Supporting learning through epistemic scaffolds embedded in a highlighter tool

Jan Erik Dahl

Department of Education
University of Oslo
Email: j.e.dahl@iped.uio.no

Abstract

This article explores the use of epistemic scaffolds embedded in a digital highlighter tool that was used to support students' readings and discussions of research articles. The use of annotation technologies in education is increasing, and annotations can play a wide variety of epistemic roles; e.g., they can facilitate a deeper level of engagement, support critical thinking, develop cognitive and metacognitive skills and introduce practices that can support knowledge building and independent learning. However, research has shown that the actual tool use often deviates from the underlying knowledge model in the tools. Hence, the situated and mediated nature of these tools is still poorly understood. Research also tends to study the tools as a passed on resource rather than being co-constructed between students and teachers. The researcher argues that approaching these resources as co-constructed can be more productive and can create new spaces for teacher–student dialogues, students' agency and self-scaffolding.

Keywords: Social annotation technology, mediation, scaffolding, co-construction/co-learning, design-based research.

Introduction

This paper is based on a study of a master's course in technology-enhanced learning. The central artefact explored in this study was an annotation highlighter tool. As Novak, Razzouk, and Johnson (2012) pointed out, social annotation technology is an emerging educational technology, but it has not yet been extensively used and examined in education. They called for more studies on the effect of annotation tools, which is a legitimate request. However, insufficient understanding of the situated and mediated nature of these tools complicates this endeavour. Particularly if we see these tools not as isolated artefacts (a *tool in itself*), but rather as an integrated part of continuously *evolving practices* that are co-constructed and result from the interaction between the tool, the task, the learner, the teacher and institutional factors that influence this practice. The article wishes to investigate the

situated interaction and how it affects the annotation practices, and the analysis will be based on a sociocultural stance (Vygotsky, 1978; Wertsch, 1991).

Research question

The paper first describes how the tool adapted to the evolving practices and some issues that prompted these changes. Second, it will explore students' annotation practices. Finally, the paper investigates the annotation practices through the analytical lenses of scaffolding. This will further the discussion of what epistemic role annotations can play and some of the obstacles, challenges and opportunities to consider when moving ahead—which ultimately also impact the question regarding the 'effect' of annotation technologies. The research questions are as follows:

1. How did the annotation script evolve in response to the annotation practices?
2. How did the tool mediate the annotation practices broadly speaking, and with regard to a) cognitive scaffolding and b) metacognitive scaffolding in particular?

A conjecture map has been created to elucidate the design ideas behind the experiment, and these conjectures will provide further guidance to the analysis of the research questions.

Tool and theoretical stance

The research–teaching nexus and co-learning

The studied master's course had a strong focus on research as part of the teaching and learning processes, and the researcher thus tried to adapt the design interventions to this focus. Such a focus constitutes a complex learning environment that potentially alters both the relation between students and teacher/researcher (students become *co-researchers* and researchers become *co-learners*), and changes the focus of the learning from content knowledge to engagement in research problems and the knowledge-building processes. To help explore these issues, the researcher drew in Dahl (2016) on Healey's (2005) distinction between four different ways of perceiving the research–teaching nexus (cf. Figure 1).
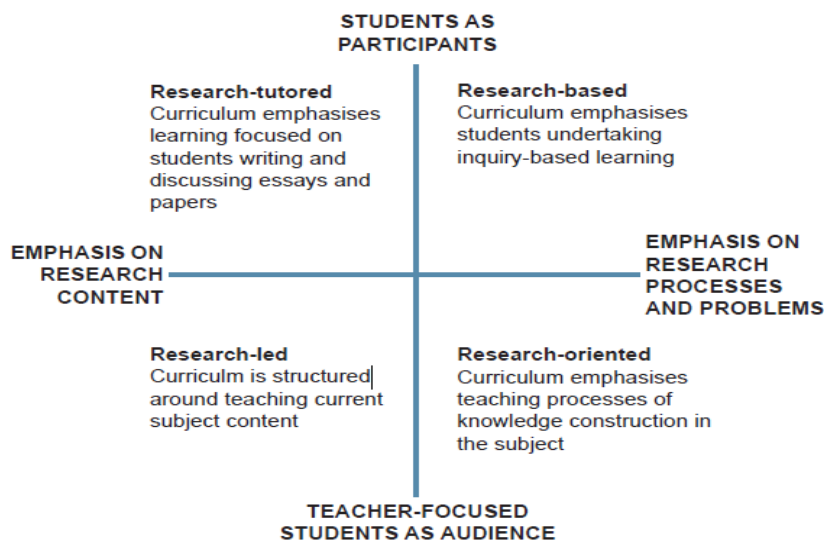

STUDENTS AS PARTICIPANTS

Research-tutored
Curriculum emphasises learning focused on students writing and discussing essays and papers

Research-based
Curriculum emphasises students undertaking inquiry-based learning

EMPHASIS ON RESEARCH CONTENT

EMPHASIS ON RESEARCH PROCESSES AND PROBLEMS

Research-led
Curriculm is structured around teaching current subject content

Research-oriented
Curriculum emphasises teaching processes of knowledge construction in the subject

TEACHER-FOCUSED STUDENTS AS AUDIENCE

Figure 1. Curriculum design and the research–teaching nexus (Healey, 2005, p. 70).

Healey's conceptualisation of the research–teaching nexus provides a relevant background for the design conjectures and how the tool is part of researcher/teacher–student co-learning and co-construction.

Annotation technologies

In Dahl (2016) the researcher reviewed related technologies and emphasised 1) standard annotation tools related to reading of documents (Ovsiannikov, Arbib, & McNeill, 1999; Wolfe & Neuwirth, 2001) and 2) knowledge-building environments that rely on annotation technologies more in the context of scaffolding students' writing processes as part of collaborative inquiries modelled after researchers' knowledge-building processes (Muukkonen, Hakkarainen, & Lakkala, 1999; Scardamalia & Bereiter, 2006):

1) The conclusion drawn from the reviewed annotation studies in Dahl (2016) was that annotation could serve a wide variety of purposes (Novak et al., 2012; Wolfe & Neuwirth, 2001), also depending on the ingenuity of the users, but that more research was needed to contextualise the use conditions (Ben-Yehudah & Eshet-Alkalai, 2014; Razon, Turner, Johnson, Arsal, & Tenenbaum, 2012; Schugar, Schugar, & Penny, 2011). Johnson, Archibald, and Tenenbaum (2010) have emphasised the need to shed more light on the intermediate processes (e.g. students' annotation practices) and how this intermediate level affects outcome measures (e.g. reading comprehension, critical thinking or metacognition). They proposed design-based research (DBR) for further exploration of their findings. DBR has generally not been applied in traditional annotation studies. One exception is Samuel, Kim, and Johnson (2011). Design-based research can help detect the contingent relationship between annotation and learning, especially if combined with methods for microanalysis of students' annotation practices. This study applies DBR, and hence takes advantage of the potential identified by Johnson and colleagues (Johnson et al., 2010; Samuel et al., 2011). Additionally, it incorporates microanalysis to study the situated and mediated nature of the students' annotation practices. This sets the study apart from Samuel et al. (2011) who based their DBR on a survey study.

