# Developing Formative Assessment Designs on Evidence for A-Level History

**Celine Oon**

*Curriculum Planning and Development Division, Ministry of Education (Singapore)*

**Bertrand Tan**

*Curriculum Planning and Development Division, Ministry of Education (Singapore)*

## Abstract

*Pre-University History teachers often use A-Level History examination questions and general formative assessment strategies (e.g. questioning and student reflection) as formative assessments. Such approaches to formative assessment provide limited information about students' understanding of historical concepts and skills to inform teaching and learning. This article outlines the process of developing a formative assessment that assesses students' understanding of historical evidence. It uses ideas from the Stanford History Education Group's Historical Assessments of Thinking and the affordances of the Singapore Student Learning Space to expand the range of formative assessment tools available to teachers. The use of short assessments designed to make students' thinking visible provides teachers with valuable and timely information on students' learning to inform their teaching for deeper historical understanding.*

## Introduction

Identifying students' learning gaps is often a challenge for Pre-University History teachers. Besides generic formative assessment strategies such as teacher questioning, think-pair-share and student reflection, formative assessments carried out at the A-Levels also involve getting students to discuss or write essays in response to past year history examination questions. While these tasks provide teachers with some sense of how students are able to manage question items in the A-Level History examination, how much do these essays or Source-based Case Study (SBCS) assignments tell teachers about students' understanding of historical concepts and skills? [i] Furthermore, how helpful are these assignments in informing the next steps of instruction?

Generally, many pre-University history teachers recognize the value of formative assessment in supporting teaching and learning. Knowing where students 'are at' at significant junctures of the learning process can help teachers decide what to do to close students' learning gaps (Wiliam, 2011). However, in the absence of formative assessments that can be quickly implemented and targeted to elicit information on students' knowledge of historical concepts and skills, teachers often end up using summative assessment for formative purposes.

Yet, to meet formative assessment objectives, dealing mainly with A-Level History examination questions may have limited utility. The first issue is that lengthy essays make it difficult for teachers to

quickly identify particular skills or concepts that need further attention (Breakstone, 2014). The second issue relates closely to the purpose of the assessment. Specifically, A-Level History examination questions require students to synthesize component skills in the course of answering them. Yet, a student's response to A-Level History examination questions offers little or limited information on precisely where the student's strengths and weaknesses lie and do not serve as an effective compass pointing teachers towards appropriate instructional interventions. As put forth by the National Research Council (NRC) in *Knowing What Students Know: The Science and Design of Educational Assessment*, "…the more purposes a single assessment aims to serve, the more each purpose will be compromised" (National Research Council [NRC], 2001: 2).

Assessments designed specifically for formative use and targeting disciplinary learning outcomes in the A-Level History Syllabus could yield more useful information to impact teaching and learning, compared to the use of summative assessment tasks. Effective formative assessment tasks need to be (a) short, (b) carefully constructed to specific historical learning outcomes and (c) designed to make relevant student thinking visible (Breakstone, 2014). These assessments broaden the range of tools teachers can use to elicit information on students' learning.

With this objective in mind, we embarked on a small-scale project to develop formative assessments that would help A-level history teachers identify strengths and gaps in students' understanding of historical concepts and skills (see Appendix A). The project focused on assessing students' understanding of evidence – a historical concept fundamental to understanding how historical knowledge is constructed. This article outlines the assessment development approach, resources used and observations on this process of assessment design.

## Adapting History Assessments of Thinking (HATs) to the Local Context

The process of developing a specific formative assessment approach did not begin from scratch. Instead, it built on the significant work of the Stanford History Education Group (SHEG), which has developed a series of formative assessment tools on historical thinking. These assessments, known as History Assessments of Thinking (HATs), were designed to assess students' abilities in evaluating evidence. The HATs that were adapted for this formative assessment design focused on students' use of evidence, including contextualization and corroboration.[ii]

From the outset, two modifications were made to the original HATs.[iii] HATs typically require students to exhibit their understanding of evidence through historical knowledge. As the selected HATs were designed for a topic that was not in the A-Level History Syllabus, adapting the assessment involved melding syllabus content with the HAT's design features. In addition, the construct of close reading was incorporated into the assessment as teachers had indicated that this was a challenge for some students.

