FORMATIVE ASSESSMENT IN INQUIRY-BASED SCIENCE EDUCATION USING INTERACTIONS ON-THE-FLY

Pasi Nieminen[1], *Catarina F. Correia*[2], *Markus Hähkiöniemi*[1], *Natasha Serret*[2], *Jouni Viiri*[1], *and Christine Harrison*[2] [1] University of Jyväskylä (JYU)

[2] King's College London (KCL)

Abstract:

This paper is focused on the characterization of informal formative assessment conversations (i.e., interactions on-the-fly) from a methodological perspective. Interactions on-the-fly are unexpected teachable moments in which the teacher tries to probe students' understanding and use that information to support their inquiry process. One of the coding systems used (ESRU framework) was adopted from the research literature (Ruiz-Primo & Furtak, 2007). Two other systems described here were developed and previously reported in two papers written by the authors (Correia, Nieminen, Serret, Hähkiöniemi, Viiri, & Harrison, in press; Nieminen, Hähkiöniemi, Leskinen, & Viiri, 2016). This paper presents these three coding systems together and discusses their features to characterize interactions on-the-fly. Examples for coding interactions on-the-fly from Finnish and English physics lessons are presented. For example, on the one hand, on-the-fly discussion proceeds as follows: The teacher initiates an on-the-fly episode by eliciting information using an open-ended (divergent) question, as the discussion progresses the teacher starts narrowing down the discussion using a closed-ended question (convergent), and closes the discussion by giving a short mini lecture. On the other hand, when guiding students' inquiry, the teacher can collect information rapidly (quick interpretation) or he/she can use series of probing questions (further probing). In the former, the guidance is more authoritative as it is based on correspondence of the students' and the teacher's ideas. In the latter, guidance is based more on the student's ideas representing more dialogic way to support learning.

Keywords: Formative assessment, teacher-students informal assessment conversations, inquiry-based learning, science education.

INTRODUCTION

Over the last decade there has been a substantial policy push towards the adoption of Inquiry Based Science Education (IBSE) in Europe (Rocard et al., 2007). IBSE has been associated with increased students' motivation and interest in science, supporting the development of inquiry competences and conceptual understanding, and motivating teachers in teaching science and about science (Minner, Levy, & Century, 2010). Within Europe several EU funded projects have been focusing on the development of inquiry-based science and mathematics teaching practice utilizing a diversity of approaches (Scientix website). Surprisingly little attention has been given to supporting teachers in developing their assessment practice in the context of IBSE. Assessment and pedagogy go hand in hand (Black & Wiliam, 2009), and teachers practice is often driven by what is assessed (Harlen, 2013). In this context, there is an obvious need to clarify the links between IBSE pedagogy and assessment, and to develop reliable and valid ways of assessing inquiry, for both formative and summative purposes.

EU FP7 project Assess Inquiry in Science, Technology and Mathematics Education (ASSISTME/2013-2016) is contributing in different ways to developing our understanding of the links between assessment and pedagogy in inquiry-based education. ASSISTME has been

focusing on developing and trialing different assessment methods to assess inquiry competences. The links between formative and summative assessment are also being explored. The aim of the project is to issue guidelines, providing a roadmap for the assessment of inquiry competences at both national and international levels.

This paper describes the methodology used to study informal formative assessment conversations (i.e., interactions on-the-fly) in inquiry physics lessons in Finnish and English classes. On-the-fly interactions are unexpected teachable moments in which the teacher tries to probe students' understanding and use that information to support their inquiry process. The paper is synthesis and extension of two conference papers (Correia, Nieminen, Serret, Hähkiöniemi, Viiri, & Harrison, in press; Nieminen, Hähkiöniemi, Leskinen, & Viiri, 2016). All these studies are part of ASSIST-ME in which we have collected a large body of audiovideo data from Finnish and English physics and mathematics lessons in order to study onthe-fly interaction. Our preliminary results are based on a small sample of the whole data. Thus, we are still working with the data and we will continue analysing and publishing in collaboration between King's College London (KCL) and University of Jyväskylä (JYU).

Interactions on-the-fly provide teachers with a significate amount of evidence of student learning, that is collected in real time as students engage with inquiry activities. This creates opportunities for teachers to make informed decisions (in real-time) regarding next steps to support students in their learning (Harrison, 2015). The evidence of learning that can be collected through interactions on-the-fly depends on how these conversations are set up, on what type of questions are used, and to what extend are the conversations based on students' ideas. Studies on classroom talk show that dialogic teaching contributes to enhanced student participation and learning (Mercer & Dawes, 2014).

