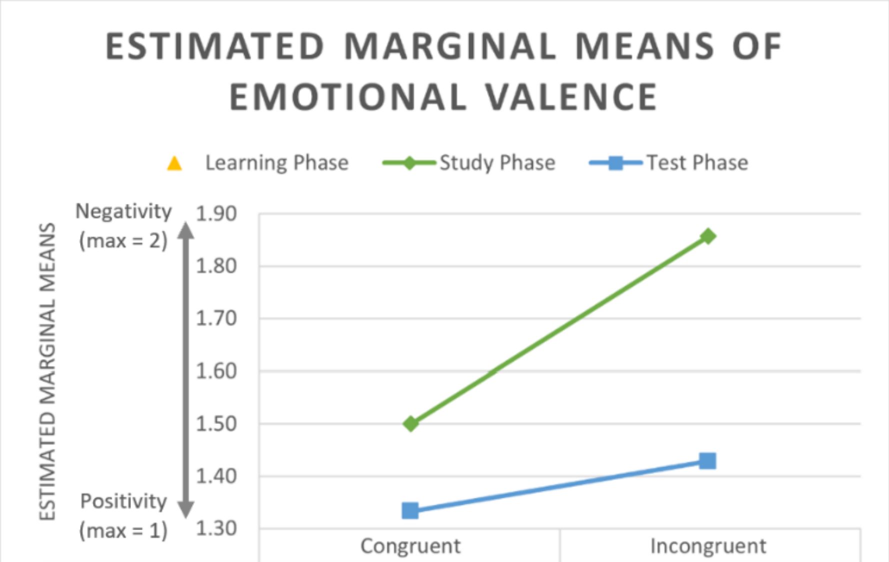


Figure 8. Schema and recall point on response time

Statistical analysis of emotion data

Repeated-measures mixed ANOVA (i.e., spilt-plot ANOVA) was applied to examine the emotional valence (positivity and negativity) of schema (congruent and incongruent) and learning phase (study phase and test phase). Moreover, the schema was between-subjects, and the learning phase was within-subjects.

Emotional valence
The 2 x 2 mixed ANOVA investigated the impact of schema and learning phase on emotions exhibit by participants. There was no statistically significant main effect of schema on emotions, $F(1,11) = 1.16$, $p = .304$, $\eta_p^2 = .10$. Also, there was no statistically significant main effect of learning phase on emotions, $F(1,11) = 2.77$, $p = .124$, $\eta_p^2 = .20$. Moreover, there was no statistically significant interaction effect between schema and learning phase on emotions, $F(1,11) = .54$, $p = .479$, $\eta_p^2 < .05$ (refer to Figure 9).



	Congruent	Incongruent
Learning Phase		
Study Phase	1.50	1.86
Test Phase	1.33	1.43

Figure 9. Schema and learning phase on emotional valence

Discussion

Memory retrieval accuracy of schema incongruent group was significantly higher than schema congruent group during delayed recall. This finding was somewhat surprising. We expected that the schema congruent participants would score more than the schema incongruent participants. Moreover, it is astounding to find out that the memory accuracy (i.e., test scores) had improved more during delayed recall compared to the immediate recall, though we expected the opposite. The reason for the former behaviour was noticed on the detailed time-logs of the web app. The schema incongruent participants have paused and replayed the learning video several times compared to the schema congruent participants before answering the MCQs. This may be due to a more elaborative form of processing done by the brain when it receives inconsistent information, which may lead to superior memory and this observation is consistent with the findings in Prull (2015). Regarding the latter behaviour (i.e., better memory accuracy during delayed recall), a possible explanation could be that the breakout session that comprises the psychological distractor tasks might not be long enough, which lead to rehearsing the currently studied material inside participants working memory (i.e., short-term memory). Furthermore, the web app time-logs prove this opinion as the data shows each participant had taken 4 minutes and 4 seconds on average in the breakout session, when the materials for the breakout activity were initially prepared for a duration of 10 to 15 minutes to ensure that nothing is stored in their memory before answering the delayed recall (i.e., long-term memory) questionnaire.

