A Critical Summary of: 'Using Situated Learning and Multimedia to Investigate Higher-Order Thinking' (Herrington and Oliver)

Colin McCarty

MA in ICT

Research and Development in ICT in Education
(ICT.02C2)

Word count: 3,974

Introduction

This paper by J. Herrington and R. Oliver is an attempt to show that a situated learning environment can be created using a multimedia program. Furthermore they hope to establish that participants in this environment will use substantial amounts of higher order thinking, as a result of the situated learning structure. They feel this is important research, because it has been assumed that higher order thinking is present in a situated learning environment, with little empirical evidence to back up this position (Herrington and Oliver, 1999, 4). How successful the authors are in this pursuit will be examined in four areas: the establishment of their theoretical position, the nature of their empirical setting, the collection and analysis of their data, and the conclusions that are reached.

The Theoretical Position

The theoretical position that this paper adopts is that multimedia programs based on situated learning environments promote substantial levels of higher order thinking. Immediately a question should be asked of this theoretical position; what do the authors mean by the terms ‘higher order thinking’ and ‘situated learning’?

The authors define situated learning, in part, around a framework of six criteria that they themselves developed: “(a) an authentic context; (b) complex authentic activities; (c) multiple perspectives; (d) expert performances; (e) coaching and scaffolding; (f) opportunities for collaboration, reflection and articulation; and (g) authentic assessment”

(Herrington and Oliver, 1999, 5). However, these criteria themselves need further definition in order to shed light on what is meant by situated learning. For example, what are 'complex authentic activities'? What is 'authentic'? How does one achieve 'authenticity' within an activity? How 'complex' do these authentic activities need to be, in order to qualify as complex? The elaboration needed on the terms used is not forthcoming from Herrington and Oliver. They do attempt to use the antecedent work to clarify situated learning, but fail to fully engage with this work. In most instances they offer a quote from a theorist in place of developing their own explicit definition of the terms. Brown and Dowling (1998) point out that the development of the theoretical field involves "making explicit a nebula of debates and theories and, indeed, empirical findings about the area of concern" (p. 20). Engagement with the theoretical field is more pervasive when the authors develop their framework for analysis. However, in the development of the problem in this study, the engagement with theory is lacking. Thus we are left with a vague definition of situated learning as "apprenticeship" and learning "embedded in the social and physical context within which it will be used" (Herrington and Oliver, 1999, 4).

The same problem occurs with the term, 'higher order thinking'. The authors imply, by quoting Newmann, that they are attempting to build a "common conception" of what higher order thinking is to attract "professional consensus"(ibid, 4). In the attempt to develop a consensus a clear definition is never given. Instead a table is developed to classify higher order thinking into six categories. This classification scheme for higher order thinking is based on the work of Resnick and Henri, with the work of additional researchers used to corroborate each unit of classification. While each unit of the

classification system is defined, higher order thinking itself is left undefined. So the common conception that they wish to develop seems to involve quoting several researchers in the theoretical field, without elaborating on how these researchers are related to the problematic in which Herrington and Oliver are working.

The substitution of a list of criteria for the clear definition of situated learning and higher order thinking leaves any conclusion that is generated from this research in doubt. As Brown and Dowling state, "the concepts involved in the problem should be developed to a degree which enables their empirical measurement or operationalization" (Brown and Dowling, 1998, 139). The reader is given no such degree of development in Herrington and Oliver's problem. Clearly without the definitions of the key concepts behind the problem, any data generated from their measurement in the empirical setting is questionable.

Empirical Field and Empirical Setting

The focus for the research is the impact of situated learning in classroom practice. While the classroom is never specifically mentioned as the empirical field, it is implied by the author's citation of researchers such as Brown, Collins and Duguid who "were among the first to use the ideas to produce a proposal for a model of instruction that has implications in classroom practice" (Herrington and Oliver, 1999, 4). The authors' choice of empirical setting indicates that educational practice is the field they wish to make statements about. Thus the classroom, or any traditional educational setting, can be construed as the empirical field for the research. The operationalization of their

theoretical position in this empirical field involved the researchers in designing an interactive multimedia program for a group of eight pre-service mathematics teachers.

There is much left to guess about these pre-service maths teachers. The researchers fail to indicate why they were chosen as the sample, or provide any background information such as gender, class, educational experience, relationship to each other or the researchers, etc. All of these aspects of the research subjects could be seen as confounding variables that are not being addressed or controlled (Brown and Dowling, 1998, 146). Do any of these confounding variables affect the amount of higher order thinking the subjects may use? There is no attempt to remove or account for this potential bias in the findings from the empirical setting other than one instance where the authors indicate that the students knew each other before the research began and were comfortable working together. This they admit may have affected the results of the study (Herrington and Oliver, 1999, 17). This acknowledgement that the students knew each other before the study is, however, later ignored when the researchers are analysing their data. This failure to elaborate on the makeup of the research subjects and rationale for their use, of what can be assumed to be an opportunity sample, limits the generalizations that can be made from the study findings. This will be touched on again when the analysis of the information collected is examined.

