

A Historical Text-Based Game Designed to Develop Critical Thinking Skills

Ivo de Vero, University of Glasgow, UK

Matthew Barr, University of Glasgow, UK*

 <https://orcid.org/0000-0002-5147-0673>

ABSTRACT

Designing an intervention that can effectively develop critical thinking skills is challenging because of the problems of transfer and domain specificity. The authors describe the design and development of a text-based game that could teach players important critical thinking skills in the domain of history. This was achieved by combining Schon's reflective practitioner model with game-based learning principles. The work contributes to the existing literature because the combination of the models employed allowed the game design to address the problem of transfer, as well as developing critical thinking skills. The instrument used to evaluate the effectiveness of the game was a questionnaire based on the reflective practitioner model. The gathered qualitative data were analysed through affinity diagramming. The results show that the game that was developed has the potential to encourage advanced levels of historical thought, as well as critical thinking skills.

KEYWORDS

Critical Thinking, Higher Education, History, Reflective Practitioner, Schon, Historical Thought

INTRODUCTION

The aim of the work described here was to develop a game harnessing game-based learning (GBL) principles to strengthen critical thinking (CT) skills and dispositions in the context of the discipline of history.

The experience shared in this report makes two contributions to the existing literature on GBL. The first crucial differentiator is the employment of the Reflective Practitioner model of behaviour (Schon, 1984) in the context of a GBL-based CT intervention. The insight Schon brings is that because practitioners learn by drawing on previous experience, then a game explicitly designed to boost CT skills and dispositions in a practitioner could form part of that experience.

Second is the key theoretical link between the idea of internal motivation as expressed by Malone (1980) as a factor which encourages people to play games and the idea of motivation as a set of consistent internal behaviours described by Facione (2000) in the context of developing CT skills. Most academic work which discusses GBL and CT together highlights the advantages games offer to skills development but somewhat neglect the equally important side of CT dispositions (see McCall, 2013 and 2016, for a review of such literature in the field of history).

DOI: 10.4018/IJGBL.323138

*Corresponding Author

This article published as an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>) which permits unrestricted use, distribution, and production in any medium, provided the author of the original work and original publication source are properly credited.

BACKGROUND

The Challenges of Teaching Critical Thinking

There is an abundant literature on the importance of critical thinking (CT). Hunt (1995) argued that the global economy had entered a new phase which required a new kind of worker, the “knowledge worker”, who would be capable of manipulating “abstract and complex symbols and ideas” and could “remain flexible enough to recognise the need for continuing change” (Hunt, 1995). As a result of surging demand for the knowledge worker, any country which wished to remain competitive in the global marketplace would need to reform its education system to emphasise CT: that combination of skills and attitude that would form the knowledge worker (Hunt, 1995). Two decades later, Western educators continue to emphasize the importance of CT for the success of the next generation and lament current institutional barriers to teaching it (Davies, 2016 and Staton, 2021). Halpern (1998) noted that an absence of the CT also has negative repercussions for people in their personal lives, not just their career. A substantial part of the American population, Halpern pointed out, spent more money than they could afford to do on psychics and pseudo-scientific remedies (Halpern, 1998). This has not appeared to improve substantially in the past two decades: McLaughlin and McGill (2017) note that belief in pseudo-historical narratives around ancient civilisations, such as the Maya, can negatively impact the well-being of their contemporary descendants. Conspiracy theories have been part and parcel of American political discourse for several years now with potentially dire implications for the health of its democracy (Venkataramakrishnan, 2020). Thus, the difficulty of teaching CT appears to be a pervasive problem with global reach and sufficient complexity to continue to challenge educators.

In a comprehensive literature review of the major academic contributions to the subject of CT, Lai (2011) describes three areas of agreement among scholars of the different contributing fields. First, that CT has a skill component. These are abilities such as analysis, inference, evaluation and problem solving (Lai, 2011). Second, CT has a dispositional or attitudinal component. CT dispositions are “consistent internal motivations to act toward or respond to persons, events, or circumstances in habitual, yet potentially malleable ways” (Facione, 2000) and include traits like open and fair-mindedness, inquisitiveness, and the desire to be well-informed (Lai, 2011). Finally, there is agreement on the fact that background knowledge is a prerequisite for effectively employing CT. This is because the evidence, reasonings and explanations that are considered to be examples of CT are domain-specific (Lai, 2011).

The interplay between these three components is what makes teaching CT so challenging. Facione (2000) notes that having CT skills does not correlate with a disposition towards applying CT, although there is a correlation between disposition and proficiency with CT skills. Attention must be paid to both for there to be an improvement in students. Furthermore, Willingham (2008) highlights that once students have been trained to think critically in one domain, they do not automatically do so in others. Even within the same broad domain such as mathematics or history, if students are taught to apply critical thinking to one specific type of problem, and then encounter a problem of a different type, there is no guarantee that those CT skills will transfer over to the new problem (Willingham, 2008). The importance of domain knowledge is such that the most effective CT interventions are those where it is taught in the context of a specific discipline, an observation also supported by Lai (2011).