Response time duration of the schema incongruent group was significantly higher than the schema congruent group in both immediate and delayed recall points, as we expected. Inline to the explanation mentioned above, when incongruent participants' brain needs to process inconsistent information more elaboratively, it leads to longer response times than congruent participants' brain (Prull, 2015). Additionally, these response time results in terms of recall point (immediate and delayed) perspective were consistent and support the findings of Galfano et al. (2012), in which the response times for both congruent and incongruent groups have shown no significant differences.

Emotional valence findings of this study were interesting. First, both schema groups exhibited more negative emotions (69%) than positive emotions (31%) during the study phase. When analysing participants behaviour during study phase we noticed that the participants were paying more attention and focused during the learning video session. This concentrating and immersed behaviour exhibited facial expressions similar to serious-minded human nature; hence AI-based emotion analysis algorithm correctly classified these as negative emotions. Second, we examined the emotions exhibited during the test phase. Both schema groups exhibited more positive emotions (57%) than negative emotions (43%) in this phase. This finding was fascinating; we expected vice-versa – the test phase to have more negative emotions among participants since it consists of answering MCQ tests, compared to watching the video (study phase), which is a minor commitment. Our experience suggests the reason for this behaviour was the questionnaire tool (i.e., Quizizz Inc.) used for acquiring participants responses. It is a game-based (i.e., gamified) platform widely used to review and assess students' knowledge. The gamification nature of the tool motivates and increases participant engagement in such a way that they are answering a questionnaire without being aware of it. Conversely, the learning video session did not maintain a proper participant engagement.

Finally, it is important to consider that the above observations are not statistically significant. We believe that the smaller sample size (n = 16) used in our experiment was the main reason; hence the statistical signal-to-noise ratio was too large. Thus, repeating this same study in a broader sample size can decrease the statistical noise by adding more variance to the data.

Limitations

While this research provides compelling evidence, caution should be used in generalising the findings beyond the scope of the study. The results should be viewed considering the following limitations of this study. The first is the non-existence of a control group (i.e., traditional face-to-face learning group); this limits the external validity of results. Moreover, the study did not vary the learning materials according to the schema congruent and incongruent groups (i.e., designing learning materials with different difficulty levels) and explored the behaviour of each focused group. Furthermore, it would be better if the criteria for the schema congruent and incongruent groups were based on pre-test followed by a short interview with the participant. This gives a holistic view of the pre-existing knowledge base.

Additionally, increasing the delayed recall length from minutes to days would reflect a more real-world setting. However, researchers should be cautious that participants do not perform further reading or gain knowledge from any external environment relating to learning material topic (i.e., topic cryptocurrency according to this study). This will invalidate the delayed recall results to be compared with immediate recall. Another limitation is connected to the fact that the emotions were classified only using the ResNet-34 neural network, thus the classification results might be more solid if computed and validated using an additional machine-learning algorithm. Last but not least, the concept of flipped learning includes pre-class, in-class and post-class activities; however, this study focused only on two subsets of activities (watching learning videos and answering questionnaires) within the scope of the flipped learning framework. Future research is required to address the above limitations.

Conclusion

This research explored the behaviour of two focused groups (schema congruent and schema incongruent) of participants when experiencing video-based materials as a medium of learning within the frame of a flipped learning environment. The criteria for evaluating the effectiveness of video-based learning included (a) memory retrieval accuracy, (b) response time duration and (c) emotional valence. The findings indicate that retrieval accuracy for the schema incongruent group was better than schema congruent. Response time for schema congruent group was quicker than schema incongruent. Both groups exhibited more negative emotions during the study phase but more positive emotions during the test phase.

In summary, the findings presented in this research provide beneficial information for informed decision-making when designing educational courses that rely on video-based materials as the learning medium. For course designers, understanding the emotions and perceptions of students are helpful when planning future course modules. All course materials, both instructional videos and knowledge acquisition tools used in a video-based approach, should be carefully designed to motivate the students to learn the subject matters and contribute to a fun and positive learning experience. One might argue that this adds more time and resources to academics and higher education institutes for course preparation compared to the traditional face-to-face approach. However, considering the recent COVID-19 pandemic and the rate of digital learning adoption, this approach is more instructionally efficient and scalable than the traditional classroom approach. Finally, in a broader sense, the emotional insights derived from this study can be tailored to a particular industry, such as smart factories and smart healthcare, which pose many applicable scenarios for developing governance strategies, managing communication and policymaking.

