Continuous Assessment of Historical Knowledge and Competence

Challenges, Pitfalls, and Possibilities

KG HAMMARLUND

Introduction

It is not uncommon for history lecturers to find that assignments handed in for marking reflect students’ imagination of what the task is about, rather than the teacher’s hopes and expectations of what they have learnt. The gap between students’ and teachers’ perceptions of what it is to know history can be described as the gap between viewing history as a body of knowledge and as a form of knowledge (Shemilt, 1983). Although a body of knowledge – information, data, or ‘facts’ – is an indispensable prerequisite for developing historical knowledge, the latter – comprising the ability to handle information of the past and of understanding how pieces of information can relate to each other – is what transforms information into knowledge. Andreas Körber (2007) has suggested that historical knowledge can be seen as dependent on the competence to formulate historically relevant questions, to (re-)construct and de-construct historical narratives, and to make use of such narratives when orienting oneself in the present and for the future – signs of which are often lacking in essays and written assignments handed in by students.

The challenge of bridging the gap – or ‘decoding the students’ – has been discussed at length by Arlene Díaz, David Pace, and their colleagues at Indiana University (Díaz et al., 2008). That students lack a more developed understanding of what constitutes historical knowledge can probably partly be explained by their experience of school history – a factor which, alas, cannot be altered. What the
university lecturer can do, however, is to consider the form and content of lectures and seminars and choose reading materials (textbooks, articles, source material) that elucidate the ambiguity, the temporality, and thus the provisionality of our narratives of the past. And, of course, to construct assignment tasks that ask not for dates and names but offer the opportunity to demonstrate a more profound understanding of history.

Assessing whether and to what degree students possess a certain body of knowledge is a relatively straightforward business. Assessing whether they also possess a form of knowledge is not quite as simple. Bruce VanSledright (2013) points out that tests commonly given in US schools produce a narrow and biased gauge of students’ historical knowledge, since those tests fail to capture the ability to ‘do history’ or what VanSledright calls ‘strategic knowledge’. In Swedish schools the picture is very much the same, as shown by David Rosenlund (2011). Although the National Curriculum includes learning outcomes that deal with the thinking processes involved in historical understanding, those learning outcomes are rarely addressed in assignment tasks given in secondary schools. Instead, a disproportionately large part of the tasks focus on the reproduction of facts. However, Fredrik Alvén (2011) and Lars Andersson Hult (2012) have convincingly argued that it is possible to construct assignment tasks that give secondary school students the opportunity to display a more profound understanding of history, not least how our interpretations of the past affect our understanding of the present.

In the following I will discuss the potential of continuous assessment in higher education as compared with traditional assessment models, drawing on experiences presented in recent literature as well as experiences from introducing continuous assessment at Halmstad University. I finally point out the challenges that may arise from presenting students with unfamiliar assessment forms and the possible need for a ‘decoding process’ following the model developed by the History Learning Project at Indiana University.
History in Swedish higher education – learning outcomes and assessment

How, then, does Swedish higher education tackle the challenge? Over the last fifteen years, undergraduate history courses (or, at least, course syllabi) have undergone fundamental changes, with a marked shift from substantive towards procedural knowledge and ‘historical thinking’ already in first year modules.

At Uppsala University, the syllabus stresses the ability to reason constructively about historical issues and the role of history in society is thus among the learning outcomes (Uppsala University 2014). At Lund University, students are expected to ‘formulate historically interesting questions’ (Lund University 2014), while at Linköping University they should be able to ‘create explanatory narratives built on a critical analysis of relevant facts’ (Linköping University 2014). At Halmstad University, students are expected to gain an ‘understanding of historical contexts and how contexts may be defined differently depending on the perspective chosen’ (Halmstad University 2014). Within academic history, knowledge is clearly understood as consisting of more than the acquisition of a wide range of information. The ability to use this information is of crucial importance. How, then, is this aspect of historical knowledge – ‘strategic’ or ‘procedural’ knowledge – assessed?

The dominant end-of-module assessment in Sweden, well-known to students who have passed through the system during the last four decades or so, has been in the form of a written exam, to be completed in 3–4 hours with no books or notes allowed during the exam. It is often made up of 10–12 shorter questions, worth 2 or 3 points each, and 4–5 longer ‘essay questions’, worth 10 points each.