Critical Thinking in the Domain of History

Because of the fundamental challenge of transferring CT skills from one domain to another (Lai, 2011, and Willingham, 2008), the authors restrict the scope of the work described here to one domain, the discipline of history. There is a literature which shows that history can engage both the dispositions and skills necessary to think critically, making it a useful domain to focus on.

Yogev (2013) argues that CT should be considered a fundamental part of the discipline because of the role it plays in forming citizens who actively engage with their democracy. Societal trends like the consumption of information through scattered media sources and the fragmentation of society

into smaller communities further increases the challenge of developing “historical consciousness” in students (Yogev, 2013). Clear in this argument is the implicit belief that teaching students history will encourage them to think critically about broader issues, or in other words, increase student disposition to think critically. McLaughlin and McGill (2017) view history as a particularly appropriate domain to boost the CT skills of students because of the type of thinking the field demands of its scholars. Indeed, the example Halpern (1998) gives to illustrate a critical thinking task is of weighing the relative credibility of two sources, which is an integral part of the work of the historian, as the rest of this section will illustrate.

The authors’ understanding of the process of the historian at work is based on *The Pursuit of History*, a book by Tosh and Lang marketed as the “essential introduction to the practice of history” (Tosh and Lang, 2006). Therein, Tosh and Lang (2006) spend the majority of their pages describing the process of the historian. This is something which sets the work apart from other renowned books on the subject by historians such as Carr (2018), which focus primarily on providing a definition of the discipline.

The first step Tosh and Lang (2006) outline in the historian’s process is what they term “external criticism”. This involves checking the authenticity of a historical document in the sense of verifying the truthfulness of its geographical and temporal details, to guard against forgery. The second step of the process is “internal criticism”, or actually interpreting the document the historian is working with, which the authors note is “usually much more demanding”. Here, the historian must consider what influenced the author of a document at the time of writing, as well as the concept of bias; that is, the intended purpose of the document at the time and the deliberate or unconscious distortions in the text that could have caused. The historian must thus weigh sources against each other, analysing their relative merit to the question they are investigating or constructing around the historical documents at their disposal.

There should be no doubt at this point that historians must exercise CT to be successful. Tosh and Lang (2006) summarise the process of the historian by calling it “common sense applied very much more systematically and sceptically than is usually the case in everyday life, supported by a secure grasp of the historical context and, in many instances, a high degree of technical knowledge”. Thus, the process here outlined is the domain-specific CT that is the focus of this work, what Willingham (2008) noted might be called “thinking historically”.

The Reflective Practitioner

The theoretical framework the authors selected as a basis for this work is provided by Schon (1984) in his highly influential book *The Reflective Practitioner*. The core of the ‘Reflective Practitioner’ (RP) model is the concept of “knowing-in-action” (Schon, 1984). Schon describes this as spontaneous “actions, recognitions and judgements” performed by the practitioner. The practitioner is often unaware of having learned how to perform, and is typically unable to describe the knowledge implicit in their actions. Schon showed through his studies of a broad range of practitioners, ranging from architects to therapists, that this concept of “knowing-in-action” was common across all of them.

The RP model argues that problem solving is a combination of experience and technical know-how. A practitioner does not assume a problem as given. Instead, they aim to identify the particular features of the specific problem they are working on by reflecting on what they can observe and on their previous experience solving similar problems. Schon (1984) defines this as “seeing-as”, in other words, framing a new problem in the context of everything they have learned solving past problems. Once a practitioner can relate a new problem to past experience, then they can “do-as” (Schon, 1984), applying what they know to the new problem, even if they cannot describe what it is that they’re seeing-as and doing-as.

Although Schon was not writing in the context of CT education, the parallels between the RP model and the effective application of CT are striking. That the RP model captures the work of the historian is also evident. The process outlined by Tosh and Lang (2006) has no formula to be

systematically applied for a given set of similar historical problems. Subjecting unique sources to rigorous external and internal criticism, as outlined above, implies treating each problem as new, then critically observing and reflecting. The project described here, then, aims to investigate history-specific CT skills and dispositions by applying the Reflective Practitioner model to Tosh and Lang's overview of the historian's process.

Game-Based Learning and Its Advantages

Traditional classroom interventions focussing on CT tend to concentrate on developing skills as opposed to increasing the disposition to think critically (Facione, 2000, Willingham, 2008, and Lai, 2011). A robust example of this type of intervention is reported by Reed and Kromrey (2011), who designed a college history module to explicitly teach CT skills together with history content, finding positive results. However, as Facione (2000) and Willingham (2008) highlight, CT dispositions need to be considered alongside skills for maximum impact.

There is a literature which has consistently observed that videogames are an effective medium for developing skills while maintaining a high level of motivation to do so in their players, of which Krath et al. (2021) provide a comprehensive summary. Chowanda and Chowanda (2016) find statistically significant evidence, of particular interest to the authors of this report, that playing a history-themed game helped a group of secondary school students remain motivated throughout the class and remember names and timelines more effectively than their control group utilising textbooks. A valuable perspective is further offered by Barr (2019), who ran a randomised controlled trial to measure learning gains of more abstract skills such as communication and adaptability in a game-playing intervention group composed of university students. Barr finds statistically significant learning gains, and offers qualitative results that describe an engaging and entertaining learning experience (Barr, 2019). Skills such as adaptability, and the factor of motivation, are crucial components in CT disposition, considering their definition as "consistent internal motivations" to exhibit certain behaviours (Facione, 2000).