Acknowledgements

We would like to thank the support given by voluntary participants and members in the School of Information Technology and the School of Psychology, Monash University, for this multidisciplinary research to be a success.

References

Anderson, J. R. (1981). Effects of prior knowledge on memory for new information. *Memory & Cognition*, *9*(3), 237–246. https://doi.org/10.3758/bf03196958

Arnau, D., Arevalillo-Herráez, M., Puig, L., & González-Calero, J. A. (2013). Fundamentals of the design and the operation of an intelligent tutoring system for the learning of the arithmetical and algebraic way of solving word problems. *Computers & Education*, *63*, 119–130. https://doi.org/10.1016/j.compedu.2012.11.020

Bartlett, F. C. (1932). Remembering: A study in experimental and social psychology. *British Journal of Educational Psychology*, *3*(2), 187–192. https://doi.org/10.1111/j.2044-8279.1933.tb02913.x

Betihavas, V., Bridgman, H., Kornhaber, R., & Cross, M. (2016). The evidence for 'flipping out': A systematic review of the flipped classroom in nursing education. *Nurse Education Today*, *38*, 15–21. https://doi.org/10.1016/j.nedt.2015.12.010

Bolton, K., Saalman, E., Christie, M., Ingerman, Å., & Linder, C. (2008). SimChemistry as an active learning tool in chemical education. *Chemistry Education Research and Practice*, *9*(3), 277–284. https://doi.org/10.1039/b812417p

Borah, R. (2021). Blended learning and flipped learning: Challenges and opportunities for the 21st-century students to create a digital environment. In S. Pal, T. Q. Cuong, & R. S. S. Nehru (Eds.), *Digital education for the 21st century* (1st ed., pp. 235–253). Apple Academic Press. https://doi.org/10.1201/9781003180517-10

Cervi, C. R., Brock, L. A., Galante, R., & de Oliveira, J. P. M. (2013). A computational tool to analyze the evolution of student learning. In *Proceedings of the 5th International Conference on Education and New Learning Technologies* (pp. 301–308). IATED. https://library.iated.org/view/cervi2013aco

Chaiyo, Y., & Nokham, R. (2017). The effect of Kahoot, Quizizz and Google Forms on the student's perception in the classrooms response system. In *Proceedings of the 2017 International Conference on Digital Arts, Media and Technology* (pp. 178–182). IEEE. https://doi.org/10.1109/icdamt.2017.7904957

D'Mello, S., Lehman, B., Pekrun, R., & Graesser, A. (2014). Confusion can be beneficial for learning. *Learning and Instruction*, *29*, 153–170. https://doi.org/10.1016/j.learninstruc.2012.05.003

Dalgleish, T., & Power, M. (2000). *Handbook of cognition and emotion*. John Wiley & Sons.

Darwin, C. (1872). *The expression of the emotions in man and animals*. John Murray. https://doi.org/10.1037/10001-000

Davis, R. A., & Moore, C. C. (1935). Methods of measuring retention. *The Journal of General Psychology*, *12*(1), 144–155. https://doi.org/10.1080/00221309.1935.9920093