The review in Dahl (2016) also found that the sociocultural history of the annotation resource was missing in the annotation studies. For instance, were the tool and the annotation practices *passed on* from teacher to students, or were they in some way a *co-constructed* resource, and if so, how were they co-constructed? The studies showed that annotations have a double role that can be quite problematic. They are strongly connected to the students' personal habits and preferences (Marshall, 1997, 1998). However, they are also subject to the teachers' interventions due to the value annotation practices are perceived to have for learning processes (Sung, Hwang, Liu, & Chiu, 2014; Yang, Yu, & Sun, 2013). This tension is likely to deeply affect the outcome of the reviewed annotation studies. The study will shed further light on this tension and the co-constructed nature of annotation resources, which is rarely discussed in the traditional annotation studies.

2) The conclusion in Dahl (2016) regarding the use of annotations as category scaffolds in *knowledge-building environments* was that these studies have paid more attention to the sociocultural dimensions of these tools. This is particularly true for the critical follow-up studies which, for instance, have shown that the actual tool use often deviates from the underlying knowledge models in the tools and have explored these practices (Arnseth & Säljö, 2007; Ludvigsen, 2012; Ludvigsen & Mørch, 2003). However, the students have been more the object of the focus rather than participating in this inquiry—the previous studies in this field have studied students' *tool use* rather than students' *research* on the tool they use. Hence, Dahl (2016) concluded that these studies also tend to perceive the tool more as a *passed on* resource,

rather than seeing them from the perspective of being a co-constructed product of the interaction between the students and the teacher/researcher.

The researcher argues that the distinction between a passed on and a co-constructed resource becomes particularly important when we want to scaffold self-regulated learning and therefore need to engage the students in their own scaffolding and in reflections on how the scaffolding they experience works (Ben-Yehudah & Eshet-Alkalai, 2014; Chen & Huang, 2014).

Sociocultural approach

The sociocultural perspective grew out of a discontent with behavioural and cognitive research, particularly how these approaches failed to account for the situated and mediated character of learning (Rasmussen & Ludvigsen, 2010). Hence, the unit of analysis in this perspective addresses learning as situated in particular social, historical, cultural and institutional settings, and as being mediated by language and material artefacts deriving from these settings (Vygotsky, 1978; Wertsch, 1991). Rasmussen and Ludvigsen (2010, p. 403) emphasise that 'to understand learning, we need to capture how humans interact and make use of different kinds of cultural tools in different kinds of settings, or, to put it differently, how the social organisation of learning is played out'. The sociocultural perspective seeks to theorise the social *interdependency* of learning through relational concepts that can explore this interdependency. One of these key concepts is *scaffolding* (Rasmussen & Ludvigsen, 2010). However, the concept originated from the cognitive sciences, and the sociocultural expansion of the scaffolding metaphor has challenged the theoretical notion of scaffolding (Stone, 1998).

Scaffolding

The word *scaffolding* was coined by Wood, Bruner and Ross (1976) and originally referred to a tutor-assisted '"scaffolding" process that enables a child or novice to solve a problem, carry out a task or achieve a goal which would be beyond his unassisted efforts' (p. 90). The concept was inspired by Vygotsky's development theory, and in particular, his concept of the *zone of proximal development* (ZPD) (Holton & Clarke, 2006). Current discussion of the scaffolding metaphor often link scaffolding to ZPD (Stone, 1998). ZPD refers to

> the distance between the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers (Vygotsky, 1978, p. 86).

Both concepts focused on the interaction between a novice/learner and an expert providing assistance so that the learner can perform at a higher level, and on the relationship between instruction and psychological development. According to Sharma and Hannafin, scaffolding provides further operationalisation of this relationship: 'The ZPD thus supplies a conceptual framework for selecting individual learning tasks, while scaffolding provides a strategic framework for selecting and implementing strategies to support specific learning' (2007, p. 28).

The research on scaffolding has mostly focused on cognitive and metacognitive scaffolding (Lajoie, 2005). A basic distinction between the two is that 'while cognition can be considered as the way learners' minds act on the "real world", metacognition is the way that their minds act on their cognition' (Holton & Clarke, 2006). There are two important aspects to keep in mind regarding metacognition. The first is that metacognition is key to *self-scaffolding*, i.e. to empower the learner to take control over their own learning processes, also referred to as 'self-regulated learning' (SRL) (Azevedo, Cromley, Winters,

Moos, & Greene, 2005). The second is that metacognition plays a prevalent role when cognition becomes problematic (Holton & Clarke, 2006). Hence in addition to structuring, scaffolding also plays an important role in problematizing in order to trigger metacognitive reflection (Reiser, 2004). Although cognitive and metacognitive scaffolding continue to be the primary focus of research, there has also been a trend towards giving the conative, affective and motivational scaffolds more attention (Lajoie, 2005).

While sharing the focus on expert support, scaffolding is distinguishable from other forms of instructional assistance due to the principle of fading, i.e. the gradual reduction and eventual elimination of scaffolds (Sharma & Hannafin, 2007). According to Van de Pol, Volman and Beishuizen (2010), no consensus regarding the definition of scaffolding exists, but some common characteristics in the various definitions can be found: 1) Contingency: the principle that the support needs to adapt to the learner's current level of performance; this also implies diagnostic strategies to identify students' level of learning. 2) Fading: the gradual withdrawal of the scaffolding. 3) Transfer of responsibility: when the scaffolds are faded the students gradually take control of their own learning; this responsibility refers to all kinds of scaffolding, e.g. cognitive, metacognitive, affective or conative support. Key questions in the research on scaffolding include what to scaffold, when to scaffold, how to scaffold, when to fade scaffolding, who or what should do the scaffolding, how to determine the effectiveness of scaffolds and identifying the mechanisms in scaffolding (Lajoie, 2005).

Some of the important criticisms of the scaffolding metaphor include that it underestimated the complexity of the learner's knowledge construction process, including how active the learner is in the learning process; the complexity of the participation structures and learning trajectories and the situated character of this learning (Stone, 1998, p. 354). Seeing the scaffolding as embedded in the broader social interaction, and in the activities and artefacts, concurs with the tenets of ZPD and sociocultural theory, but has created a conundrum for the theoretical notion of scaffolding (Palincsar, 1998; Puntambekar & Hubscher, 2005; Stone, 1998). From its initial narrower focus on one-on-one interaction between a novice and an expert, the notion of scaffolding has expanded to include a higher variety of participation structures and interactional patterns, as well as scaffolding through other means (e.g. artefacts, learning resources, curriculum design) within the broader learning environment (Puntambekar & Hubscher, 2005; Sherin, Reiser, & Edelson, 2004).

Case description

The subjects in this study were ten female and six male students in their early to mid-20s participating in the master's course in technology-enhanced learning. Fourteen of the students came from the Department of Informatics and two came from the Department of Education.

The experiment took as its starting point a natural occurring activity in many types of tertiary education—students applying some sort of reading tip/guidance from teachers regarding how to approach and analyse research literature, e.g. locating research questions, contributions, theoretical perspective, methodological concerns etc. (cf. Appendix 1). The researcher designed an annotation/categorisation experiment around this activity, and integrated the reading guidelines as scaffolds in an annotation/highlighting script in a wiki.