The process whereby a videogame can teach skills while keeping a high level of motivation is elaborated in a seminal article by Malone (1980), which outlines the three concepts of goals, fantasy and curiosity. Goals are the objective of the game. A game goal which is clear and compelling is an intrinsic motivator because the player will be driven to complete it for its own sake. Fantasy refers to the concept of suspension of disbelief, when a game presents images and situations of any kind which are "not present" in reality (Malone, 1980). Malone divides the third concept, curiosity, into sensory and cognitive curiosity. Sensory curiosity is a technique to attract and retain the attention of players by changing the sensory stimuli of the game environment (Malone, 1980). Cognitive curiosity, however, is defined as "a desire to bring better form to one's knowledge structures" (Malone, 1980). This type of curiosity is a powerful intrinsic motivator because it relies on an instinctive need of the player to uncover information. The game designer can harness this by presenting information which challenges the existing game knowledge of the players – a technique that will be utilized here.

DEVELOPMENT

Game Content

The game described here is titled *Evenir Case Files*. In the game, the player takes on the role of a Keeper of Records serving the Mage Council of the fantasy kingdom of Evenir. The game is entirely text-based, the only illustrations present are menu icons. The first screen of the game introduces the player to the fantasy world of the game and explains their task. As Keeper of Records, the Council asks the player to investigate the sealing of an ancient witch and establish whether it was the correct thing to do, as the court historians of the king of Evenir claim. The first screen also hints that the player is descended from the witch, another reason why the Council selected the player specifically.

After that introduction, the game is played by navigating to different screens, representing locations in the world, which contain documents to be read. There is a timer, initially set at thirty days, and visiting a location as well as reading documents removes days from the timer. Once the player has no more days to act with, they must drag individual sentences from the documents they have read into three sections which require them to think about what that sentence contributes to their overall argument. Once the player presents their case, they are taken to a final screen where the Council makes a statement on the quality of their case. The rest of this section will show that the presentation and mechanics of the game are grounded in the literature this research project draws on, and are specifically constructed with the aim of developing CT skills in history.

The first element of the game worth discussing is its objective. Tosh and Lang (2006) note that it is very common for historians to have a specific question they want to address before they begin their research. Although ideally the professional historian should allow their sources to take them in an unexpected direction, the constraints of academia are such that it is far more common to have a question at the outset which remains unchanged throughout their work (Tosh and Lang 2006). As well as closely reflecting the actual practice of history, Malone (1980) also notes the importance of a clearly defined goal, such as a question to be answered, as an intrinsic motivator in a GBL context. Given the overlap of the two literatures, the authors decided to focus on a historical question that was clearly defined from the outset. Thus, as well as having their goal introduced in the first screen of *Evenir Case Files*, the player also is presented with the question they must answer.

The next important element of the game is its setting. Fictional worlds are typically not the purview of the historian. However, the *Evenir Case Files* was designed to focus on CT skills and dispositions, not history content skills. Setting the game in a context which was clearly divorced from real history was a deliberate design choice to prevent players from focussing excessively on the content, and conversely to avoid other players potentially feeling like they could not play the game properly due to their lack of content knowledge. From a game design perspective, this also allowed the authors to design scenarios without having to distort real historical events in order to satisfy game design requirements, a problem which a number of successful historical games have had to deal with (McCall, 2016).

Despite being fictional, the introductory screen of the game presents a detailed world. It establishes multiple factions in the King, the Council of Mages, and the player themselves, who is hinted to be connected to both. This, too, was deliberate. Fullerton (2004) notes that the premise of a game, defined as that which establishes “the action of the game within a setting or metaphor”, is key to enable the players to be emotionally invested in the events of a game. Furthermore, Fullerton (2004) argues that bringing together the “formal elements” of a game, that is its action, with the “dramatic elements” which include the premise, can improve the overall play experience. This is also supported by Malone (1980), who suggests that the goal of a game becomes more compelling, and thus more effective at teaching, when it is supported by the fantasy it presents. Furthermore, increasing player investment in the game by establishing a detailed setting as recommended by Fullerton (2004) increases player motivation through “cognitive curiosity” (Malone, 1980). This is because a player is more likely to want to fill gaps in their understanding of a fictional story if they care about its outcome. In the case of a game designed to strengthen CT, this should have the additional benefit of motivating the player to apply their CT skills in the context of the game, thereby increasing their disposition towards CT.

The next element of the game to examine is its action, the game mechanics. There are three locations in the game the player can visit, each containing a number of documents. The objective of the game, addressing the question in the Case page, is achieved through reading these documents, so this is the most important action in the game. The choice of documents the player encounters during the game is also a deliberate design decision. Tosh (2006) divides the types of sources available to the professional historian as “primary” and “secondary”, depending on their relative closeness to the events the historian is interested in studying. There are many categories within these two types, including recollections of past events, autobiographies, pieces of visual media and many more. Tosh

notes that each category is subject to a slightly different process of external and internal criticism, as different types of sources have different relative merit depending on the question the historian is trying to address.