Dawkins, M. S. (2000). Animal minds and animal emotions. *American Zoologist*, *40*(6), 883–888. https://doi.org/10.1093/icb/40.6.883
Ekman, P. (Ed.). (1973). *Darwin and facial expression: A century of research in review* (1st ed.). Academic Press.
Faul, F., Erdfelder, E., Buchner, A., & Lang, A.G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. https://doi.org/10.3758/brm.41.4.1149
Faul, F., Erdfelder, E., Lang, A.G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/bf03193146
Findlay-Thompson, S., & Mombourquette, P. (2014). Evaluation of a flipped classroom in an undergraduate business course. *Business Education & Accreditation*, *6*(1), 63–71. https://ssrn.com/abstract=2331035
Fleming, M. (1987). Displays and communication. *Instructional Technology Foundations*, 233–260. https://doi.org/10.4324/9781315060248
Fredrickson, B. L. (1998). What good are positive emotions? *Review of General Psychology*, *2*(3), 300–319. https://doi.org/10.1037/1089-2680.2.3.300
Fredrickson, B. L. (2001). The role of positive emotions in positive psychology: The broaden-and-build theory of positive emotions. *American Psychologist*, *56*(3), 218–226. https://doi.org/10.1037/0003-066x.56.3.218
Gagne, R. M. (1986). Instructional technology: The research field. *Journal of Instructional Development*, *8*(3), 7–14. https://doi.org/10.1007/bf02906263
Galfano, G., Dalmaso, M., Marzoli, D., Pavan, G., Coricelli, C., & Castelli, L. (2012). Eye gaze cannot be ignored (but neither can arrows). *The Quarterly Journal of Experimental Psychology*, *65*(10), 1895–1910. https://doi.org/10.1080/17470218.2012.663765
Ghosh, V. E., & Gilboa, A. (2014). What is a memory schema? A historical perspective on current neuroscience literature. *Neuropsychologia*, *53*, 104–114. https://doi.org/10.1016/j.neuropsychologia.2013.11.010
Gilboy, M. B., Heinerichs, S., & Pazzaglia, G. (2015). Enhancing student engagement using the flipped classroom. *Journal of Nutrition Education and Behavior*, *47*(1), 109–114. https://doi.org/10.1016/j.jneb.2014.08.008
Göksün, D., & Gürsoy, G. (2019). Comparing success and engagement in gamified learning experiences via Kahoot and Quizizz. *Computers & Education*, *135*, 15–29. https://doi.org/10.1016/j.compedu.2019.02.015
Gray, L. M., Wong Wylie, G., Rempel, G. R., & Cook, K. (2020). Expanding qualitative research interviewing strategies: Zoom video communications. *The Qualitative Report*, *25*(5), 1292–1301. https://www.proquest.com/openview/c264828516f288b941ad22c63c576706/
Greve, A., Cooper, E., Tibon, R., & Henson, R. N. (2019). Knowledge is power: Prior knowledge aids memory for both congruent and incongruent events, but in different ways. *Journal of Experimental Psychology: General*, *148*(2), 325. https://doi.org/10.1037/xge0000498
Guo, P. J., Kim, J., & Rubin, R. (2014). How video production affects student engagement: An empirical study of MOOC videos. In *Proceedings of the first ACM Conference on Learning @ Scale* (pp. 41–50). ACM. https://doi.org/10.1145/2556325.2566239
Halawi, L. A., McCarthy, R. V., & Pires, S. (2009). An evaluation of e-learning on the basis of Bloom's taxonomy: An exploratory study. *Journal of Education for Business*, *84*(6), 374–380. https://doi.org/10.3200/joeb.84.6.374-380
Hascher, T. (2010). Learning and emotion: Perspectives for theory and research. *European Educational Research Journal*, *9*(1), 13–28. https://doi.org/10.2304/eerj.2010.9.1.13
He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition* (pp. 770–778). IEEE. https://doi.org/10.1109/cvpr.2016.90
Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., & Adam, H. (2017). MobileNets: Efficient convolutional neural networks for mobile vision applications. *arXiv*. https://doi.org/10.48550/arxiv.1704.04861
Huang, G., Mattar, M., Berg, T., & Learned Miller, E. (2008, October). *Labeled faces in the wild: A database for studying face recognition in unconstrained environments* [Workshop presentation]. Faces in 'real-life' images: Detection, alignment, and recognition, France.