To what extent did being the subject of the study affect the behaviour of the pre-service maths teachers? Again the reader lacks the necessary information to make a judgement. The research subjects were video taped while using the program, and this may have been

an effort to limit the effects of being studied, but little else is indicated about the actual research setting. If, for example, this group were students of Herrington and Oliver, how might being observed by their professors affect their engagement in the activity? It is difficult to control for the Hawthorne effect outside the laboratory (Brown and Dowling, 1998, 40). If however the researchers had elaborated more on their sampling procedures and nature of the sample, aspects of its effect on the findings could have been taken in to consideration in relation to the findings. If this is indeed an opportunity sample, "the validity of the generalization relies on the researcher marking out the continuities and discontinuities between the setting and the empirical field in an ad hoc manner" (ibid, 30). In this case the reader is not given the information necessary to mark out differences between the study and the empirical field of the classroom.

The multimedia application was developed around the previously mentioned seven criteria, which the researchers claim are present in a genuine situated learning environment. The students were organized into four pairs and given "a complex and sustained authentic task to investigate" (Herrington and Oliver, 1999, 6). This 'authentic' task involved responding to two documents, one written by a 'parent', complaining about assessment practices in the schools mathematics class and a memo by the 'mathematics coordinator' to the maths teachers of the 'virtual' school. The research subjects then took this information and examined it with the aid of the multimedia program, in order to produce a new assessment plan to be implemented in the 'virtual' school.

Herrington and Oliver state that their multimedia program consists of "(a) short video clips of assessment strategies being used in classrooms, (b) interviews with teachers and

students on the assessment strategies, and (c) a variety of text documents, such as a description of each strategy, sample resources and mathematics experts' views" (Herrington and Oliver, 1999, 5). However, the authors do not state how the contents of the program relate to the criteria that they state are necessary for a situated learning environment. Even if the readers assume which portions of the program relate to the criteria, they are not given any examples of the various aspects of the program. Instead the readers are given a figure, which represents the graphical interface for the multimedia program. Although the buttons in the interface are labelled, this does not satisfactorily explain how the contents of this program relate to the criteria necessary for a situated learning environment. Without any elaboration as to the contents of the program the readers are left wondering if this application meets the standards the authors set for a constructivist situated learning environment. As a result the authors fail to demonstrate the authenticity of the task.

The empirical setting constructed by Herrington and Oliver also needs to be examined in relation to the key antecedent work on situated cognition, particularly that of Lave and Wenger. Lave and Wenger (1991) acknowledge that examining schooling from the perspective of situated learning or legitimate peripheral participation "will turn out to be a fruitful exercise" (p. 41). However, situated cognition is, according to them, "an analytical viewpoint on learning", rather than a pedagogical strategy, educational form, or teaching technique (Lave and Wenger, 1991, 40). It can be argued that Herrington and Oliver are attempting to use a situated learning environment as a teaching technique by constructing this empirical setting. They are "attributing a prescriptive value" to situated learning by implementing or operationalizing it "for educational purposes" (Lave and

Wenger, 1991, 41) through their empirical setting. This is contrary to Lave and Wenger's view. In these terms Herrington and Oliver's interpretation of situated cognition, and its examination in this empirical setting, does not fit with some of the key work on the subject.

Information Collection and Analysis

The research subjects were videotaped while using the multimedia program, with their 'talk' used to produce the empirical findings. The researchers looked for each "instance of a type of talk" (Herrington and Oliver, 1999, 12). Then, using the classification chart developed from the theoretical field, they categorised these types of talk into three non-higher order types, social, procedural and lower order, and six sub categories of higher order types. While the focus of the study was on the higher order types of talk "it was also necessary, however, to draw up similar criteria for the classification of talk which could not be considered higher order"(ibid, 10). Thus we are given the three "non higher order thinking"(ibid, 10) categories.

The higher order categories were "(a) uncertainty, (b) path of action, (c) judgement, (d) multiple perspectives, and (c) imposing meaning or metacognition" (Herrington and Oliver, 1999, 16). Each sub category, both with higher and lower order talk, was then given a definition. The information, collected from the videotapes, was coded with this classification scheme. This analysis of the information resulted in the data that is the basis for the conclusions the authors reached. It is therefore important to examine the

validity of this classification scheme in order to assess the conclusions that Herrington and Oliver assert.