A typical short question could be ‘When and how did Sweden lose Finland?’, and the expected answer something like ‘1809, having lost the war against Russia’, giving the year (1809, 1 point), what happened (lost the war, 1 point) and the actor involved (Russia, 1 point). An essay question would, on the face of it, look as if it asked for a much more complex understanding of historical change and/or continuity, e.g. ‘Give a detailed summary of political change in England 1640–1688’. However, if we take into account the conditions...
under which the students work out their answer: limited time (maybe half an hour if all questions are to be addressed) and no access to notes or literature, it is clear that the answer given will not display the student’s ability to search, find and assess information from various sources, or to compare and evaluate contrasting or conflicting narratives. Instead, we may expect an answer that faithfully and exhaustively reproduces the narrative given in the textbook and has been dutifully memorised by the student.

Even if this type of exam question still exists, it no longer dominates as it did twenty-five years ago. Today, students engage in a variety of assignment tasks that allow them to use books, notes, and relevant source documents, and are more often aimed towards procedural or strategic knowledge. End-of-module assessment still dominates, however, which also means that a number of problems related to the single, high-stake assessment form remain. Students often study to the test, and with end-of-module assessment the given touchstone for relevance and significance will be ‘will it come up in the exam?’ Students may skip lectures and seminars deemed irrelevant for this assignment task (cf. Ludvigsson, 2012: 69). And even if the assessment may contain formative qualities (‘for the future it would be great if you could...’), its summative character will dominate. A possible way of circumventing those problems is offered by continuous assessment.

Reported experiences of continuous assessment

There are a number of articles and reports on experiences and outcomes of models for continuous assessment. Most of them, however, build on experiences from trying out continuous assessment in one course during one semester. Longitudinal studies or studies involving more than one department or institution are rare, and the conclusions that can be drawn are therefore, at best, preliminary and tentative. The number of common and recurring traits across the various studies nevertheless suggests that continuous assessment actually improves student learning. Although it cannot be ascertained that they learn more than from a course with traditional assessment
methods, it does seem as if they develop a deeper understanding of their own learning. It is thus the learning experience, rather than the learning in itself, the learning 'as such', that is affected and stimulated by continuous development.

Jennifer Frost, Genevieve de Pont, and Ian Brailsford (2012) introduced continuous assessment in a course on the history of African-American freedom struggles given at the University of Auckland in 2008. Here the continuous assessment took the form of shorter assignments, one per week, linked to the weekly tutorials. Some assignments were completed during the (group) tutorials, while others required that students brought shorter texts to the tutorial where they received instant feedback from their tutor and from fellow students.

From the viewpoint of lecturers and tutors, the model had one obvious drawback: it increased their overall workload, and time allocated for grading was exceeded by 20 percent. The big advantage, however, was that the instructors could follow the process of student learning; the how, when, and why of advancement. The continuous assessment helped the students to develop a familiarity and facility not only with the subject matter dealt with but also (and maybe more important) the conceptual framework of the course. The model also improved student attendance at tutorials. In most courses at the University of Auckland, students are neither required to attend, nor rewarded for attending tutorials with the result that attendance can drop down to 25 percent. In this course, however, 81 percent of the students attended at least 80 percent of the tutorials. Also, during the tutorials students were less inhibited and spoke more freely. Since they had been encouraged to reflect as part of their preparation they could easily form reasoned opinions, whereas in courses where they encountered such questions in the tutorial itself they had to do their thinking and orient themselves during tutorial time. Discussions, and student learning through them, thus benefited from the assessment model.

Student views were positive. In the evaluation at the end of the course almost one third of the students singled out continuous assessment as the most helpful factor. A typical remark was 'The
tutorial assignments really forced me to keep reading in-depth, to keep reflecting, so I feel more prepared’. A small group – less than 5 percent – was critical of the assignments, finding them repetitive and stressful. Continuous assessment was experienced as being under continual assessment.

A similar model was tried out in a BSc module on Business Taxation at a British University (Trotter, 2006). Continuous assessment was built upon ‘tutorial files’, i.e. a carefully prepared set of texts that students were expected to work with, analyse/comment, and finally to hand in a report at the tutorial. Again, continuous assessment tended to enhance student activity throughout the course. As one student stated in a follow-up interview: ‘It changed your behaviour ’cause instead of leaving stuff to the last minute and not doing any work through the semester I was working constantly’.

However, the students also admitted that they probably would not have spent so much time on the tasks if their work had not contributed to their final grade. Trotter’s conclusion is that this made the students more inclined to consider their tutor’s comments: Students who had done well tried to keep up the standard of their work, and those who had performed less well than anticipated strived to improve their result.

Similarly overwhelmingly positive experiences of continuous assessment have been reported by, among others, Sven Isaksson (2008), Jorge Pérez-Martínez et al. (2009), and Naomi Holmes (2014).