Huang, R., Spector, J. M., & Yang, J. (2019). *Educational technology: A primer for the 21st century*. Springer. https://doi.org/10.1007/978-981-13-6643-7
Inan, F. A., Lowther, D. L., Ross, S. M., & Strahl, D. (2010). Pattern of classroom activities during students' use of computers: Relations between instructional strategies and computer applications. *Teaching and Teacher Education, 26*(3), 540–546. https://doi.org/10.1016/j.tate.2009.06.017
Izard, C. E. (1971). *The face of emotion*. Appleton Century Crofts.
Izard, C. E. (1991). *The psychology of emotions*. Springer.
Izard, C. E. (2009). Emotion theory and research: Highlights, unanswered questions, and emerging issues. *Annual Review of Psychology, 60*(1), 1–25. https://doi.org/10.1146/annurev.psych.60.110707.163539
Izard, C. E. (2013). *Patterns of emotions: A new analysis of anxiety and depression*. (Rev. ed.) Academic Press. https://doi.org/10.1016/c2013-0-10907-5
Johnson, M. (1987). *The body in the mind: The bodily basis of meaning, imagination, and reason*. University of Chicago Press. https://doi.org/10.7208/chicago/9780226177847.001.0001
Jonassen, D. H., & Grabowski, B. L. (1993). *Handbook of individual differences, learning, and instruction* (1st ed.). Routledge. https://doi.org/10.4324/9780203052860
Kant, I. (1899). *Critique of pure reason* (J. M. D. Meiklejohn, Trans.). Willey Book. https://doi.org/10.1037/11654-000
King, D. E. (2009). Dlib-ml: A machine learning toolkit. *The Journal of Machine Learning Research, 10*, 1755–1758. https://doi.org/10.5555/1577069.1755843
Kolås, L. (2015). Application of interactive videos in education. In *Proceedings of the 2015 International Conference on Information Technology Based Higher Education and Training* (pp. 1–6). IEEE. https://doi.org/10.1109/ithet.2015.7218037
Kolekar, S. V., Pai, R. M., & Pai, M. M. (2018). Adaptive user interface for Moodle based e-learning system using learning styles. *Procedia Computer Science, 135*, 606–615. https://doi.org/10.1016/j.procs.2018.08.226
Kristjánsson, K. (2016). *Aristotle, emotions, and education*. Routledge. https://doi.org/10.4324/9781315567914
LaBar, K. S., & Cabeza, R. (2006). Cognitive neuroscience of emotional memory. *Nature Reviews Neuroscience, 7*(1), 54–64. https://doi.org/10.1038/nrn1825
Laura, C., & Olivier, R. (Eds.). (2020). Introduction: Why Plato comes first. In *Emotions in Plato* (Vol. 4, pp. 1–14). Brill. https://doi.org/10.1163/9789004432277_002
Leff, J. P. (1973). Culture and the differentiation of emotional states. *British Journal of Psychiatry, 123*(574), 299–306. https://doi.org/10.1192/bjp.123.3.299
Lin, Y., & Nguyen, H. (2021). International students' perspectives on e-learning during COVID-19 in higher education in Australia: A study of an Asian student. *Electronic Journal of e-Learning, 19*(4), 241–251. https://doi.org/10.34190/ejel.19.4.2349
Little, J. L., & Bjork, E. L. (2015). Optimizing multiple-choice tests as tools for learning. *Memory & Cognition, 43*(1), 14–26. https://doi.org/10.3758/s13421-014-0452-8
Lowenthal, P., & Wilson, B. G. (2010). Labels do matter! A critique of AECT's redefinition of the field. *TechTrends, 54*(1), 38–46. https://doi.org/10.1007/s11528-009-0362-y
Lu, W. (2014). Using case study research as an active learning tool for demonstrating the ability to function on multidisciplinary teams. In *Proceedings of the 121st ASEE Annual Conference & Exposition* (pp. 1–12). ASEE. https://doi.org/10.18260/1-2--23256
Mandler, G. (1959). Stimulus variables and subject variables: A caution. *Psychological Review, 66*(3), 145–149. https://doi.org/10.1037/h0043276
Marsh, E. J., Roediger, H. L., Bjork, R. A., & Bjork, E. L. (2007). The memorial consequences of multiple-choice testing. *Psychonomic Bulletin & Review, 14*(2), 194–199. https://doi.org/10.3758/bf03194051
Moussiades, L., Kazanidis, I., & Iliopoulou, A. (2019). A framework for the development of educational video: An empirical approach. *Innovations in Education and Teaching International, 56*(2), 217–228. https://doi.org/10.1080/14703297.2017.1399809
Mühler, V. (2019). *Face-API.JS* (Version 0.17.1) [JavaScript API]. GitHub. https://github.com/justadudewhohacks/face-api.js.git
Myllymäki, M., Hakala, I., Härmänmaa, T., & Laine, S. (2017). Flipped learning experiment in video-based education. In *Proceedingsof the 9th International Conference on Education and New Learning Technologies* (pp. 2415–2424). IATED. https://doi.org/10.21125/edulearn.2017.1502

Pekrun, R. (1992). The impact of emotions on learning and achievement: Towards a theory of cognitive/motivational mediators. *Applied Psychology*, *41*(4), 359–376. https://doi.org/10.1111/j.1464-0597.1992.tb00712.x

Pekrun, R. (2014). *Emotions and learning*. UNESCO.