The classification scheme was “specifically developed for use with multimedia” (Herrington and Oliver, 1999, 20). It was not mentioned how many coders were employed, or whether the schedule was piloted. Brown and Dowling contend that a new schedule should be piloted, so that problems that may exist in the schedule can be changed before it is used in actual research (Brown and Dowling, 1998, 53). The lack of testing for “intercoder reliability” may account for the lack of “explicitness, coherence, and clarity of the framework” (ibid, 53) that was used to classify the information. The given framework fails to provide “ a range of categories that enable all forms of the phenomenon in which you are interested to be easily classified”(ibid, 53). Nor does “each coded event fall into just one category” (ibid, 53). The researchers themselves acknowledge that their “interpretation of higher order thinking may be too liberal” (Herrington and Oliver, 1999, 20). They go on to provide the example of how the categories can be confused and misinterpreted. “For example, many of the comments and statements classified as *Uncertainty* and *Path of action* may actually be better defined as *Lower-order* thinking, simply because such comments may require little mental effort” (ibid, 20). The ability to classify the information gathered from the students into several different categories of thinking, both higher and lower, on the schedule, is more prevalent than the researchers are willing to admit. For example, the following was cited in the study as students showing *metacognition*: “G: What was that? I didn’t get any of that. E: I don’t know what to do. Where is that piece of paper I had

before?" (ibid, 14). *Metacognition* was defined by Herrington and Oliver (1999), as "comments which showed that the students were aware of their own thinking and performance, and comments related to the use of this awareness to improve performance" (p. 14). However, this exchange could also be classified as *Task*, which the authors define as "any exchange of information related to the task (the formal requirements of the oral or written report)" (ibid, 13).

The authors not only fail to make the categories explicit, but also fail to identify which section of the passage they are coding from the examples given. They explain that in some of the transcribed passages "more than one instance or utterance is provided to enable the reader to grasp the context of the comment rather than view it as an isolated statement" (Herrington and Oliver, 1999, 14). So the reader could mistakenly code the wrong section of the examples. This could be what has led to the coding confusion that was mentioned above. Herrington and Oliver are using what Brown and Dowling (1998) term as "event coding", that is recording "specific events as they happen, rather then waiting for a fixed period between the coding events" (p. 52). In this case the event is defined as "each unit of meaning, that is, each instance of a type of talk as it occurred" (Herrington and Oliver, 1999, 12). This unit of analysis was determined to be the best of the others described by the authors. They considered coding the complete passage but decided it would be too difficult to decide where the passage began or ended. They also examined the possibility of coding each utterance, but decided it would be too difficult to code utterances in which several types of talk occurred. The choice to code each instance of a type of talk was made, because it allowed them to code utterances by the students in

which more than one type of talk was used (ibid, 12). However the difficulties described by the authors in the first two types of coding were not eliminated by the choice they made, because they fail to adequately define the unit of analysis, which is to say, to explain their principles of recognition of a unit of analysis and its categories.

So where in the transcript of the conversation between G and E is the instance of metacognative talk? When does it begin, and end? How many different types of higher order thinking are there in this passage? The authors do not tell us. Even if the reader assumes that this passage only contains metacognition it is difficult to say where the unit of analysis begins or ends, because the reader lacks a clear definition of the unit. If the reader follows the research design indicated in the article, the instance of metacognative talk is when this type of talk is used, according to the definition provided. The result is that for each researcher, using this observational schedule, the unit of meaning could be different. This again brings into question the intercoder reliability of this schedule. When using event coding "closer attention has to be paid to the definition of the unit that is being recorded and to recognition of the beginning and end of the codable event" (Brown and Dowling, 1998, 52). In this case we have overlapping definitions within the schedule resulting in confusing units of analysis that cannot be coded. The authors themselves even allude to this by stating, "such frameworks inevitably have different features which may lead to different interpretations of meaning" (Herrington and Oliver, 1999, 7). If different interpretations of meaning can be made using this schedule then it is an unreliable schedule, and should be re-developed.

A more fundamental question also needs to be addressed, when considering the validity and reliability of the analytic frame. This is the assumption that Herrington and Oliver make in regards to talk. The reader is asked to assume that talk indicates thought (Herrington and Oliver, 1999, 7). The authors contend that there is little option but to "count talk as the outward representation of thought" in order to measure cognition (ibid, 7). To support this position the researchers cite the work of von Wright, who states that "the advantage of social contexts for learning is that they elevate thinking to an observable status" (Herrington and Oliver, 1999, 7). The author's contention is that in a social setting "the sharing of thoughts is critical to communication" (ibid, 8). Therefore the talk that occurs between the maths teachers is their thoughts made observable because of the social setting.