*Aligning the Assessment Construct with the Assessment Design*

Modifying the HATs made it necessary to verify if the formative assessment design was successful in eliciting information on students' understanding of evidence. This alignment between the assessment construct and the assessment design was achieved through (a) a strong understanding of the assessment construct (in this case, the

concept of evidence), (b) the observable behaviors expected from students who hold strong understandings of the assessment construct and (c) the corresponding assessment items that could generate those behaviors (NRC, 2001; Seixas, et al., 2015). [iv] The relationship between these three components contribute to the validity of the assessment in eliciting relevant information on students' understanding of evidence, and anchored the assessment development process.

*Incorporating Multiple Choice Questions into Formative Assessment*

Another adaptation to the HATs concerned the incorporation of Multiple-Choice Questions (MCQ) in the formative assessment. HATs do not contain MCQs because the question type does not provide opportunities for students to make thinking visible (Breakstone, 2014). MCQs, in assuming definitive answers, appear to run contrary to the nature of history which allows for different interpretations depending on how the historical issue and evidence are assessed (Breakstone, 2014; VanSledright, 2015). However, well-designed MCQs can provide teachers with quick information on students'

understanding and misunderstandings of historical concepts.

The MCQs developed as part of the formative assessment design recognized that the outcomes of historical thinking are not definite. Students are tasked to weigh the strengths of various options, and to select the option that was the "best" in the list of possible choices. Each MCQ item choice was also designed to reflect the strength of the respondent's understanding about evidence. For example, for Question 2 (see Figure 1), both the first and third options were designed to be reasonable responses to the question. However, the third option is a stronger explanation compared to the first because it reflected a more sophisticated understanding of the nature of historical evidence and an awareness of historians' treatment of sources. The first option, in contrast, suggests that sources can have a fixed property of being political and does not recognize the historian's role in analyzing sources as evidence. The MCQ choices did not simply involve identifying the correct option in a midst of wrong answers. Instead, they reflect a range of responses, from most-to-least defensible[v].

**Figure 1: MCQ Options for Question 2 of the Formative Assessment**



However, a public radio address may not accurately reveal all the considerations Truman had in mind when he made his decision. How would you decide with a greater degree of certainty, what was Truman's real reason for involving the US in the Korean conflict?

Q2: **Instruction**: Choose the option that you think best completes the statement.

**Truman's actual reasons for American involvement are difficult to figure out because...**

○ Public speeches are always made for political reasons and are not always useful as evidence.

○ The Korean War happened a long time ago and we cannot go back in time.

○ We only have one source and cannot conclude on Truman's claims one way or another.

Framing each MCQ item choice to reflect understandings or misconceptions about evidence enables the provision of specific feedback for each option selected by the student. For example, a student who chooses the option "The Korean War happened a long time ago and we cannot go back in time" likely has misconceptions about evidence and its relationship with the past. The student would subsequently receive feedback clarifying how analyzing sources allow us to make sense of the past even though we were not present when the Korean War took place. Developing these MCQs on the Singapore Students' Learning Space (SLS) platform has the added advantage of giving students such targeted feedback immediately through the automated feedback function. In addition to feedback on their answer choice, students will also be shown the "best option" and an explanation of why it may reflect a stronger understanding of historical evidence. This mechanism promotes metacognition as students reflect on the given feedback and perform self-correction.

Students' responses to the MCQs will also inform teaching. Students who were unable to identify the best answer in Question 1 (see Appendix A) probably ignored evidence from the source excerpt and the context of the American public's disillusionment with an adventurous foreign policy. This learning gap could be addressed by students observing the process of "close reading" of sources. Teachers could also engage students in groups to clarify, summarize, answer questions and/or make predictions from sources (Nokes, 2011).