Phelps, E. A. (2004). Human emotion and memory: Interactions of the amygdala and hippocampal complex. *Current Opinion in Neurobiology*, *14*(2), 198–202. https://doi.org/10.1016/j.conb.2004.03.015

Piaget, J. (1952). *The origins of intelligence in children*. Norton & Co. https://doi.org/10.1037/11494-001

Prull, M. W. (2015). Adult age differences in memory for schema-consistent and schema-inconsistent objects in a real-world setting. *Aging, Neuropsychology, and Cognition*, *22*(6), 731–754. https://doi.org/10.1080/13825585.2015.1037821

Redmon, J., & Farhadi, A. (2017). YOLO9000: Better, faster, stronger. In *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition* (pp. 6517–6525). IEEE. https://doi.org/10.1109/cvpr.2017.690

Richey, R. C., Klein, J. D., & Tracey, M. W. (2010). *The instructional design knowledge base: Theory, research, and practice*. Routledge. https://doi.org/10.4324/9780203840986

Riedesel, M. A., & Charles, P. (2018). Learning any time, anywhere: Big educational data from smart devices. In J. M. Spector, V. Kumar, A. Essa, Y. M. Huang, R. Koper, R. A. W. Tortorella, T. W. Chang, Y. Li, & Z. Zhang (Eds.), *Frontiers of cyberlearning: Emerging technologies for teaching and learning* (pp. 1–31). Springer. https://doi.org/10.1007/978-981-13-0650-1_1

Righi, S., Marzi, T., Toscani, M., Baldassi, S., Ottonello, S., & Viggiano, M. P. (2012). Fearful expressions enhance recognition memory: Electrophysiological evidence. *Acta Psychologica*, *139*(1), 7–18. https://doi.org/10.1016/j.actpsy.2011.09.015

Rumelhart, D. E. (2017). Schemata: The building blocks of cognition. In R. J. Spiro, B. C. Bruce, & W. F. Brewer (Eds.), *Theoretical issues in reading comprehension* (Rev ed., pp. 33–58). Routledge. https://doi.org/10.4324/9781315107493-4

Sá, M. J., & Serpa, S. (2020). COVID-19 and the promotion of digital competences in education. *Universal Journal of Educational Research*, *8*(10), 4520–4528. https://doi.org/10.13189/ujer.2020.081020

Saito, A. (1996). Social origins of cognition: Bartlett, evolutionary perspective and embodied mind approach. *Journal for the Theory of Social Behaviour*, *26*(4), 399–421. https://doi.org/10.1111/j.1468-5914.1996.tb00299.x

Sampath, K. (1981). *Introduction to educational technology*. Sterling Publishers.

Seli, P., Wammes, J. D., Risko, E. F., & Smilek, D. (2016). On the relation between motivation and retention in educational contexts: The role of intentional and unintentional mind wandering. *Psychonomic Bulletin & Review*, *23*(4), 1280–1287. https://doi.org/10.3758/s13423-015-0979-0

Sentis, K. P., & Burnstein, E. (1979). Remembering schema-consistent information: Effects of a balance schema on recognition memory. *Journal of Personality and Social Psychology*, *37*(12), 2200–2211. https://doi.org/10.1037/0022-3514.37.12.2200

Sloan, T. W., & Lewis, D. A. (2014). Lecture capture technology and student performance in an operations management course. *Decision Sciences Journal of Innovative Education*, *12*(4), 339–355. https://doi.org/10.1111/dsji.12041

Smilkov, D., Thorat, N., Assogba, Y., Yuan, A., Kreeger, N., Yu, P., Zhang, K., Cai, S., Nielsen, E., & Soergel, D. (2019). Tensorflow.js: Machine learning for the web and beyond. In A. Talwalkar, V. Smith, & M. Zaharia (Eds.), *MLSys 2019—Proceedings of Machine Learning and Systems 1* (pp. 309–321). https://proceedings.mlsys.org/paper/2019/file/1d7f7abc18fcb43975065399b0d1e48e-Paper.pdf

SooHwan, K., HyeonCheol, K., & Han, S. (2013). A development of learning widget on m-learning and e-learning environments. *Behaviour & Information Technology*, *32*(2), 190–202. https://doi.org/10.1080/0144929x.2011.605907