It should be noted that in all these cases the assignments that were continually assessed had limited weight in relation to the final grade. Lurking at the end of the course was a larger assignment, often some kind of final essay, perceived by the students as a high-stake task. While there are indications that the continuously assessed tasks actually help the students to perform well in the final assignment, one cannot disregard the risk that students tend to view the smaller tasks as less important. If the links between the two types of tasks are obscure, so that the students cannot see how the feedback received will help them to tackle the final assignment, the value of continuous assessment remains limited.
At Indiana University, historian Andrew M. Koke took the innovative step of replacing every high-stake, end of the course assignment with a large number of low-stake assignment tasks, given continuously throughout the course (Koke, 2011a, b). His main objective was to steer clear of what he saw as a drawback of high-stake assignments: they give very little room for students to experiment or try out new and unfamiliar lines of reasoning, since stumbling may jeopardise the student’s final grade. An assessment model built on high-stake assignments is, in Koke’s words, a model that ‘punishes failure’.

Koke used a broad variety of assignments – one-minute sentences, quizzes, oral presentations, research, and tests – 45 (!) in total. This meant that students had to produce something for every class, two-three times a week. On the other hand, failing one assignment (or a few) was not disastrous – there were plenty of opportunities to hand in assessments that could offset a few failures. Students who failed an assignment were also offered a second opportunity to submit it.

From the students’ point of view the model was well received. 73 percent preferred many small assignments over a few larger. ‘It helps me to not forget what I have learnt’, was one telling comment. ‘3–6 assignments would put more pressure on you because you have to get everything right’ was another. 93 percent of the students said that the rewrite policy (second effort) was a good one.

Did the model also improve learning? Koke found that the students learnt at least the same. However, their experience, as expressed in the evaluation process, was that they had learnt more than they had expected. They also found themselves having improved competencies beyond a narrowly defined field of historical knowledge. A significant number of students stated that the model had helped them improve both their writing skills and their reading comprehension. Students were able to develop an incipient meta-understanding of their own learning. They also became less anxious to search for the kind of answers they believed that the teacher expected, allowing them to elaborate on what they themselves had found interesting and/or important.
As in the Auckland case, the flip side was that the workload increased significantly. Koke had to handle, comment, and grade close to 900 assignments. Even if some of the tasks were small and could be commented upon briefly, grading was an ever-present task that had to be done during each and every week of the course.

As a tentative summary, the aforementioned reports indicate that students appreciate continuous assessment since it helps them to keep focus throughout the course and not postpone reading until the last week, that it allows them to deal with course material in a more reflective way, and, finally, that it helps them to shift focus from what they believe is the expected learning outcome towards the learning process itself, thus developing a meta-understanding of their own learning. From a teacher viewpoint, these gains are, of course, most valuable. However, documented experience reminds us that the gains in terms of student learning must be weighed against the increasing workload.

There are other important benefits, not mentioned in the reports referred to above. If properly designed, the assignment task will make it apparent that the assessment is not just about giving an account of what has been learnt previously but part of the learning process. And since continuous assessment consists of a large number of varied tasks they can each focus on a different aspect of knowledge and competence: one assignment may focus on criteria of significance, another on judging and evaluating claims and conclusions. All aspects of knowing and understanding history need not be crammed into one single final task.

Continuous assessment at Halmstad University: Examples and experiences

At Halmstad University, my colleagues and I have over the last two years modified the assessment of history courses, gradually moving from a traditional final test/essay model towards continuous assessment. The ambition is to set weekly assignment tasks that address ‘proper’ competency of historical thinking, thus creating opportunities for students to train and display their capability in
handling analysis, interrogation, interpretation, and argumentation. Normally the tasks given at the beginning of a module are smaller and ‘easier’ than the one given at the end, for which the students might be allowed more time (two weeks) for completion. As in Andrew Koke’s model, unsatisfactory efforts can be rewritten. Also, even if failure might affect the final grade, a single failure might well be offset by other tasks that are carried out more successfully.

A particularly successful task, given to first-year students, deals with demographic change during the eighteenth and nineteenth centuries. Using local census records from their own home parishes for the period 1750–1850, students are asked to compare birth/mortality rates, household size, age distribution and means of support (e.g. the number and proportion of landowners, tenant farmers, and sharecroppers). Students then compare their findings – similarities and differences over time – to the established view of changes in demography, family patterns, socio-economic strata etc.

Possibly the most rewarding experience has been to note the eagerness with which students have unearthed information on their own local history. Immigrant students are no exception from this pattern. They can, for obvious reasons, not work with census records from Iraq or Bosnia – they have to choose the Swedish parish they regard as their home, where they spent most of their youth and formative years and where their parents live. They are just as enthusiastic as their native Swedish course mates, sometimes even more, to gain a fuller view and a deeper understanding of their local history.