Similarly, Question 2 seeks to gather evidence on students' understanding of the historian's role in analyzing sources as evidence. The option selected by students could reveal that they have a good understanding of how historians use sources to construct history (Option 3). It could also reflect their misconceptions about evidence, including confusion between history and the past (Option 2) or perceiving that the role of the historian is largely to read sources at face value (Option 1). To correct these misconceptions, students could be instructed on how historians approach sources as evidence or be involved in conducting authentic historical inquiries to deepen their appreciation of historical methods and processes (Nokes, 2011).

*Use of Short-Answer Questions as Formative Assessment*

While MCQs provide quick indications of students' understanding, short-answer questions make students' thinking in analyzing evidence visible. Compared to long essays, short-answer questions can be more targeted in assessing students' understanding of historical concepts and skills because it removes the cognitive load of conceptualizing and sustaining an argument. The short-answer questions adapted from the HAT example (on "Edison and the Kansas Housewife") further managed the cognitive demand of the formative assessment through its design. The questions provided short descriptions of sources for corroboration with the main source instead of source excerpts (see Figure 2). This was intended to free up students' cognitive resources to assess evidence rather than expend effort reading and inferring from additional excerpts.

**Figure 2: Managing Cognitive Load by Providing Key Information about Sources for Corroboration with Truman's Radio Address (Main Source)**



The short-answer questions provide students with a short description of three sources (see Figure 3). Students are asked to assess the utility of each source in providing more certainty about the reasons for US involvement in the Korean conflict indicated by Truman's radio address (the main source), and to explain why. The questions ascertain if students are able to compare information provided by two sources, identify their similarities and assess the extent to which one source is able to support the conclusions drawn from the other source. For example, a strong response to Question 3 (Figure 2) would recognize that Acheson's speech was not useful for corroboration with Truman's radio address. This is because the speech did not directly address the situation in Korea. Furthermore, the speech was made before North Korea's invasion of South Korea in June 1950, which influenced US foreign policy towards Korea and shaped Truman's radio address.

Each short-answer question requires a response of less than 100 words. Yet, they allow teachers to make several observations on students' understanding of evidence and ability to corroborate sources. Students' answers may demonstrate their unfamiliarity with the use of corroboration to analyze sources and construct knowledge in history. The answers could also indicate limited background knowledge, preventing students from identifying similarities between the sources. To address these learning gaps, students could revise how historians approach sources as evidence, observe the thinking process of corroboration or engage in authentic inquiries.

The choice of the three sources and their descriptions influence the strength of evidence that teachers can obtain from the short-answer question. For example, *NSC-68* was listed as a possible source for corroboration because of the document's significance on US actions in the Korean

War. In contrast, a description like "US National Security Documents from the late 1940s to 1950s" would be too vague and does not provide students with sufficient

focus to draw similarities with Truman's radio address. Such descriptions are likely to generate responses based on common sense and not historical understanding.

**Figure 3: Sources featured in the short-answer questions for students to corroborate to Truman's radio address**

> A. US National Security Council Planning Document 68 (NSC-68)
>
> B. Classified Soviet sources detailing interactions between Kim and Stalin prior to
>
> June 1950
>
> C. US Secretary of State Acheson's Jan 1950 speech outlining US commitment to
>
> Taiwan and Japan

The formative assessment that was developed comprised both MCQs and short answer questions. Both question types were included for a specific purpose in eliciting students' understanding of evidence. MCQs provide quick indicators of students' understanding, while short-answer questions provide opportunities to observe students' thinking. The affordances of SLS provide teachers with access to the answers of specific students and aggregated data of students' answers. The latter can reflect variations in students' understanding within a class or cohort. Teachers could also use the MCQ responses to identify which students' short-answer responses they want to focus on.

### A Model of Formative Assessment on Historical Evidence

Teachers can implement this formative assessment design in the course of teaching the Korean War, and when developing students' understanding of historical evidence and skills. The assessment also provides a design model that teachers can

use to develop similar formative assessments on evidence for other topics. These assessments should feature the following:

a) A meaningful inquiry question;
b) A rich source that allows students to make inferences, and presents a tension between the rhetoric of its creator and historical context;
c) MCQs that ask students to comprehend the source and make judgments on its usefulness; and
d) Short answer questions that require students to assess the usefulness of other sources in relation to the main source and explain their analysis.