Thong, L. P., Stewart, C., Arnab, S., & Lameras, P. (2016). Virtual designer: Digital role-playing game for knowledge transferal in design education. In T. Connolly & L. Boyle (Eds.), *Proceedings of the 10th European Conference on Games Based Learning* (Vol. 1, pp. 862–869). http://toc.proceedings.com/31939webtoc.pdf

Tilkov, S., & Vinoski, S. (2010). Node.js: Using JavaScript to build high-performance network programs. *IEEE Internet Computing*, *14*(6), 80–83. https://doi.org/10.1109/mic.2010.145

Um, E., Plass, J. L., Hayward, E. O., & Homer, B. D. (2012). Emotional design in multimedia learning. *Journal of Educational Psychology*, *104*(2), 485–498. https://doi.org/10.1037/a0026609

van Kesteren, M. T. R., Ruiter, D. J., Fernández, G., & Henson, R. N. (2012). How schema and novelty augment memory formation. *Trends in Neurosciences, 35*(4), 211–219. https://doi.org/10.1016/j.tins.2012.02.001
Vikan, A. (2017). *Fast road to the study of emotions*. Springer. https://doi.org/10.1007/978-3-319-52313-2
Vuilleumier, P. (2005). How brains beware: Neural mechanisms of emotional attention. *Trends in Cognitive Sciences*, *9*(12), 585–594. https://doi.org/10.1016/j.tics.2005.10.011
Willey, K. (2016). Combining a collaborative learning framework with an e-learning tool to improve learning and professional development in blended learning environments. In *Proceedings of the 2016 Future Technologies Conference* (pp. 1303–1304). IEEE. https://doi.org/10.1109/ftc.2016.7821769
Willey, K., & Gardner. (2014). Combining flipped instruction and multiple perspectives to develop cognitive and affective processes. In *Proceedings of the SEFI Annual Conference*. European Society for Engineering Education. https://www.sefi.be/wp-content/uploads/2017/10/0092.pdf
Winn, W. (2004). Cognitive perspectives in psychology. In D. Jonassen & M. Driscoll (Eds.), *Handbook of research on educational communications and technology* (2nd ed., pp. 79–112). Routledge. https://doi.org/10.4324/9781410609519
Wu, Y., Hassner, T., Kim, K., Medioni, G., & Natarajan, P. (2017). Facial landmark detection with tweaked convolutional neural networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 40*(12), 3067–3074. https://doi.org/10.1109/tpami.2017.2787130
Yang, S., Luo, P., Loy, C.C., & Tang, X. (2016). WIDER FACE: A face detection benchmark. In *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition* (pp. 5525–5533). IEEE. https://doi.org/10.1109/cvpr.2016.596

Corresponding author: Anjana Junius Vidanaralage, junius.kandevidanaralalage@monash.edu

Please cite as: Vidanaralage, A. J., Dharmaratne, A. T., & Haque, S. (2022). Schema and emotion in memory retrieval following video-based learning: An artificial intelligence study. *Australasian Journal of Educational Technology*, *38*(5), 109-132. https://doi.org/10.14742/ajet.7473

Appendix A: Pre-test questionnaire

Q1 – What best describes your gender?
- Male
- Female
- Prefer not to say

Q2 – Which age group describes you?
- 19 or under
- 20 – 24
- 25 – 29
- 30 – 34
- 35 or above

Q3 – What is the highest education level you have completed?
- No schooling completed
- High school or college graduate
- Bachelor’s degree holder
- Master’s degree holder
- Doctorate degree holder

Q4 – Are you left or right handed?
- Left handed
- Right handed

Q5 – Are you currently wearing glasses?
- Yes
- No

Q6 – Any history of learning disability?
- Yes
- No
- Prefer not to say

Q7 – Any history of psychotic disorders?
- Yes
- No
- Prefer not to say

Q8 – Any alcohol consumption in the last 24 hours?
- Yes
- No
- Prefer not to say

Q9 – How would you describe your cryptocurrency knowledge?
- Nil - No Crypto Knowledge
- Novice - Crypto Theoretical Knowledge Only
- Competent - Crypto Theoretical + Practical Usage
- Expert - Crypto Trader or Investor

Q10 – Have you ever bought or invested in cryptocurrencies?
- Yes
- No

Q11 – What is the ticker symbol for Bitcoin?
- BIT
- BTC
- BC
- BI
- I don’t know