One of the assigned and graded tasks in the course was that students in groups of two should do a close reading of a course article and present this article for the rest of the class. As an additional task—which was designed to observe the

collaborative use of the designed scaffolding tool—the same student groups was asked to jointly categorise/annotate the article they were going to present by using a highlighting script introduced by the researcher. The students were also able to suggest changes to this highlighting script. The experiment entailed that students in groups of two used a set of categories to analyse and categorise the content in a course article. The highlighting technology resembles the highlighting pens or the colour markers in Microsoft Word, but with the extra feature that each colour was linked to a specific category (cf. Figure 2). Each of the categories was also connected to a question or a reflection prompt (cf. Appendixes 1, 2 and 3).
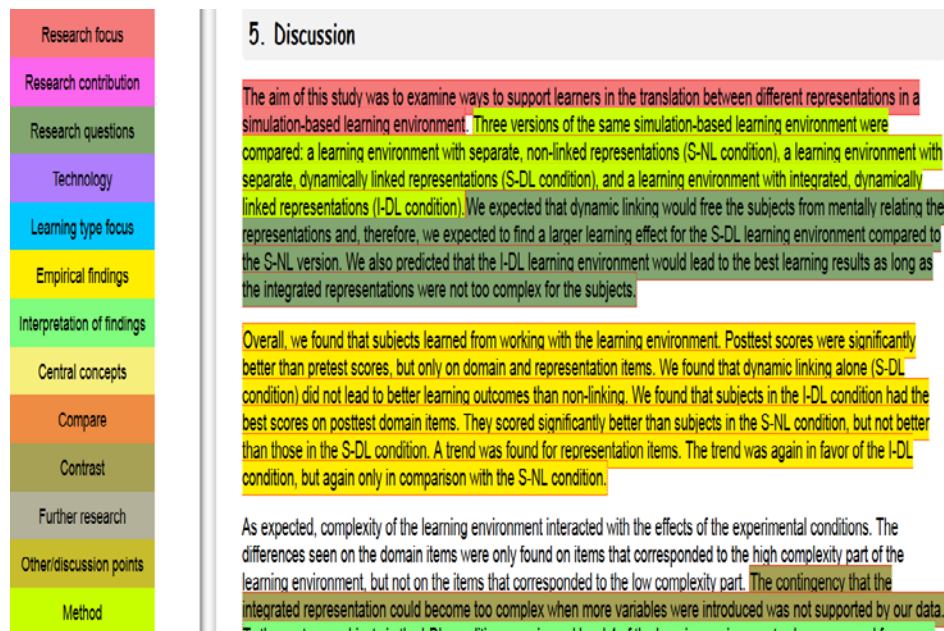

Figure 2. The highlighter/annotation script is to the left. A part of a curriculum article that students have categorised with the help of the highlighters is to the right.

The design experiment was associated with several contexts: 1) The annotation experiment; 2) the students' class presentation (the script was tailored for the deeper reading this presentation task required); 3) a semester project where students critically investigated the tool and argued for improvements and 4) the technology development this research aimed for. Potential impacts of the different contextual settings on students' inquiries were explored in Dahl (2016). This paper conducts a more detailed inquiry into the first context, i.e. the annotation experiment.

Methodology

Analytical procedure

The students' interaction with the categories was observed and video filmed with a stationary camera and a screencast. The whole interaction was transcribed and the qualitative data analyses software NVivo was used for a more comprehensive analysis. Interaction analysis will be employed to analyse the students' inquiries (Jordan & Henderson, 1995). This is a methodological framework for studying sociocultural embedded moment-to-moment interaction, i.e. microanalysis of meaning production. The unit of analysis is the embedded interaction between humans, including interaction with the socio-material world, e.g. artefacts mediating the interaction.

Design experiment

The study is based on the principles of design-based research (DBR). DBR uses theory-driven designs to generate complex interventions that are further improved through empirical studies, which also provide a more basic understanding and refinement of the theory (Design-Based Research Collective, 2003). An important area for DBR is to explore possibilities for novel learning and teaching environments (Design-Based Research Collective, 2003), and novel treatments of instructional resources, e.g. new topics, new technologies or novel forms of interaction (Confrey, 2006).

DBR's novel and experimental approach can pose a problem for the assessment of these experiments, particularly when the issue of what to expect is ignored. Kelly has criticised DBR for lacking an argumentative grammar, meaning '"the logic that guides the use of a method and that support reasoning about the data"' (Kelly quoted in; Sandoval, 2014, p. 19). To address this criticism, Sandoval proposes making a conjecture map to explicate the logic of the design, and exemplifies this in a generalised model (cf. Figure 3). The model is used to articulate the 'what', 'how' and 'why' that guided this experiment.
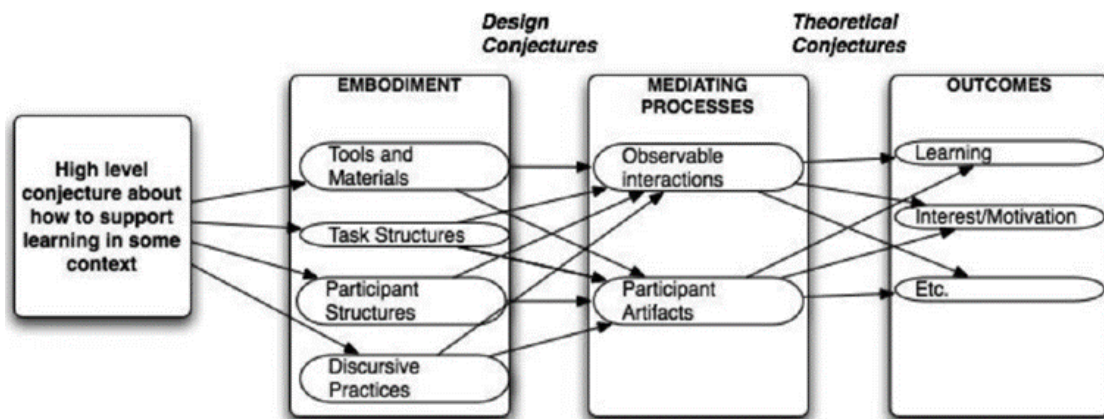

Figure 3. Generalised conjecture map for educational design research (Sandoval, 2014, p. 21).

According to Sandoval (2014), the *high-level conjecture* articulates the theoretical principle of how to support some desired form of learning, but it is on a very general level and needs to be operationalised in order to inform design and interpretation of findings. *Embodiment* refers to how the high-level conjecture has been reified in the learning design. *Tools and materials* refer to the designed tangible learning resources. *Task structure* characterises the task the learners are asked to do. *Participant structure* refers to participants' role and responsibility, and how they are expected to participate. *Discursive practices* refer to the intended ways of talking. *Design conjectures* refer to how the embodiment of the design is presumed to affect *mediating processes*, e.g. participants' interaction (*observable interactions*), or artefacts participants construct through the activity (*participant artefacts*). *Theoretical conjectures* postulate what desirable learning *outcomes* the mediating processes will help produce.

The study's *high-level conjecture* was that using annotation techniques to support students' reading and discussion of research literature could enhance students' learning in the following two respects. *First theoretical conjecture*: A greater focus on research processes and problems (*mediating process*) will

help develop knowledge-building strategies (e.g. to analyse, criticise and create knowledge) relevant for the twenty-first century (Healey, 2005; Scardamalia & Bereiter, 2006). *Second theoretical conjecture*: Personal experience with the epistemic scaffolds embedded in the annotation tool can help students explore related themes in the course, such as scaffolding, ZPD and appropriation. This can enhance students' understanding of the cognitive, metacognitive and affective functions of scaffolds, and pave the way for self-scaffolding/self-regulated learning.