The authors also sought to gain a better understanding of the relationship those learning to be practicing historians have with historical sources by surveying current History students at the host institution. These were a convenience sample of eight respondents from across different year groups. One respondent explicitly brought up the link between utilising primary sources well and developing CT skills, and three more respondents noted that the greatest challenge with using sources for them was understanding their context and possible biases. As a result, the authors deliberately sought to incorporate a range of sources in the game, to encourage that key CT skill of the historian: validating sources and weighing them against each other.

The final element of the game worth discussing at this stage is the timeline mechanic. Visiting one of the three locations, and reading a new document, subtracts in-game time from the player. They can only read as long as they have enough days to do so. There were two considerations which led to the implementation of this mechanic. Firstly, Tosh (2006) notes that the extent of an argument a historian may make is limited by the sources at their disposal. Modelling this real-life limitation by restricting the number of sources the player could access was one of the reasons for implementing the timeline mechanic. Secondly, there were also specific game design considerations to be made. Salen et al. (2004) note that a play experience is considered meaningful when there is a “discernible and integrated” relationship between player actions and game outcomes. In other words, players must be aware of what happens after they take an action and that an action will have repercussions later in the game. The timeline here serves both of these purposes: a player reading a document is made aware that their remaining time has diminished, and they know that as a result they have less time to read more documents, which makes their next choice more important than the last.

Game Development as Iterative Design

Though the elements of the game discussed in the previous section are grounded in academic literature and their substance did not change throughout the development process, that is not to say that specific details within those elements remained fixed. Salen & Zimmerman (2004) argue that it is crucial for a game to be developed in an incremental, iterative approach early on in its lifecycle in order to ensure the best possible final product.

Following a User-Centered System Design approach (Gulliksen, 2003), think-aloud evaluations were carried out, in which the user was asked to interact with the system and voice their thoughts out loud. As the user is asked to voice their thoughts in context, the feedback a designer can capture is quite detailed, uncovering insights which might otherwise have gone unvoiced. These think aloud evaluations lasted for thirty to forty minutes each, with an additional twenty minutes of follow-up questions based on a game-specific usability survey designed by Fullerton (2004).

Iteration 0: Paper Prototype

Following the advice of Fullerton (2004), the authors judged that creating a paper prototype would be an appropriate first step. The paper prototype was a collection of slides which together represented every possible screen in the game. The researcher would manually tab through these screens at the request of the evaluator.

This evaluation confirmed the appeal of the storyline, and the participant was able to successfully assemble their case, giving early evidence of the potential of the game to develop historical CT skills around questioning documents. However, the initial user interface (UI) sketches were thought to be confusing, and the lack of an automatic timer was identified as a problem. A software-based prototype, complete with an on-screen timer and revised UI, was therefore the next logical step.

Iteration 1: Software Prototype

The software prototype was developed to the point where all the functionality of the game was complete, before moving on to its visual presentation. The authors decided to retain the “medieval” aesthetic through the beige colours and serif fonts which can be seen in Figure 1, but otherwise opted for a more modern, “flat” interface that is the contemporary trend in web applications.

Iteration 2: User Experience Improvements

This iteration was the first set of think-aloud evaluations the authors conducted to improve the game, which yielded both reassurance that certain aspects of the game were strong and suggested directions for improvement. Participants noted that the story was the strongest aspect of the game. However, participants also noted that they did not recognize how aspects of the UI were intended to work: the operation of a slider, for example, was not well understood, and the clickable nature of certain areas of the screen was missed. These issues related to discoverability of features (Norman, 1988). The UI was therefore revised to denote all interactive elements with a specific blue colour, which would stand out from the non-interactive elements (Loranger, 2015).

Iteration 3: Further User Experience Improvements

Three new participants evaluated the game at this stage. There was, again, consistency in the remarks on what the game was doing well and on areas for improvement. All three participants noted that the story was the element which drew them in most, and that they found learning more about the world and the story as they read documents an appealing mechanic. Crucially, at this stage, none of the participants had significant issues navigating through the game. There was no hesitation as they clicked the now blue text anchors, for example. This suggested that the problem of discoverability had been addressed.

The Final Iteration

At this stage, the game was evaluated with one think-aloud evaluation and five remote evaluations, at the end of which the authors compiled the survey (based on Fullerton, 2004) in written form.

Figure 1. The Case screen of the Evenir Case Files game in its final iteration

My Case

Was sealing the witch Callisto necessary for the good of the kingdom?

Strongly disagree | Neutral | Strongly agree

My Documents

Read documents used to build your Case

- Extracts from King Richard's Letters
- Treasury Reports from 1 YoU - Present
- The Kingdom of Evenir, vols. 1-3
- Population Census 1 YoU - Present

Case Sections

Drag sentences from documents into one or more of the sections below

Challenge

Drop sentences here if you think your argument challenges them

Defend

Drop sentences here if you think your argument supports them

Emphasize

The feedback was mostly positive. The five remote participants all expressed a clear understanding of the objective of the game, ranging from “travel around to do research and produce a compelling argument” to more nuanced answers like “analyzing information to find contradictions in the story to potentially prove the innocent or guilt of the witch”. All respondents also stated that they had no issues with the controls of the game.