They do admit to this position being tenuous and provide the reader with evidence that this may be questionable. This evidence, which they provide in the form of cited works from Halliday and Young, goes quite far in invalidating the tenuous assumption of the authors. According to the authors, Halliday contends that in everyday conversation, speech is smooth and articulate, "because the speakers are not having to think all the time about what they are saying", while in an academic social setting, thinking, to form an argument interrupts this smooth flow (Herrington and Oliver, 1999, 7). In other words thinking becomes silent and unobservable in an academic setting. Although the situated learning environment designed by the authors is not necessarily an ordinary, everyday situation, it is a social situation involving two people who knew each other before the study. Therefore, it could be argued that the subjects are not required to think very much

about what they are saying, because they are already comfortable working together. If, however, we consider it an academic situation where the subjects have to think about what they are saying, then the actual thinking is done silently, and only the outcome of that thinking is observable in the form of speech.

Young, according to Herrington and Oliver, also contends that students will fall quiet while engaging in problem solving "possibly due to cognitive workload" (Herrington and Oliver, 1999, 7). Young, in this case, is examining studies in which the subjects are asked to articulate their thoughts out loud while solving problems. The authors feel they overcome this problem by creating a social situation, forcing the subjects to share their thoughts. Does the social setting overcome the problems with measuring thought that Halliday and Young describe? Herrington and Oliver seem to admit that they don't have the evidence to make a strong argument in favour of accepting speech as thought, when they state that the social situation only vindicates their position to some extent. They cannot and do not make a strong argument for their case, so the reader is left wondering why they chose the unit of analysis they did.

Conclusion

The data produced from the classification schedule does indeed indicate that in each of the four groups the majority of talk was higher order. Herrington and Oliver therefore feel quite justified in saying that their study shows "that all the students used a substantial proportion of higher order thinking in the situated learning environment" (Herrington and Oliver, 1999, 19), and that "the constructivist nature of the learning environment

provided greater opportunities for students to engage in higher order thinking" (ibid, 21). The multimedia program, built around the tenets of situated learning provided a "learning environment capable of supporting and maintaining substantial levels of higher order thinking" (ibid, 21). The findings also contradict "many previous studies exploring studies exploring students' cognition and thinking" using multimedia (ibid, 22). The net result of this is that the original theoretical position of the authors is validated and extended to all learning environments that incorporate situated learning. This generalization can be accepted only if substantial aspects of the research design and its implementation are ignored.

In order to examine the validity of the argument in this article the reader first needs to consider the relationship between the concept variables and the indicator variables (Brown and Dowling, 1998, 26). In this case the indicator variables that are being looked for, using the classification schedule, are based on the concept variable of higher order thinking, a concept that is never clearly defined. Even if the reader assumes a definition for the concept variables, the schedule used to indicate them lacks reliability. The categories are not mutually exclusive. The unit of analysis is confusing and ill-defined. Furthermore the reliability does not appear to have been verified by piloting the schedule.

There are additional problems with the generalization of this study. The authors failed to provide information about the nature of the sample, and the researchers ignore potential confounding variables within the sample. Only in one instance do Herrington and Oliver address one of these variables. As they examined the types of talk used by the groups, it

became apparent that group 3 used more of the *path of action* and *uncertainty* types. It was suggested that this might be because the rest of the groups, but not group 3, "had all worked with their partners on several previous occasions" (Herrington and Oliver, 1999, 18). When the authors go on to discuss that there is an apparent lack of hierarchy in the students' talk, no acknowledgment that the subjects of the study knew each other and had worked together before, is made. This lack of hierarchy in the student's talk is seen as supporting the work of Resnick and Newmann, and counter to the theorists Bloom and Gagné. If a hierarchical structure was to be found in the conversations of the students, "it might be expected that the students would begin with a little social talk to establish their working relationship" (ibid, 21). However, the authors have already stated that the students in all the groups except one, knew each other and had worked with each other previously. Why would they need to establish a working relationship in this case? The author's willingness to ignore variables within the sample, that may have affected the outcomes of the study, makes it difficult to accept that this sample is truly representational of the empirical field about which they wish to generalize.

The net result is that enough doubt can be thrown on this study in the areas of the sample used, how the theoretical position was operationalized into the empirical setting, the use and construction of the observation schedule, and how the theoretical position itself was glossed over, to make any conclusion reached by this study, questionable.

Bibliography

Brown, A.J. and Dowling, P.C. (1998). Doing Research/Reading Research: a mode of interrogation for education. London: Falmer Press.

Herrington, J. and Oliver, R. (1999). Using Situated Learning and Multimedia to Investigate Higher-Order Thinking. Journal of Interactive Learning Research. 10(1), 3-24.

Lave, J. and Wenger, E. (1991). Situated Learning: Legitimate Peripheral Participation. Cambridge: Cambridge University Press.