Tools and materials: The annotation tool and the research article were the key resources. The tool had constraints that could either impede the scaffolding effect or increase it by making the conditions for the success and failure of scaffolds more visible. One example is some students' dissatisfaction with having to use a predefined script, which made them reflect on the personal nature of annotation strategies and the potential challenge this type of scaffolding constitutes (Dahl, 2016). Another constraint observable in this set of data is that each segment of text could only be attached to one category. This may create a false conflict between categories when they are, in fact, complementary and could both be used. How the students reason in the face of this dilemma shows whether these constraints trigger or impede their learning.

Task structure: The task was quite open. The students were expected to use appropriate categories to analyse the article. The other parts of the task was left more to students' discretion, e.g. what to tag and how to tag. However, they were encouraged to approach the highlighting as a way of engaging with interesting aspects of the article rather than trying to categorise as much as possible.

P*articipant structure*: The researcher introduced the experiment and asked the students to familiarise themselves with the different categories and propose any changes to the script. The students were asked to collaborate in the tagging, but apart from that, they decided themselves how they wanted to organise the task. The expectation was that the researcher mainly would take the role of an observer. However, successful use of annotations regularly depends on scaffolding from the teacher, so the researcher also had to intervene when necessary to facilitate more productive use of annotations.

Discursive practices: The goal was that students' talk would oscillate between the annotation categories and instances/exemplars in a way that would lead to elaborations/exploration of both the categories and the exemplars.

First design conjecture is concerned with cognitive scaffolding: The hope was that the embodied design could lead students to problematize the text in general and to explore the research process and problems covered by the article in particular (*observable interaction*), and that the discourse indicated improved conceptual understanding according to this focus (*participant artefacts*). *Second design conjecture* is concerned with metacognitive scaffolding: The hope was that the epistemic annotation categories themselves would come into focus, not only as categories that they need to know in order to use them, but also as epistemic scaffolds (*observable interaction*). Moreover, that indication of improved conceptual understanding of scaffolding could be traced in the students' discourse (*participant artefacts*).

Analysis and results

The evolvement of the tool (RQ 1)

Three scripts were explored during the design experiment (cf. Appendix 4). Before doing an empirical analysis of two students that used the third script, a few potential problems with the two prior scripts will be addressed in order to

provide some context for the script (cf. first research question). Two issues in particular with the first script led to the creation of the second script. The first issue was that several of the categories simply became too general to trigger discussions and elaboration of category decisions. During an interview after the experiment, one student note that, 'I felt at least in cases where one sits and marks absolutely everything, then there is no point anymore, it is more a case of being able to categorise this information a little. The annotation experiment came in conflict with the students' own annotation practice, and in particular the use of annotation for the purpose of indexing important issues for easy retrieval. Regular highlighting is often associated with making some key parts of the text *stand out*. When asking the students to categorise with a script that potentially can cover a major part of the text, one drawback is that it changes the ratio between highlighted and non-highlighted text, with the consequence that the highlighted text stands less out. A second concern with the first script was that it lacked a critical focus on the why and how aspects of the research, and thus resulted in an emphasis on research content rather than also trying to engage the students in the research process and the problems that the researchers tried to solve.

A goal with the second and third script was to develop a greater focus on the research process and problems as emphasised in the right side of Figure 1. Some of the means used to try to achieve this goal included placing greater emphasis on research contributions, on how the study positioned itself in relation to previous research and on the interpretational process.

The annotation practice evolving from Script 2 indicated several problems, which led to the changes in Script 3 (cf. Appendixes 2 and 3). The space only permits a few examples. These examples indicate the considerations behind the third script.

- *General focus* was removed. There was too much discussion about whether something was a general focus or the research focus. On the positive side, this discussion drew attention to the research contribution of the study.
- *Research focus* and *Research contribution* were separated into two categories. Since the students seemed to be concerned with finding instances of all the highlighter tags, this could give the research contribution more attention. When two categories shared a highlighter tag, the students tended to focus on the easiest category to decide on, which in this case was research focus.
- *Learning type focus* was added because there was little reflection regarding what sort of learning the technology-enhanced learning environment targeted.
- *Theoretical perspective* was used a lot in the tagging process, but it was not elaborated. For example, what does the theory explain? How does the theory explain the phenomenon? Which inferences can be drawn from the theory? Is it validated to the extent that we will call it a theory or is it more a framework for the systematisation of empirical findings? Theoretical perspective was replaced by *Central concepts* to force students to elaborate on whether the concept was part of a theoretical framework as well as what the concept denotes.
- *Compare and Contrast* was separated into two categories to attract more attention to the study's positioning with regard to previous research, either drawing an analogy (compare) or criticising and adding something missing in previous research (contrast). The hope was that this could lead to more exploration of research problems.
- The category descriptions were on a separate wiki page. To ease the access to these prompts, they were integrated as popups appearing when the marker was placed above a category.

Empirical analysis of students' use of categories

The selected data focuses on students' discussions of whether to tag with *Empirical findings* or *Interpretation of findings* (in the students' discourse this was often shortened to *Findings* and *Interpretation*). The category descriptions/prompts (cf. Appendix 3) are central resources in the students' discussions, particularly with respect to the following two categories:

- *Empirical findings*: What are the empirical findings of the study? These are observation data—NOT to be confused with the interpretations drawn from the data, which can be erroneous. The empirical findings can also stem from reviews/other studies.
- *Interpretation of findings*: Do the interpretations emerge naturally from the data, or can they be questioned?

A generic script was used, and the researcher therefore wanted the *Empirical findings* category to be broad. Besides primary data, the desire was to include studies based on secondary data, so the script could also fit the more review-based articles in the course. In the following case, the students are working with a review-based article.

The case will be illustrated through five excerpts. Excerpt 1 shows the students' interaction with the category descriptions to establish an understanding of the categories and their relevance prior to annotating. Excerpt 2 illustrates how verbal cues pervaded the students' category decisions. Excerpt 3 shows how the interpretation work escapes students' attention. Excerpt 4 gives an example of scaffolding by the teacher. Excerpt 5 analyses students' engagement with the categories after the scaffolding intervention.

Prior to this excerpt, the researcher suggests that the students go through the categories and category descriptions in the highlighter script to find out what categories they missed:

Excerpt 1:

33	Student B:	((*Puts the mouse above 'Empirical findings' so the popup prompt appears and reads:*)) '*Empirical findings*' (.) he has a lot of those.
34	Student A:	Umm.
35	Student B:	((*Switches to the description of 'Interpretation of findings' and reads:*)) *Interpretation=*
36	Student A:	=Wait a little ((*Returns to description of 'Empirical findings' and reads quietly:*)) '[*…NOT to be confused with the*] Interpretations drawn from the data, which can be [*erroneous*]' (.) okay (.) well, yeah, alright.
37	Student B:	((*Reads from the latter part of the description:*)) 'can also stem from other studies' (.) Well that's everything.
38	Student A:	Um.
39	Student B:	((*Switches to 'Interpretation of findings' and reads selected words:*)) '[*Do the interpretations*] emerge naturally [*from the*] data, [*or*] can [*they*] be questioned?'
40	Student A:	Um.
41	Student B:	Ummm.
42	Student A:	I'm thinking about the table with the development of research.
43	Student B:	Yes (.) that's quite=
44	Student A:	=very natural from the data.
45	Student B:	It is a very quantitative study, so there is (.) perhaps (.) not much to question.
46	Student A:	Umm.