EVALUATION

One of the most widely employed tools to measure CT skills is the Watson-Glaser Critical Thinking Appraisal (WGCTA) (Bernard et al, 2008). Although the WGCTA has generally shown to be reliable under some statistical aspects (Bernard et al, 2008, Bawuens and Gerhard, 1987, Crouch 2015, and Hassan and Madhum, 2007), authors including Crouch (2015) and Hassan and Madhum (2007) find that it can struggle to act as a predictor of the ability to think critically in the future. Hassan and Madhum (2007) suggest this may be due to factors such as the difficulty with transferring CT skills from one domain to another. Because of the focus of this project on the specific domain of history, a study design based around the WGCTA was deemed inappropriate here. Gelerstein (2016) show that it is possible to develop an ad hoc instrument to measure CT skills based on the specific circumstances of the study. They validate their instrument through a variety of methods, including think-aloud evaluations and statistical tests such as Cronbach’s alpha (Gelerstein, 2016). However, their model of CT is similar to that which underlies the WGCTA, and the instrument therefore has similar limitations to the WGCTA when it comes to the problems of domain specificity.

The model at the core of this work is the Reflective Practitioner conceptualisation put forth by Schon (1984), so the authors sought to design a tool which would reflect this model. Crucial to this was the notion of “virtual worlds”. Schon defines virtual worlds as “contexts for experiment within which practitioners can suspend or control some of the everyday impediments to rigorous reflection-in-action” (Schon, 1984). This allows the practitioner to focus only on certain parts of the unique problem they face, which facilitates their ability to “see-as”, since it is easier to recognise a problem with fewer variables as something similar to an experience they have already had. Because “seeing-as” is the prerequisite for “doing-as” in the Reflective Practitioner model, a well-constructed virtual world is a fundamental problem-solving tool in the arsenal of the practitioner, and therefore to this study.

The authors developed the *Evenir Case Files* game as a virtual world for the historian, in that it excludes certain parts of the process of the historian outlined by Tosh and Lang (2006) from the gameplay. The process of external criticism, checking for the authenticity of the sources, is not part of the gameplay, for example. Instead, the game is entirely focussed on the process of internal criticism of historical sources. The aim here was thus to ensure the creation of a virtual world that faithfully mirrored this part of the process of the historian in an engaging way. Consequently, according to the Reflective Practitioner model, all that was necessary to evaluate our approach was to identify a way to check that the game was indeed encouraging its players to “think historically”, as outlined by Willingham (2008).

Based on the precedent of an effective ad-hoc instrument offered by Gelerstein (2016), the authors designed a set of questions that asked the participants to reflect on what they were doing as they were playing the game and why they took certain actions. These were administered in the form of a semi-structured interview right after the participant played through the game. Because of the difficulty of conducting in-person studies imposed by the pandemic, the evaluations were conducted over the videoconferencing platform Zoom. Full, anonymised transcripts of the study can be found at [data URL to be inserted]. Nine participants took part in this study. The participants were selected on the basis that they had not taken any university-level history courses prior to the study being conducted. This was done to limit the confounding effect that previous experience in subjecting historical sources to internal criticism would have had on the results. For the same reason, none of the participants had

seen or played through the *Evenir Case Files* game before. The questions were based on the process outlined by Tosh and Lang (2006).

DISCUSSION

Results were analysed along the lines of the grounded theory methodology detailed in Charmaz (2007). First, for each question, all nine responses were grouped together and viewed side-by-side. This allowed the authors to have an initial overview of similarities across participants. Then, categories were identified: broader themes which emerged from analysing responses across questions. This shortened form of grounded theory analysis is referred to as “affinity diagramming” in industry, and was selected because it is the qualitative analysis technique best suited to quickly identifying themes in the data (Rosala, 2019). The results are presented here in narrative form as recommended by Charmaz (2007).

Critically Assessing Bias

One of the key skills a historian must employ is dealing with bias in their sources (Tosh and Lang, 2006). This relates to the skills component of CT such as analysis and evaluation. Assessing the bias of the sources participants encountered in the game is one of the first emergent themes worth discussing here. All participants who encountered the Royal Palace location stated that the sources there were biased in some way. On the personal letters of King Richard, P1 stated “you have to take [them] with a bit of a pinch of salt, because they’d obviously be subjective in nature (...) in his mind he’s the good guy.” On the official court histories, P2 noted that they were likely “(...) pre-written books, extremely edited (...)”, while P3 stated they believed them to be “propaganda”. Considering the purpose a text was written for is the first step to critical analysis in history (Tosh and Lang, 2006). On the official court histories, P3 noted that they would have considered them a useful source if they could have found a document on the context or reason for which they were produced. The idea that a heavily biased source can still be useful to a historical argument is something which Tosh and Lang (2006) also note, and the fact that this player voiced such an opinion after playing the game suggests it has the potential to encourage advanced levels of historical thought.