Part of the intended outcome of this task is of course to make students more confident in analysing unfamiliar source material. Another is the conviction that the relation between generalised (and thus more or less ‘abstract’) concepts is easier to grasp if this relationship, or parts of it, becomes apparent on a micro level. In this case, students hopefully gain a deeper understanding of the generalised statement that a growing agrarian proletariat is one factor that must be taken into account when explaining the industrial revolution when they can see how population, land ownership and living conditions changes in a specific parish. As a bonus, this
localisation of context not only in time but also in space apparently triggers their curiosity and makes them feel that the task was worthwhile; the more so since they have a personal relation to the area in question.

In another assignment students are asked to read a selection of primary sources – letters, reports and articles – stemming from Paris and the month of July 1789. Among the sources is an extract from a report from Thomas Jefferson to John Jay, US Secretary of Foreign Affairs, a letter from a M. de Sevelinge, banker to the Marquess de Lostanges (illegitimate cousin of king Louis XVI), and a few issues of The London Gazette. Besides being an exercise in careful reading and interpretation of primary sources (taking into account who wrote and who received the narratives), the assignment is also designed as an attempt to make the students think about significance. No course in modern history, whether in secondary school or university, leaves out the French revolution – its significance is beyond doubt. But was the significance of the events obvious for those who, although they had experienced (or heard of) the storming of the Bastille, hardly could foresee the abolishment of the French monarchy, the revolutionary wars, the reign of terror and the rise of Napoleon?

Many students have been surprised, even astonished, to find that there were people – informed and educated people – living in Paris who apparently did not find the situation (including events such as the storming of the Bastille and the lynching of its governor) particularly alarming or significant, but instead expected that ‘business as usual’ soon would resume. Students have commented on the heuristic experience of realising that the French revolution, as seen from the streets of Paris in 1789, was not the well-rounded, coherent, and unequivocal narrative offered by their old school textbooks. Rather, it was a kaleidoscopic array of conflicting and contradictory stories, some of them depicting the events as almost trivial. However, not all students are comfortable having their view of the textbook as the source of ‘what really happened’ questioned. Although seldom explicitly stated, it was clear that they preferred indisputable facts over tentative discussions on how different viewpoints may result in different perceptions.
A third example, still under development, is an inverted version of Peter Lee’s and Denis Shemilt’s synoptic frameworks (Lee & Howson, 2009; Shemilt, 2009). Instead of starting with a broadly sketched outline or framework that gradually becomes intelligible by being filled with details, the starting point is a ‘synoptic skeleton’, an individual life trajectory as depicted in an obituary. The obituary in question describes the life of a man who was born out of wedlock in 1913 and was raised as foster-child in a stone mason’s family. As a young man he found his way to the temperance movement, the trade unions, and the Social Democratic Party. His professional career went from the stone mason’s yard to the news desks of various newspapers and journals affiliated to the labour movement and eventually to the position of Assistant Under-Secretary at the Ministry of Health and Social Affairs. This life trajectory, used as a recurrent Leitmotif for a module on twentieth-century history, may show how ‘big pictures’ are intertwined with ‘little stories’ and facilitate an understanding, as Alison Kitson has put it, ‘that the past did actually happen and that the people in it were real’ (Kitson, 2004: 2).

Formative or summative: a remaining challenge

The experiences of myself and my colleagues at Halmstad University during 2013 and 2014 are similar to those of others discussed above. Attendance at lectures has improved, almost all students work conscientiously and finish their tasks on time, and quite a few students suggest that the model helps them to focus and also that they enjoy doing ‘real history’, not just reading textbooks over and over again preparing for a final test. There is, however, room for improvement. An issue that remains to be resolved is how to strike a balance between formative assessment (feedback) and summative assessment (marking). Hitherto, students have received feedback on each assignment but no marks. At the end of the module, students receive their grade, which rests on all the assignments weighed together. Although strengths and weaknesses have been pointed out in the weekly feedback, students have no straightforward guidelines
to tell them if their work is ‘good enough’ or ‘hits the mark’. Some students are perfectly happy with this. Others complain – most likely out of insecurity.