Transcript notations:

The transcript notations are based on the Jefferson system, but follow a simplified and more accessible version:

[ ]	Text in square brackets represents clarifying information regarding the discourse.
(( ))	A description inserted between double brackets denotes contextual information.
=	Indicates that there is no discernible pause between two speakers' turns.
?	Rising intonation.
:	Indicates prolongation of a sound.
Underlined:	Emphasis in speech.
(.)	Short pause in the speech.
[…]	Utterances removed from the original dialog.
-	Single dash in the middle of a word denotes that the speaker interrupts herself.
--	Double dash at the end of an utterance indicates that the speaker's utterance is incomplete.

A prevalent focus was whether the students could find instances of the categories (cf. line 33). They further emphasise different aspects of the category *Empirical findings*, which may provide different directions for their further explorations. Student A (line 36) highlights the contrast between *Empirical findings* and *Interpretation of findings*, while Student B (line 37) addresses the category property that broadens it to include the secondary data. The comment that follows, 'Well that's everything', indicates that he perceives the category to be so general that it leaves nothing out. The students' discussion of *Interpretation of findings* (line 39–46) relates the category to some tables they later discuss in Excerpts 3 and 5. Here, they regard the tables as interpretations, but interpretations that emerge naturally from the data, and that it is a quantitative study, so it might be little to question about the study.

The next excerpt illustrates the central role of verbal cues.

Excerpt 2:

341   Student B: And then it is some *Interpretation.*
342   Student A: Yeah.
343   Student B: ((*Reads:*)) 'Might be'=
344   Student A: ((*Reads the next words in the sentence:*))
                   ='due to the…trajectory', well, yes.
345   Student B: ((*Tags part of the sentence*)) I don't know if we should take more.
346   Student A: It's perhaps natural to include the whole sentence (.) I don't know.
347   Student B: ((*Tags the whole sentence*)) Let's do it.
348   Student A: Yes.
349   Student B: It's perhaps the next as well; ((*reads:*)) 'might be', then it's 'interpretation'
                   ((*tags*)).

Excerpt 2, shows how particular words ('might be') become associated with a category (*Interpretation of findings*). The words become a *verbal cue* for category identification. The verbal cues are further used as heuristics for making swift category decisions.
A second pattern that constituted a lot of the students' talk concerns the question of how much of the text to tag. On one side, students were concerned with making sure that they managed to include all relevant instances. On the other side, the annotation was associated with highlighting a smaller part of the text so that this part could *stand out.*

In the next excerpt, the students address the table discussed in Excerpt 1. Their category decisions involve few elaborations and are more associative.

Excerpt 3:

350 Student B: ((*Scrolls to a table displaying the yearly percentage of empirical studies*))
Shall we tag all of this as *Findings*? ((*both laugh*)).
351 Student A: But what exactly is (.) yes, no.
352 Student B: Hm?
353 Student A: ((*Reads table heading:*)) 'Percentage of empirical studies'.
354 Student B: This must be the *Findings.*
355 Student A: But is it the <u>*Empirical*</u> *findings*?
356 Student B: Isn't it?
357 Student A: I don't know (.) Yeah?
358 Student B: ((*Accesses the 'Empirical Finding' prompt so they can read it*)).
359 Student A: Yes.
360 Student B: It has to be.
361 Student A: Yes (.) okay.
362 Student B: Although he didn't observe anything, but he retrieves the data from (.)
363 Student A: Umm.
364 Student B: Um, um, um.
365 Student A: Yes there's nothing (.) he is not an interpreter, he is just=
366 Student B: ((*Interrupts and reads selected words:*))='[*This finding was, by and large,*]
<u>consistent</u> with the claim that the exploration [*of SL in education*] started
[*from*] 2005 [*(Foster, 2007; Luo and Kemp, 2008).*]' (.) this is a *Finding* (.)
or?
367 Student A: Yes, he studies the data here (.) all, all is=
368 Student B: =yes all is=
369 Student A: =*Empirical Findings*=
370 Student B: =indeed=
371 Student A: =and then he interprets it as ((*reads selected words:*)) '[*consistent with the*]
claim [*that the exploration of SL in education*] started [*from 2005*]'.
However, this is also a *Comparison.*
372 Student B: Yes, it is that as well.
373 Student A: He mentions another one who said that it started in 2005=
374 Student B: =Umm=
375 Student A: =two others (.) three.
376 Student B: It is both *Interpretation* and *Compare.*
377 Student A: Yes.
378 Student B: But it isn't possible to tag two things ((*the tool functionality prevented the
students from using overlapping tags, i.e. using several categories on the
same text*)).
379 Student A: No.
380 Student B: But if we say that this ((*points*)) is a little more *Interpretation*, and the first
one is a little more *Compare*, perhaps.
381 Student A: Umm.
382 Student B: ((*tags second sentence as 'Interpretation'*)) Do you agree?
383 Student A: Umm.
384 Student B: ((*tags first sentence as 'Compare'*)).

Contrary to their discussion of the tables in Excerpt 1, Student B now asks if they should tag the table as *Findings* (line 350). Student A questions whether these findings are '*Empirical*' (line 355), but does not elaborate, and finally agrees that it is. The article uses the concept 'findings' sixteen times, but in a broader meaning that refers to both the empirical findings and their interpretations. This conflict between the script and how the article used the concept did not seem to be noticed, but may still have had an impact on their understanding.

In line 362, Student B draws on the category description, and implies that although the findings do not meet the first criterion (primary data), it meets the second criterion (secondary data). Student A agrees (line 365) and argues that it must be *Empirical findings* because there are no interpretations, and hence uses a logic that reaches a category decision by excluding the alternative categories.

In the interaction following line 366, the meaning of the table is explored through two sentences next to the table, and here the phrase 'consistent with...' becomes a verbal cue. Student B associates 'consistent with [other studies]' as a cue for *Empirical findings,* which follows his understanding in Excerpt 1, line 37 that when secondary data is included in *Empirical findings,* the category covers 'everything'. Student A concurs, but also refers to the sentences as interpretations that can fall under the category *Compare* (line 371). Student B's response (line 376) is that the text is both *Interpretation* and *Compare.* Since the tool did not support affiliation to more than one category, Student B suggests dividing the text between the categories (line 378–384).

Shortly after this excerpt, the researcher/teacher intervenes to interrogate about the students' understanding of the two categories and to provide some scaffolding if needed.

Excerpt 4:

403	Researcher:	So how did you understand the difference between *Empirical understandings* and *Interpretation of findings*?
404	Student B:	*Findings* are just very straightforward; here is the data I collected, without any further description. While *Interpretation* is=
405	Student A:	=is what they extract from the data (.) what they understand from the data. […]
411	Researcher:	Is it possible to problematize whether some of the empirical findings you have tagged are so straightforward, taken directly from the data, or if there is a lot of coding involved? […]
417	Researcher:	Is it so that a study calls itself empirical and another conceptual ((*reads:*)) 'pure conceptual discussions' ((*the students have tagged these categories as 'Empirical findings', hence assuming that there is no interpretation involved*)).

The researcher concluded from the students' answer that they understand the distinction between the categories. However, he doubted that their practical skills of using the categories were on the same level, and he asks a couple of follow-up questions to scaffold their further tagging.