Q12 – Which is a non-fiat currency?
- Dollar
- Bitcoin
- Pound
- Euro
- I don’t know

Q13 – What is the notation symbol for Bitcoin?
- $
- €
- ₿
- £
- I don’t know

Q14 – Where is the Bitcoin central server located?
- Undisclosed Location
- United States
- European Union
- None – It’s Decentralized
- I don’t know

Q15 – The process of encrypting and decrypting information is known as?
- Cryptocurrency
- PIN
- Cryptography
- Password
- I don’t know

Q16 – The Blockchain is same as Bitcoin?
- True
- False
- I don’t know

Q17 – Blockchain transactions are recorded automatically?
- Yes
- No
- I don’t know

Q18 – Ethereum Virtual Machine is powered by Bitcoin?
- Yes
- No
- I don’t know

Q19 – What does P2P stand for?

- Password to Password
- Product to Product
- Peer to Peer
- Private Key to Public Key
- I don’t know

Q20 - Which of the following is used to access cryptocurrency?

- Wallet
- Pocket
- Bridge
- Gate
- I don’t know

Appendix B: Post-test questionnaire

Q1 – What is a fiat currency?
- Currency backed by gold standard
- Currency backed by physical commodity
- Currency declared by a government
- Currency backed by cryptocurrency

Q2 – What is a cryptocurrency transaction?
- Encrypting a digital asset using public key
- Transfer of private key to access a digital asset
- Decrypting digital asset using private key
- Transferring the value of a digital asset

Q3 – Does cryptocurrency requires a bank or trusted third-party for a transaction?
- Yes, need both
- No, it is decentralized
- Not a third party, but need a central bank
- Not a bank, but need a trusted third-party

Q4 – What is a node?
- A joint in a Blockchain
- A block in a Blockchain
- A computer in a distributed network
- A point in a network connection

Q5 – When a cryptocurrency transaction is confirmed, can it be altered or chargeback?
- Can chargeback only if receiver approve
- It is irreversible cannot be altered
- Can be chargeback by the sender
- Can only be altered by central bank

Q6 – Who created Bitcoin?
- Anonymous
- Elon Musk
- Satoshi Nakamoto
- Nickoi Szakabo

Q7 – How Bitcoin obtain its value?
- National banks demand
- Rarity to find
- Trading demand
- Limited supply

Q8 – A decentralized distributed digital ledger that keeps and tracks all transactions is known as?
- Blocklink
- National database
- Blockchain
- Ledgerchain

Q9 – How many Bitcoins there ever will be?
- 1 million
- 21 million
- 140 million
- Unlimited

Q10 – According to current projections, when is last Bitcoin be issued?
- 2024
- 2040
- 2140
- 2224

Q11 – How are blocks in a Blockchain linked?
- Using cryptographic hash
- Using public and private keys
- Using nodes
- Using Blocklink

Q12 – How does Blockchain achieve its high availability?
- Blockchain integrated a complex algorithm
- Nodes always online 24/7 x 365 days
- Decentralized distributed network
- Satellite coverage around the globe

Q13 – What does Blockchain digital ledger contain?
- The mining algorithm
- List of transactions
- List of blocks in Blockchain
- List of cryptocurrencies

Q14 – Which mining algorithm does the Bitcoin use?
- BTC-256
- HASH-256
- SHA-256
- CASH-256

Q15 – Output string of numbers and letters with same length, generated by an algorithm is called?
- PIN
- Key
- Cache
- Hash

Q16 – Computers that verify transactions and updates the Blockchain by adding new blocks are called?
- Miners
- Excavators
- Hashing
- Workers

Q17 – What is a block reward?
- All transaction fees in a block
- Bitcoins generated by adding new blocks
- All transaction fees in a block (+) Bitcoins generated by adding new blocks
- Bitcoins generated by adding new blocks (-) all transaction fees in a block

Q18 – Why all miners in a Blockchain needs to verify a transaction?
- To increase network speed
- To generate new Bitcoins
- To distribute transaction fees
- To achieve consensus

Q19 – The process of solving a complex mathematical equation to find the ‘Nonce number’ is?
- Encrypting difficulty
- Decrypting difficulty
- Mining difficulty
- Hashing difficulty

Q20 - Each block contains the hash of?
- Next block
- Previous block
- Each block
- All blocks