### Conclusion

This article reflects a modest attempt to develop formative assessment tools for the Pre-University History classroom. It aims to help teachers make more accurate inferences about students' learning gaps and make better use of limited curriculum time. The assessment allows teachers to

observe students' understanding of evidence and ability in analyzing sources more quickly and more clearly compared to an SBCS or essay question. The use of such versions of formative assessment in the history classroom will allow teachers to make evidence-based decisions on whether they should focus on historical concepts, argumentative writing or other related aspects in their next lessons. The evidence could also offer indications as to whether students have reached an expected level of conceptual understanding. Teachers can then proceed more confidently to the next topic, and consider how to build on students' prior knowledge. Equally important is the value of such assessments in reminding teachers of the core learning outcomes of a curriculum that focuses on developing historical understanding.

## References

Breakstone, J. (2014). Try, Try, Try Again: The Process of Designing New History Assessments. *Theory & Research in Social Education, 42(4)*, 453-485. doi:10.1080/00933104.2014.965860.

National Research Council. (2001). *Knowing What Students Know: The Science and Design of Educational Assessment.* Washington, DC: The National Academies Press. https://doi.org/10.17226/10019.

Nokes, J. (2011). Recognizing and Addressing the Barriers to Adolescents' "Reading Like Historians". *The History Teacher, 44(3)*, 379-404. Retrieved from http://www.jstor.org/stable/41303991

Seixas, P., Gibson, L., & Ercikan, K. (2015). A Design Process for Assessing Historical Thinking: The Case of a One-Hour Test. In Ercikan, Kadriye & Seixas, Peter (Eds.) *New directions in assessing historical thinking (pp. 102-106).* New York, NY: Routledge.

Seixas, P. C., & Morton, T. (2013). *The big six historical thinking concepts*. Toronto: Nelson Education.

Stanford History Education Group. (n.d., a). Edison and the Kansas Housewife. Retrieved from https://sheg.stanford.edu/history-assessments/edison-and-kansas-housewife

Stanford History Education Group. (n.d., b). Opposition to the Philippine-American War. Retrieved from https://sheg.stanford.edu/history-assessments/opposition-philippine-american-war

Story, R., & Boller, P. F. (1984). *A more perfect union: Documents in U.S. history* (Vol. 2). Boston, MA: Houghton Mifflin Company.

VanSledright, B. (2015). Assessing for Learning in the History Classroom. In Ercikan, Kadriye & Seixas, Peter (Eds.) *New directions in assessing historical thinking (pp. 75-88).* New York, NY: Routledge.

Wiliam, D. (2018). *Embedded formative assessment*. Bloomington, IN: Solution Tree Press.

Wineburg, S. (1991). Historical problem solving: a study of the cognitive processes used in the evaluation of documentary and pictorial evidence. *Journal of Educational Psychology*, 83, 73–87

---

[i] The A-Level History curriculum seeks to develop students' historical understanding. Historical concepts and skills are central to this. The teaching and learning syllabus can be found at
https://www.moe.gov.sg/docs/default-source/document/education/syllabuses/humanities/files/2016-history-syllabus-(pre-university)-h2.pdf

[ii] Evaluating sources as evidence involves the processes of sourcing, contextualization and corroboration as explained in Wineburg (1991).

[iii] Our assessments adapted the following HATs: "Edison and Kansas Housewife" (Stanford History Education Group [SHEG], n.d.a) and "Opposition to the Philippine-American War" (SHEG, n.d.b).

[iv] In developing the formative assessment, Sexias and Morton's *The Big Six Historical Thinking Concepts* (2012) was used as the primary reference to define the concept of historical evidence and identify the behaviors exhibited by students with different levels of conceptual understanding.

[v] Our MCQ design took reference from VanSledright's (2015) exploration of weighted MCQs. However, we made an important distinction. While his assessment was intended for summative purposes and incorporated a scoring element, our focus was on formative assessment. We designed the MCQs to provide teachers and students with feedback on students' learning and did not incorporate a framework of scoring.