The final excerpt looks at how this intervention may have influenced the tagging:

Excerpt 5:

471	Student B:	((*Scrolls to a section with a table that summarises the focus in the reviewed studies*)) This one is very nice=
472	Student A:	=Umm=
473	Student B:	=*Findings*=
474	Student A:	=Umm=
475	Student B:	=In a way.
476	Student A:	The question is how he made these.
477	Student B:	Yes, it is probably some *Interpretation*=
478	Student A:	=Here comes perhaps some *Central* (.) is there some *Central Concept* here?=where he describes the categories he has chosen (.) didn't he?
479	Student B:	Umm. He has analysed 'research purposes' (.) coded them (.) that can also fall under *Method--*
480	Student A:	Umm.
481	Student B:	how he has done it.
482	Student A:	Umm.
483	Student B:	((*tags with 'Method'*)).
484	Student B:	Don't know if we call this *Interpretation* or [*Empirical*] *Finding.*
485	Student A:	Yes=it is both.
486	Student B:	=It is both ((*said simultaneously*)) in a way=
487	Student A:	=Because=
488	Student B:	=the categories [*in the table*] are an interpretation.
489	Student A:	Umm.
490	Student B:	So it is maybe *Interpretation* then (.) but it's somewhat important.
491	Student A:	Umm.
		[…]
499	Student A:	But he presents it as a *Finding*= but it is an *Interpretation.*
500	Student B:	=but it is an *Interpretation* ((*said simultaneously*)), yes (.) it is? So it might be an *Interpretation*=
501	Student A:	=Umm=
502	Student B:	=all of this=
503	Student A:	=Umm.
504	Student A:	It is a little funny (.) here he presents everything as a finding (.) and here he presents how he interprets it.
505	Student B:	Umm.

Student B associates the table with *Empirical findings* (line 471–475). Student A claims the question is how the categories in the table are made (line 476). Student B acknowledges that some *Interpretation* can be involved. However, Student A's comment 'Here comes perhaps some *Central* [*Concepts*]' changes the focus and represents a pattern that the students tried to find instances of in all the categories. Aligning *Interpretation* with *Concepts* could help disclose the interpretation work and research contributions. However, Student B instead suggests that the *Method* category could be applied. In line 484–491, the discussion returns to whether the table should be regarded as *Interpretation* or *Empirical findings*; they conclude that it is both. This time it is Student B making the point that the categories in the table are *Interpretations*. In the final episode (line 499–505), Student A emphasises the paradox that this text is presented as Empirical findings when it is clear that interpretations are involved.

The further discussion focuses on the design conjectures, and will first address the design's mediating role with regard to cognitive scaffolding (RQ 2a) before exploring the mediation of metacognitive scaffolding (RQ 2b).

Mediation of cognitive scaffolding (RQ 2a)

The use of verbal cues to make swift category decisions (cf. Excerpt 2) was a pervasive pattern across the study. When the verbal cues were first established, the categorisation process could become almost mechanical, and thus often displaced the more conceptually oriented discussions. However, also when this strategy was not used, the interpretational process often escaped students' attention, cf. Excerpt 3: Although the students finally tagged

a part of the text with *Interpretation of findings*, this change of category was sudden and did not lead to elaborations or reflections on the table that was the focus of the discussion. The students' category decisions in Excerpt 5 also involved many sudden turns and few elaborations, but it also resulted in a major conceptual turn: that of seeing the presented findings as relying heavily on interpretations. The researcher's scaffolding in Excerpt 4 probably contributed to this turn.

Many factors in the embodied design and the evolving annotation practice affected the observed interaction. Only a few examples can be addressed. The discussions in Excerpts 3 and 5 seem to be seeded in two different routes for exploring *Empirical findings* that Student A and Student B chose in Excerpt 1. Student A focuses on contrasting categories, and making category decisions based on *exclusion* (excluding the alternative categories). Student B instead addressed the different types of data (primary and secondary) that *Empirical findings* applies to and hence focused on this category's *inclusion* criteria.

Although the third script improved some of the critical issues of the second script, several problems remained. For example, it was not possible to organise the categories in a folder structure, and the high number of categories complicated the script's use. Another constraint influencing the tool use is illustrated in Line 378 where Student B commented 'But it isn't possible to tag two things'. They 'solved' this by dividing the text in two, and categorising the first part with *Compare* and the second part with *Interpretation of findings*. It would have been more in line with the conjectures if the students instead would reflect on how the text can have multiple category affiliations and explore this relationship. The students seem to presume that the relationship between the categories necessarily involves competition and/or contradictions (as emphasised in the prompts of *Empirical findings* and *Interpretation of findings*). However, several of the categories were complementary; *Compare* and *Interpretation of findings* is just one example. In addition, the students' focus seems to be on finding instances of all the categories (Excerpt 1, line 33; Excerpt 5, line 478), rather than seeing the categories as alternative ways to problematize the text and explore research processes and problems. Hence, the assumption regarding the *task structure* that the students would approach the highlighting more as a way of engaging with interesting aspects of the article rather than thinking that they should categorise as much as possible proved incorrect. This was perhaps due to the fact that such a detailed script can widen the focus and counteract elaborations.

In conclusion, the above discussion provides several indications of cognitive scaffolding (Holton & Clarke, 2006), but also many examples that this scaffolding did not fully reach the conjectures that the design aimed for.

Hence, the *first design conjecture*—presuming that the embodied design could lead students to problematize the text and to explore the article's research processes and problems (*observable interaction*)—is only partly supported.

Mediation of metacognitive scaffolding (RQ 2b)

The presumed *discursive practice* was that 'students' talk would oscillate between the annotation categories and instances/exemplars'. This happened in the sense that students frequently used the category prompts in their category decisions, as well as compared different category descriptions in cases of doubt. The following part of the assumption (that this 'would lead to elaboration/exploration of both the categories and the exemplars') was partly true. The development of students' understanding of the knowledge construction, in which they gradually became more attentive to the interpretational work involved in the article's conceptualisation and coding procedures, illustrates the exploration of the exemplars.

However, their exploration of the categories was more modest. There are several examples that the annotation categories affected students' thinking and reflections on the article. However, this does not reflect back on students' discussion of the categories. The category descriptions were treated merely as definitions. This approach risks turning the categories into static entities, which can be in conflict with the more dynamic, instrumental and interpretational nature of scaffolding, particularly the metacognitive scaffolds.

The lack of metacognitive reflections was perhaps not that surprising as this was their first experience with the design. The annotation experiment became a rather isolated task, which thus made it more difficult for the students to find a relevant framing of the task. The tool had several constraints, e.g. predefined scripts that deprived students of epistemic agency (Dahl, 2016). Yet the conjecture was that the tool's constraints not only weakened the tool support, but also could be an asset, insofar as the students used their experiences with the tool to problematize key concepts and principles in the course literature. However, this required meta-reflections, where the students not only explored the functionality of the tool (used a *tool perspective*), but also reflected on the fundamental design principles and how these related to their learning (used a *tool inquiry perspective*) (Dahl, 2016). In spite of the researcher's attempt to frame the technology within a tool inquiry perspective, Dahl (2016) showed that it was difficult for the students to acquire meta-reflections based on the annotation experiment in isolation. There was a need for another context where the tool inquiry perspective could get more attention, and the students' semester project provided this context. Here, aspects of cognition, metacognition and scaffolding became key concerns (Dahl, 2016).

In conclusion, the metacognitive scaffolding was not sufficiently triggered, which impacts the opportunities of self-scaffolding and SRL (Azevedo et al., 2005; Holton & Clarke, 2006). We see no examples that the categories are approached or problematized as tools for thinking about learning; they are treated merely as means for categorising. The students seem to have little awareness of the presence of scaffolding and the scaffolding principles embedded in the design.

Hence, the *second design conjecture*—presuming that the epistemic annotation categories themselves would come into focus, not only as categories the students need to know in order to use them, but also as epistemic scaffolds (*observable interaction*)—is not supported.