There are no easy solutions to this problem. Partly it stems from the fact that historical knowledge is a multifaceted entity, and this complexity is reflected in the learning outcomes for the course. It is nigh impossible to lay down grading criteria so clear-cut and detailed that you can tick them off when reading a text where a student discusses Thomas Jefferson’s views on what happened in Paris in July 1789, using a letter that Jefferson wrote at that very time. The undesired consequence is that what lies behind the grade, (even if the grade is, in a broad sense, supported by existing grading criteria) remains obscure. It can be argued that tasks constructed to be more ‘authentic’ than a multiple choice test lead to a higher degree of validity in assessment, since they allow the examiner to evaluate the student’s ability for historical thinking, but the flip side is that this also leads to less reliability, or at least to a lower degree of transparency and predictability. If the summative assessment lacks transparency, students will find it difficult to see how formative feedback and summative grade relate to each other, and it can be expected that the feedback will have limited effect on their learning (for a further discussion on the problem of supportive assessment of complex tasks, see Jönsson, 2012: 29-39).

Since continuous assessment tends to muddle the distinction between the formative and the summative, this is an issue that requires careful consideration. Doubt has been expressed whether formative and summative assessment – feedback and grading – can co-exist in an assessment model. Although not touched upon in the articles cited above, the literature on assessment does suggest that this is not an easy relationship. Rosario Hernández (2012) discusses this in an article presenting the results from a study of assessment models at language departments at eight universities in the Republic of Ireland. He cites a number of scholars who argue that the summative element will always be perceived as the most important, so that the students notice their grade but forget about the formative feedback or regard it as nothing but a justification for the received mark. In search of a
more balanced approach Hernández cites Maddalena Taras (2006, 2008) who suggests that the summative/formative dichotomy is questionable and that formative assessment in reality consists of a summative assessment plus feedback. She prefers the concept 'learning-oriented assessment' where the focus shifts towards the appropriateness of the assignment tasks so that tasks are created that focus on the students’ learning, encouraging and supporting them to autonomously monitor their own learning. Both the task in itself and the feedback are important if this is to be achieved.

What needs to be addressed in order to take full advantage of the benefits of continuous assessment is a number of ‘decoding challenges’, a concept described by Joan Middendorf, David Pace, and their colleagues (Middendorf & Pace, 2004; Díaz et al., 2008; Middendorf et al., 2014). Even if almost all students complete all assignments and a very small number fail, it is clear that while a fairly large group finds the tasks inspiring and does very well, some students find it difficult to go beyond a literal comprehension of the task, meaning that they conscientiously produce exactly what is asked for, neither more nor less. They have, undoubtedly, understood the task and their assignments go far beyond a mechanical copying of ‘facts’, but it is questionable whether they have also seen the point of it. One of the benefits of continuous assessment – developing of a meta-understanding of one’s own learning – is then lost.

Gauging students’ experiences through Supplemental Instruction

For the upcoming academic year (2015/16), at Halmstad University we will therefore introduce ‘Supplemental Instruction’ (SI) as a supportive environment for history students (for a short description of SI, see http://sac.indiana.edu/programsservices/supplementalinstruction). In its most common form, SI consists of weekly voluntary meetings where students can ask questions and discuss issues under the leadership of a fellow student who has previously completed the same course.
Since students can be expected to discuss more freely, feeling less anxious about disclosing their shortcomings or difficulties when no teacher is present, the student leaders will probably gain a lot of information that can throw some light upon what it is that students find difficult. If the SI meetings are supplemented with regular and frequent meetings between the student leaders and the course leader, this information may help to identify ‘bottlenecks’ and, as the next logical step, to develop and improve both the content and structure of the coursework.

If students are given weekly assignment tasks and SI meetings are held weekly, it is fairly safe to predict that discussions during SI meetings will circle around the assignment. Thanks to a formalised structure for feedback from students to the course leader, filtered and condensed by the SI leaders, this model will therefore also be helpful in evaluating the assessment model. This will, in turn, give the course leader a clearer picture of both strengths and weaknesses with the task given. It is hoped that this can lead to being able to reap the full benefit of the strengths inherent in continuous assessment.

References

Literature


Perspectives on University History Teaching, Uppsala: Opuscula Historica Upsaliensia, no. 48, pp. 63-93.


**Web resources**
Halmstad University (2014) ‘Kursplan Historia (1-30)’
http://utbildning.hh.se/se_proxy/utb_kursplan.asp?kurskod=HI1002&revisionsnr=15,1&format=pdf

Linköping University (2014) ‘Syllabus History 1’
http://kdb-5.liu.se/liu/fil/kp_detail_print_en.lasso?&ID=201344

Lund University (2014) ‘Kursplan Historia: Grundkurs’
http://kursplaner.lu.se/svenska/HISA13.pdf

Uppsala University (2014) ‘Kursplan för Historia A’
http://www.uu.se/node701?kKod=5HA302