All participants who encountered the Village Elder character stated that they considered their testimony to be more reliable than the documents in the Royal Palace location. The two reasons given is that a character like the Elder would have been “closer to” the event the player was investigating, and that the Elder would not have “anything to gain” from manipulating the truth. This first intuition of closeness to the event is supported by Tosh and Lang (2006), who note that the historian will often prefer a source as close as possible to the events. The second point on having something to gain by portraying events a certain way inadvertently echoes a long-standing debate in the literature on the nature of historical truth (Carr, 2018). Both points are further evidence of historical thought being encouraged.

Most participants who encountered the statistical type documents in the Treasury Archives location considered them to be reliable. P4 stated that unlike the other documents, which were “rumours or hearsay”, the statistical reports were the only ones which could be verified by counting, and therefore the most reliable. P3 was the only participant who strongly doubted the reliability of the documents in the Treasury Archives, noting that the documents were likely not “independently sourced, right, it came from the same government that benefits from those numbers looking good, so it makes sense that they’d want to use it.” Dealing with bias and errors in statistical data is also an important part of the work of the historian as outlined by Tosh and Lang (2006). Thus, the fact that at least a few participants extended their assessment of bias to the documents in the Treasury Archive shows another example of the players thinking historically.

The final theme concerning CT skills is the importance the participants attributed to the context around the documents they encountered. No participant felt like they had enough information to be confident that the position they took was correct. Many reported that they would have liked more information on their historical sources in order to feel more confident. P3 stated they would have liked “some more authorship information about like who wrote it, what’s their background, what else they published... and it’s like to look at who’s writing what”, a sentiment echoed by P5, who said “I don’t know who wrote these, I don’t think they present any kind of robust argument.” On the one hand, this could be seen as a critique of the way the game was designed. These participants felt that they did not have enough information to make meaningful choices as they were playing, which is one of the key tenets of good game design (Salen and Zimmerman, 2004). On the other hand, this is a positive sign for the objectives of the game. The players were not only thinking about the information that was presented to them, but also about what was not said and what might have been. This interpretation is also supported by examples in the literature. For example, Squire (2005) finds similarly nuanced discussion of biases in his seminal study of GBL in history with the game *Civilisation III*. It appears, then, that playing the *Evenir Case Files* game caused participants who had never taken a university-level history course before to display many of the same CT skills a practicing historian would in an academic setting.

Forming the Argument

Moving on from CT skills, Facione (2000) notes that an important dispositional component of CT is “self-correcting”, the propensity of a critical thinker to re-evaluate their opinion in light of new evidence. In the domain of history, Tosh and Lang (2006) note that a historian can take two approaches to utilising historical sources. They can start their investigation with a set question, or they can begin without a question in mind and let the sources offer a direction. The former is the most common, and what the *Evenir Case Files* game looks to model, but Tosh and Lang note that, ideally, a historian should take a mixed approach, remaining open to avenues they had not considered at the outset. The first result worth discussing from the evaluation concerns this aspect of self-correction in the relationship between the players and the documents of the game.

Most participants stated that they tried to be selective with the locations they visited and the documents they read, within the constraints imposed by the timeline mechanic of the game. P2, for example, started in the Royal Palace location of the game, but then “once I realised I was running out of days I decided to go through the, Callisto’s Village, because I thought there might be some contradictory testimony there”. There is evidence that the participants were re-evaluating their understanding of the game objective as they acquired new information and used their updated understanding to guide their subsequent actions. Not only is this in line with the dispositional component of CT noted by Facione (2000), it is also an excellent demonstration of reflection-in-action per the Reflective Practitioner model. Another participant acknowledged that their position on the question the game asked them to solve was likely biased on account of the first document in the game they encountered, the personal letters of King Richard, as it conditioned what sources they looked for next. This is a sophisticated understanding of the relationship between the historian and their source material, which Tosh and Lang (2006) note the practicing historian must guard against in the exploratory phase of their work.

An emergent theme which supports the idea that the *Evenir Case Files* game engaged player disposition to think critically in the domain of history was the variety of arguments the participants developed. Five participants took the position of somewhat disagreeing that sealing the witch Callisto was necessary for the good of the kingdom, the position the authors designed to be the most correct in terms of how highly the game rewarded players who selected it. However, four participants read the same material but still took a neutral position or even somewhat agreed that sealing the witch was necessary, based on the other evidence they found supported that position, or not arguing they did not have enough contrary evidence to disprove it. This was likely due to the authors applying the principle of cognitive curiosity (Malone, 1980) when designing the game, by revealing information

only in parts, forcing the players to look for more. As P7 noted, “when the game doesn’t really hold your hand and you have to make up the, sort of, guesses, it becomes a bit more difficult.” Thus, the participants offered a variety of arguments and ways of reasoning about them -- evidence that there was a critical relationship between their arguments and the sources the presented by the game.