Concluding remarks

The principal issues the paper has attempted to emphasise are as follows:
- When annotations are studied from the perspectives of mediation (Vygotsky, 1978; Wertsch, 1991) and scaffolding (Vygotsky, 1978; Wood et al., 1976), there is no simple cause–effect relationship between the explored social annotation technology and the outcome of using it. The tool is not an isolated variable, but instead is constructed through its use, often in unpredictable ways. We glimpsed fragments of this pattern, but not (yet) to the extent that we can claim that we know and control the relevant variables and can assess the tool's general effect (Novak et al., 2012).
- There were several deviations from the design's conjectures. An articulation of the conjectures (Sandoval, 2014) can help accomplish the following:
    - reveal the deviations and their implications;
    - provide guidance for further design and theory improvements and

- expose the expectations of the tool so that these can become part of a researcher/teacher–student co-learning.
- The deviations are themselves interesting areas for further investigations.
- Involving the students in this investigation can provide unique learning opportunities. Hence, social annotation resources ought to be perceived as co-constructed rather than passed on resources (Dahl, 2016).
- The co-construction should include not only the use and understanding of categories, but also the epistemic scaffolding embedded in these categories. Whereas there are many examples of the first type of co-construction, the students became less involved in the construction of the scripts. Including the students in the creation and reflection on these scaffolds can enhance their metacognitive thinking (Holton & Clarke, 2006) and provide valuable resources for the further development and tuning of these scaffolds.
- Two areas for further research are as follows:
  - settings where students themselves develop their own annotation scripts to a greater extent and
  - settings where teachers and students collaboratively explore the potentials and challenges of these scaffolds.

References

Arnseth, H. C., & Säljö, R. (2007). Making sense of epistemic categories. Analysing students' use of categories of progressive inquiry in computer mediated collaborative activities. *Journal of Computer Assisted Learning, 23*(5), 425-439.

Azevedo, R., Cromley, J. G., Winters, F. I., Moos, D. C., & Greene, J. A. (2005). Adaptive human scaffolding facilitates adolescents' self-regulated learning with hypermedia. *Instructional science, 33*(5-6), 381-412.

Ben-Yehudah, G., & Eshet-Alkalai, Y. (2014). *The influence of text annotation tools on print and digital reading comprehension.* Paper presented at the Proceedings of the 9th Chais Conference for Innovation in Learning Technologies.

Chen, C. M., & Huang, S. H. (2014). Web-based reading annotation system with an attention-based self-regulated learning mechanism for promoting reading performance. *British Journal of Educational Technology, 45*(5), 959-980.

Confrey, J. (2006). *The evolution of design studies as methodology*: na.

Dahl, J. E. (2016). Students' Framing of a Reading Annotation Tool in the Context of Research-Based Teaching. *TOJET, 15*(1).

Design-Based Research Collective. (2003). Design-based research: An emerging paradigm for educational inquiry. *Educational researcher*, 5-8.

Healey, M. (2005). Linking research and teaching exploring disciplinary spaces and the role of inquiry-based learning. *Reshaping the university: new relationships between research, scholarship and teaching*, 67-78.

Holton, D., & Clarke, D. (2006). Scaffolding and metacognition. *International journal of mathematical education in science and technology, 37*(2), 127-143.

Johnson, T. E., Archibald, T. N., & Tenenbaum, G. (2010). Individual and team annotation effects on students' reading comprehension, critical thinking, and meta-cognitive skills. *Computers in human behavior, 26*(6), 1496-1507.

Jordan, B., & Henderson, A. (1995). Interaction analysis: Foundations and practice. *The journal of the learning sciences, 4*(1), 39-103.

Lajoie, S. P. (2005). Extending the scaffolding metaphor. *Instructional science, 33*(5-6), 541-557.

Ludvigsen, S. (2012). What counts as knowledge: Learning to use categories in computer environments. *Learning, Media and Technology, 37*(1), 40-52.

Ludvigsen, S., & Mørch, A. (2003). Categorisation in knowledge building *Designing for change in networked learning environments* (pp. 67-76): Springer.

Marshall, C. C. (1997). *Annotation: from paper books to the digital library.* Paper presented at the ACM Digital Libraries '97 Conference, Philadelphia, PA (July 23-26, 1997), pp. 131-140.

Marshall, C. C. (1998). *Toward an ecology of hypertext annotation.* Paper presented at the Proceedings of the ninth ACM conference on Hypertext and hypermedia:

links, objects, time and space---structure in hypermedia systems: links, objects, time and space---structure in hypermedia systems.

Muukkonen, H., Hakkarainen, K., & Lakkala, M. (1999). *Collaborative technology for facilitating progressive inquiry: Future learning environment tools.* Paper presented at the Proceedings of the 1999 conference on Computer support for collaborative learning.

Novak, E., Razzouk, R., & Johnson, T. E. (2012). The educational use of social annotation tools in higher education: A literature review. *The Internet and Higher Education, 15*(1), 39-49.

Ovsiannikov, I. A., Arbib, M. A., & McNeill, T. H. (1999). Annotation technology. *International journal of human-computer studies, 50*(4), 329-362.

Palincsar, A. S. (1998). Keeping the Metaphor of Scaffolding Fresh—-A Response to C. Addison Stone's "The Metaphor of Scaffolding Its Utility for the Field of Learning Disabilities". *Journal of Learning Disabilities, 31*(4), 370-373.

Puntambekar, S., & Hubscher, R. (2005). Tools for scaffolding students in a complex learning environment: What have we gained and what have we missed? *Educational psychologist, 40*(1), 1-12.

Rasmussen, I., & Ludvigsen, S. (2010). Learning with computer tools and environments: A sociocultural perspective. In K.Littleton, C.Wood & J.K.Staarman (eds) *International handbook of psychology in education*, (pp. 399-435. Bingley, UK: Emerald Group Publishing Limited.

Razon, S., Turner, J., Johnson, T. E., Arsal, G., & Tenenbaum, G. (2012). Effects of a collaborative annotation method on students' learning and learning-related motivation and affect. *Computers in human behavior, 28*(2), 350-359.

Reiser, B. J. (2004). Scaffolding complex learning: The mechanisms of structuring and problematizing student work. *The journal of the learning sciences, 13*(3), 273-304.

Samuel, R. D., Kim, C., & Johnson, T. E. (2011). A study of a social annotation modeling learning system. *Journal of Educational Computing Research, 45*(1), 117-137.

Sandoval, W. (2014). Conjecture mapping: an approach to systematic educational design research. *Journal of the Learning Sciences, 23*(1), 18-36.

Scardamalia, M., & Bereiter, C. (2006). Knowledge building: Theory, pedagogy, and technology. *The Cambridge handbook of the learning sciences*, 97-115.

Schugar, J. T., Schugar, H., & Penny, C. (2011). A nook or a book? Comparing college students' reading comprehension levels, critical reading, and study skills. *International Journal of Technology in Teaching and Learning, 7*(2), 174-192.

Sharma, P., & Hannafin, M. J. (2007). Scaffolding in technology-enhanced learning environments. *Interactive Learning Environments, 15*(1), 27-46.

Sherin, B., Reiser, B. J., & Edelson, D. (2004). Scaffolding analysis: Extending the scaffolding metaphor to learning artifacts. *The journal of the learning sciences, 13*(3), 387-421.

Stone, C. A. (1998). The metaphor of scaffolding its utility for the field of learning disabilities. *Journal of Learning Disabilities, 31*(4), 344-364.

Sung, H.-Y., Hwang, G.-J., Liu, S.-Y., & Chiu, I.-h. (2014). A prompt-based annotation approach to conducting mobile learning activities for architecture design courses. *Computers & Education, 76*, 80-90.

Van de Pol, J., Volman, M., & Beishuizen, J. (2010). Scaffolding in teacher–student interaction: A decade of research. *Educational Psychology Review, 22*(3), 271-296.

Vygotsky, L. S. (1978). Mind and society: The development of higher mental processes: Cambridge, MA: Harvard University Press.

Wertsch, J. V. (1991). *Voices of the mind*: Harvard University Press.

Wolfe, J., & Neuwirth, C. M. (2001). From the Margins to the Center: The Future of Annotation. *Journal of Business and Technical Communication, 15*(3), 333-371.

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of child psychology and psychiatry, 17*(2), 89-100.

Yang, X., Yu, S., & Sun, Z. (2013). The effect of collaborative annotation on Chinese reading level in primary schools in China. *British Journal of Educational Technology, 44*(1), 95-111.

Appendix

Appendix 1

The categories and prompts/category descriptions in the first script are as follows:

How to present an article

- Focus
  - What is the theme of the article/study?
  - Research questions (explicit or implicit)
- Theoretical perspective
  - What theory or conceptual framework is utilised?
  - Can you compare it with the key concepts we use in the course (ZPD, mediation, acquisition, participation etc.)?
- Locus
  - Setting and Participants
    - E.g. Informal vs. formal learning, distributed vs. collocated
  - Subject matter
  - Technology
    - E.g. Scaffolding
- Methods
  - Qualitative/quantitative?
  - Types of data
  - Analysis methods
- Findings and conclusions
- Discussion points
  - Strengths and shortcomings of the study
  - Do you agree/disagree about the conclusions reported

Figure 4. Categories and prompts for the first script.

Appendix 2

The categories and prompts/category descriptions in the second script are the following.

How to read an article?
Identify the basic elements and use these for questioning the research process.

Categories	Category descriptions/prompts
General focus	All there is vs. general focus: Even the general focus is topical, disciplinary etc., i.e. narrowed down, e.g. ICT and Learning, Technology-Enhanced Learning, Computer-Supported Collaborative Learning (CSCL). Notice that while CSCL is a narrower focus than the first two, it is still very general.
Research focus	General focus vs. Research focus: For a topic to be researchable, it needs to be narrowed down. The move from a general to a focused topic often also provides interesting perspectives regarding research contribution and the framing and implications of the research.
Research questions	Research focus vs. Research questions: The number of research questions that can address the research focus is usually very high, and thus the question we should ask is as follows: Which of the research questions are the most promising for our research? A critique we might easily miss if we group the two together is how well the research focus and the research questions are aligned.
Setting	What is the research setting? Who and what is studied? Where and in what context is it studied? For example, primary school pupils, university students or workers; informal vs. formal learning; synchronous vs. asynchronous communication; subject matter (English, science, communication etc.).
Data collection	Research questions vs. Data collection: Many different sorts of data can be used to explore the research questions, and there are also multiple methods for gathering these data. The chosen method for data collection can be critiqued based both on the research's own assumptions (e.g. for being inadequate to address the research questions) and on more general methodological concerns (e.g. methodological conventions have not been followed).
Review & Empirical findings	Review & Empirical findings vs. Interpretation of findings. For practical reasons one separates observations from the interpretations that can be drawn from the observations. There is also an important methodological reason for this distinction: research shall attempt to be verifiable and open to the scrutiny of others, and jumping directly to the interpretations and conclusions is in conflict with this objective.
Data analyses and interpretation	Data has the potential to inform (cf. empirical findings) and information has the potential to support knowledge claims (cf. interpretation of findings). But much of the knowledge we are interested in is not accessible through a simple 'reading' of the data. The data analysis is crucial to reveal the deeper patterns of possible interpretations that can be inferred from the data. Due to the guesswork and conditional interpretation involved in data analyses, and hence the constructed nature of the knowledge resulting from the analyses, some prefer to call it 'data production' instead of 'data analysis'.
Theoretical perspective	Data analyses will often rely on theories and conceptual frameworks. Like hypotheses, theories make assumptions about the world. A distinction is that theories have undergone extensive testing, whereas hypotheses conventionally often refer to assumptions that have not yet been rigorously tested and validated. Theories help explain or predict phenomena or correlations. When using a theory, we also add to the knowledge of the value of this theory, e.g. by adding or excluding phenomena or relations the theory can account for or pointing out limitations where other theories stand a better chance of explaining the phenomenon etc.
Overall conclusion/ Broader Implications	Suggestions for further research.

Figure 5. Categories and prompts for the second script.

A few other concepts that can be useful when analysing science articles as listed below:

Categories	Category descriptions/prompts
Compare and contrast (method for analyses)	Compare: What are the similarities? Contrast: What are the differences? What units and variables have been Compared/Contrasted? Are the units and variables the main focus, or are they indicators/used to represent something else? What is the purpose for the Comparison/Contrasting?
Central concepts, principles, definitions, ideas, objects	Central concepts, principles, definitions, ideas and objects that constitute the research discipline. Beside the theoretical concepts, there might be many other concepts of relevance for understanding and contextualizing the research and domain knowledge.
Human models	Human models extract particular aspects of what constitute humans and the social world and use these aspects to make assumptions, explanations, predictions or normative judgements. The intended scientific contribution of these models can be to isolate a phenomenon and help us create empirical testable assumptions/hypotheses. Our critique of these kinds of human models follow the same path; are the hypotheses validated empirically (through observation, experimentation etc.)? If not, the human models are not supported and should be corrected.

Figure 6. Categories and prompts for the second script.

Appendix 3

The categories and prompts/category descriptions in the third script are as follows:

Categories	Category descriptions/prompts
Research focus	What is the focus of the study?
Research contribution	What do the authors consider to be their unique research contribution?
Research questions	What is the research question(s) (explicit or implicit)?
Technology	What type of technology (collaboration system, ITS, domain-specific app, game, virtual world etc.)? Is it a prototype or a finished system? Which affordances and constraints are built into the technology? What are the underlying design principles and theory?
Learning type focus	What kind of learning or learning activity is supposedly enhanced through the support of the technology? How?
Interpretation of findings	Do the interpretations emerge naturally from the data, or can they be questioned?
Central concepts	Explain the meaning of the concept. Can the concept be linked to controversies? If so, how? Is the concept related to other concepts? Which and how? Is the concept a theoretical concept, or is it based on interpretations of empirical observations not yet validated to the extent of gaining a theoretical status?
Compare	Which similarities with other studies and perspectives are picked up and extended in this study? This implies an analogy—is the analogy the authors have made questionable? Does the knowledge gained from the other quoted studies transfer indisputably to this study?
Contrast	How do the authors contrast and position their study (cf. research contribution) to other studies, perspectives etc.?
Empirical findings	What are the empirical findings of the study? These are observation data—NOT to be confused with the interpretations drawn from the data, which can be erroneous. The empirical findings can also stem from reviews/other studies.
Further research	The authors' or your own suggestions regarding the need of further research as an extension of this study.
Other/discussion points	This is an open category for any other topic you want to highlight.
Method	Note: This category was added by the students.

Figure 7. Categories and prompts for the third script.

Appendix 4

Here follows a comparison of the categories in the three scripts:

Figure 8. The three scripts.