The Reflective Practitioner Model and the Question of Transfer

The Reflective Practitioner model at the core of this work argues that practitioners draw on previous experience to solve new problems they encounter by relating new problems to ones they have already solved, or part of them, by controlling variables through virtual worlds. This is the reason that the authors can make any claim to addressing the problem of transfer in the first place. *The Evenir Case Files* faithfully reproduces part of the process of the historian, so completing the game means it becomes one of the already solved problems the players could draw upon in the future. That said, the ability of a player to do so depends on whether they are able to “see-as”, relating the new problem one already solved, in this case completing the game. This is the key limit of the Reflective Practitioner model. The authors believe that if a game like *The Evenir Case Files* were to be used effectively in a classroom setting, then care would have to be taken to relate more traditional forms of historical analysis to the mechanics of the game. This would be in line with the suggestion of McCall (2016) to integrate games in the history curriculum as opposed to trying to work around them.

CONCLUSION

Here, the domain of history was chosen as a representative microcosm of the broader problem of teaching CT. Historians must employ a range of CT-related skills when dealing with historical documents and have the disposition to employ those skills throughout their process (Tosh and Lang, 2006). Many previous studies focussed on the skills component of CT and neglected dispositions, which are much harder to teach (Facione, 2000 and Lai, 2011). These same studies also struggle to say much about the problem of transfer, that is whether students who successfully apply CT in one area will continue to do so in the future (Crouch, 2015 and Willingham, 2008). The novel approach taken here applied Schon’s Reflective Practitioner model to CT in the domain of history, while harnessing game-based learning principles to the problem, thereby addressing the often-neglected dispositional side of CT.

The objectives of the game were validated with a study design based on the Reflective Practitioner model (Schon, 1984). In the final evaluation, nine participants, who had not seen the game before and had not previously taken a university-level history course, were asked to play the game and then interviewed about their thought process and decision-making as they were playing. The results suggest that the *Evenir Case Files* game was indeed able to engage critical thinking skills in participants. They discussed the bias of their sources, including sophisticated ideas like the usefulness of biased sources to constructing an argument, the possible bias of statistical data, and the faultiness of memory. There is evidence that the game also engaged player disposition to think critically, beyond just skills. Many participants were able to reflect on their own actions and demonstrate that they were re-evaluating their position as they acquired new information. There is evidence that the game might also contribute to the transfer of CT skills. However, it would be wrong to make any exaggerated claims at this stage. The results of this study are limited, partially by the small sample size and homogeneity of the participants, partially because there is no evidence of transfer, only a reasonable inference based on the Reflective Practitioner model.

FUTURE WORK

From a research perspective, the evidence for the game developing CT skills and dispositions is limited. Conducting a longitudinal study might be one effective way to verify whether the Reflective

Practitioner model actually holds up to scrutiny in this context, as would attempting to integrate this game in as real educational curriculum. Exposing the game to a larger group of players with a wider variety of experiences would also be useful to strengthen the claims made here. A larger study might also make use of pre-/post-tests to assess gains in CT skills, or changes in dispositions.

From a game design perspective, the *Evenir Case Files* is still a prototype. The game can be played fully in twenty minutes, and there are still “rough edges”. Work remains to be done to improve the user experience. More detail could also be added to the game to cover more parts of the process of the historian. As it stands, the *Evenir Case Files* game has shown that combining the Reflective Practitioner model with game-based learning opens up a wide range of possibilities for designers of serious historical games to explore.

ACKNOWLEDGMENT

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

REFERENCES

- Barr, M. (2019). *Graduate Skills and Game-Based Learning: Using Video Games for Employability in Higher Education*. Springer International Publishing. doi:10.1007/978-3-030-27786-4
- Bauwens Eleanor, E., & Gerhard Gwenyth, G. (1987). The Use of the Watson-Glaser Critical Thinking Appraisal to Predict Success in a Baccalaureate Nursing Program. *The Journal of Nursing Education*, 26(7), 278–281. doi:10.3928/0148-4834-19870901-06 PMID:2822871
- Bernard, R. M., Zhang, D., Abrami, P. C., Sicol, F., Borokhovski, E., & Surkes, M. A. (2008). Exploring the structure of the Watson–Glaser Critical Thinking Appraisal: One scale or many subscales? *Thinking Skills and Creativity*, 3(1), 15–22. doi:10.1016/j.tsc.2007.11.001
- Carr, E. H. (2018). *What is History?* Penguin Classics.
- Charmaz, K., & Belgrave, L. L. (2015). Grounded Theory. In *The Blackwell Encyclopedia of Sociology*. American Cancer Society. doi:10.1002/9781405165518.wbeosg070.pub2
- Chowanda, A., & Chowanda, A. D. (2016). Gamification of Learning: Can Games Motivate Me to Learn History? *ComTech: Computer, Mathematics and Engineering Applications*, 7(3), 225. doi:10.21512/comtech.v7i3.2503
- Crouch, S. J. (2015). Predicting Success In Nursing Programs. *Journal of College Teaching and Learning*, 12(1), 45–54. doi:10.19030/tlc.v12i1.9069
- Davies, M. (2016, November 22). *What is critical thinking? And do universities really teach it?* The Conversation. <http://theconversation.com/what-is-critical-thinking-and-do-universities-really-teach-it-69046>
- Facione, P. A. (2000). The Disposition Toward Critical Thinking: Its Character, Measurement, and Relationship to Critical Thinking Skill. *Informal Logic*, 20(1). Advance online publication. doi:10.22329/il.v20i1.2254
- Facione, P. A. (n.d.). *A statement of expert consensus for purposes of educational assessment and instruction*. Academic Press.
- Fullerton, T., Swain, C., & Hoffman, S. (2004). *Game design workshop: Designing, prototyping, and playtesting games*. CMP .
- Gee, J. P. (2003). What video games have to teach us about learning and literacy. *Computers in Entertainment*, 1(1), 20. doi:10.1145/950566.950595
- Gelerstein, D., del Río, R., Nussbaum, M., Chiuminatto, P., & López, X. (2016a). Designing and implementing a test for measuring critical thinking in primary school. *Thinking Skills and Creativity*, 20, 40–49. doi:10.1016/j.tsc.2016.02.002
- Halpern, D. F. (1998). Teaching critical thinking for transfer across domains: Disposition, skills, structure training, and metacognitive monitoring. *The American Psychologist*, 53(4), 449–455. doi:10.1037/0003-066X.53.4.449 PMID:9572008
- Hassan, K. E., & Madhum, G. (2007). Validating the Watson Glaser Critical Thinking Appraisal. *Higher Education*, 54(3), 361–383. doi:10.1007/s10734-006-9002-z
- Hunt, E. (1995). *Will We Be Smart Enough?: A Cognitive Analysis of the Coming Workforce*. Russell Sage Foundation.
- Krath, J., Schurmann, L., & von Korflesch, H. F. (2021). Revealing the theoretical basis of gamification: A systematic review and analysis of theory in research on gamification, serious games and game-based learning. *Computers in Human Behavior*, 125, 125. doi:10.1016/j.chb.2021.106963
- Lai, E. R. (2011). Critical thinking: A literature review. *Critical Thinking*, 50.
- Loranger, H. (2015). *Beyond blue links: Making clickable elements recognizable*. <https://www.nngroup.com/articles/clickable-elements/>
- Malone, T. W. (1980). What makes things fun to learn? Heuristics for designing instructional computer games. *Proceedings of the 3rd ACM SIGSMALL Symposium and the First SIGPC Symposium on Small Systems*, 162–169. doi:10.1145/800088.802839

- McCall, J. (2013). *Gaming the Past: Using Video Games to Teach Secondary History*. Routledge., doi:10.4324/9780203831830
- McCall, J. (2016). Teaching History With Digital Historical Games: An Introduction to the Field and Best Practices. *Simulation & Gaming*, 47(4), 517–542. doi:10.1177/1046878116646693
- McLaughlin, A. C., & McGill, A. E. (2017). Explicitly Teaching Critical Thinking Skills in a History Course. *Science & Education*, 26(1), 93–105. doi:10.1007/s11191-017-9878-2
- Norman, D. A. (1988). *The psychology of everyday things*. Basic books.
- React – A JavaScript library for building user interfaces. (n.d.). Retrieved 23 March 2021, from <https://reactjs.org/>
- Reed, J. H., & Kromrey, J. D. (2001). Teaching critical thinking in a community college history course: Empirical evidence from infusing Paul's model. *College Student Journal*, 35(2), 201–201.
- Rosala, M. (2019). *How to analyze qualitative data from ux research: Thematic analysis*. <https://www.nngroup.com/articles/thematic-analysis/>
- Salen, K. T., & Zimmerman, E. (2004). *Rules of play: Game design fundamentals*. The MIT Press.
- Schön, D. A. (2017). *The Reflective Practitioner: How Professionals Think in Action*. Routledge., doi:10.4324/9781315237473
- Squire, K. (2005). Changing the game: What happens when video games enter the classroom? *Innovate. Journal of Online Education*, 1.
- Staton, B. (2021, January 22). Educators around world seek to take axe to exam-based learning. *Financial Times*. <https://www.ft.com/content/9d64e479-182c-4dbd-96fe-0c26272a5875>
- Tosh, J., & Lang, S. (2006). *The Pursuit of History: Aims, Methods and New Directions in the Study of Modern History*. Pearson Education.
- Venkataramakrishnan, S. (2020, December 26). 'Rudderless' QAnon may reinvent itself after US election, warn experts. <https://www.ft.com/content/357272df-46a1-4ed9-88b2-f5188d624e62>
- Willingham, D. T. (2008). Critical Thinking: Why Is It So Hard to Teach? *Arts Education Policy Review*, 109(4), 21–32. doi:10.3200/AEPR.109.4.21-32
- Yogev, E. (2013). On the need to strengthen political-critical thinking in history education. *International Review of Education*, 59(5), 627–645. doi:10.1007/s11159-013-9360-6

Ivo de Vero is a Computer Science and Economic & Social History graduate from the University of Glasgow. Currently an Innovation Consultant at ABBYY.

Matthew Barr is Programme Director for the Graduate Apprenticeship in Software Engineering at the University of Glasgow, and co-founder of the Ada Scotland Festival (ada.scot). He is co-Director of the University's Games & Gaming Lab, and serves as Vice Chair of British DiGRA and as a Trustee and Director of the Scottish Game Developers Association. Matt also sits on the BAFTA Scotland Committee and currently serves as the Games Jury Chair. His book, *Graduate Skills and Game-Based Learning*, was published by Palgrave in 2019.