Although there is a growing body of knowledge on what good classroom talk looks like, little is known on the characteristics of classroom talk in the context of inquiry lessons, and on how these interactions contribute to good formative assessment practice. In the context of this study, we address the following research question:

How can interactions on-the-fly be charaterised taking into account teacher's formative action and the learning potential that arises from it?

THEORETICAL FRAMEWORKS

Students engagement with inquiry requires involvement in questioning, reasoning, observing, conjecturing, data gathering and interpreting, investigative practical work and collaborative discussions, and working with problems from and applicable to real life contexts (Anderson, 2002). The formative use of assessment evidence is a key pedagogical driver to support and advance students' learning (Black & Wiliam, 2009) and is as applicable to inquiry as to other science activities. Formative assessment (Assessment for Learning) in inquiry lessons requires that teachers recognize and collect the assessment data that arises directly from lessons, and use it to advance learning.

Interactions on-the-fly are informal formative assessment conversations which are not planned beforehand, but take place spontaneously when the teacher recognizes appropriate opportunities to support students in advancing their learning (Shavelson et al., 2008). These informal assessment conversations provide good opportunities for formative action. The aim of the teacher in these interactions is to probe students' understanding and thinking in real time, so as

to collect evidence of where students are in their learning, decide on the next steps that will move the learning forward, and enact a formative intervention.

In this context a simple initiation-response-evaluation (IRE) sequence would not be considered as good formative practice because evaluative feedback provides no information regarding next steps that can advance learning. In our view, good formative assessment practice relies on teachers making use of the learning evidence that was gathered to make informed decisions on next steps and acting on them. The challenge for the teacher is recognizing that evidence, deciding what that evidence indicates in terms of learning and then forming an appropriate response that takes learning forward.

Ruiz-Primo and Furtak (2007) model assessment conversations as a four-step cycle where the teacher **E**licits information from the student by formulating a question, the **S**tudent responds, teacher **R**ecognizes the student's response, and then **U**ses the information collected to student learning. This is known as the ESRU cycle, and has been empirically derived from analysis of informal assessment conversations that took place during inquiry lessons. This provides a useful framework to characterize interactions on-the-fly because it enables a differentiation between instances where teacher is *eliciting* information from the student (e.g. formulating a

question to probe student's ideas) and instances where the teacher is *using* the information that was collected to push the learning forward (e.g. building on a student's answer). Further, the authors zoom in into the Eliciting, Recognizing and Using main categories and provide a list of subcategories for each of these. These sub-categories are then clustered as conceptual or epistemic (for more details see Ruiz-Primo & Furtak, 2007, Table 2 on p. 63).

Ruiz-Primo and Furtak (2007) observed that for a group of 4 middle school science teachers (each with 12 observed lessons), the students' performance in pre-post multiple-choice questionnaire and embedded assessments were higher for classes with teachers who used more complete ESRU cycles. These authors argue that the more cycles a teacher completes, the more likely it becomes that the information gathered from students was used in a helpful way for learning purposes.

Torrance and Pryor (2001) describe formative assessment practices as placed along a continuum between convergent and divergent formative assessment. Convergent assessment is conceptualised as creating opportunities to find out if students know, understand and can do certain things. Convergent assessment is typically associated with use of close-ended or pseudo-open ended questions. Divergent assessment is conceptualised as creating opportunities to explore students' thinking and understanding. For Torrance and Pryor (2001) it is about finding out what students know, understand, and can do. Divergent assessment is typically associated with the use of open-ended questions that promote discussion.

Based on their observations, these authors developed a descriptive and analytical framework of the processes of formative assessment in the classroom (see Torrance & Pryor, 2001, Fig. 2 on p. 620). This framework lists 14 different descriptors of what the teacher does in the classroom. Each descriptor is accompanied by an interpretation of the teacher's possible intentions and the possible positive effect it may have on the students. These descriptors are clustered around four dimensions that characterise teachers' actions: questioning, observation of process and products, making explicit the task and quality criteria, and feedback and judgement. Teachers' observed formative practice is then characterised by mapping it against the descriptors and their dimensions.

Torrance and Pryor's work in 2001 is situated in the context of an action research project with teacher-researchers. Teachers found that the convergent and divergent approaches to assessment helped them conceptualise their own formative assessment practice. Teachers also recognised that a divergent approach is more likely to create an environment (both socially and intellectually) more conducive to enhance learning.

In the context of this study we focus on exploring how ESRU-framework can be completed by investigating on micro level types of ESRU moves and by characterizing on macro level interactions on-the-fly episodes as a whole. In micro level we utilize the framework by Torrance and Pryor (2001) and on macro level create the characterization through data driven analysis.

METHODS Research context

This paper presents three coding systems which were developed (excluding ESRU) and implemented in four case studies (Table 1). Cases 1 and 2 are reported in Correia et al. (in press) and cases 3 and 4 Nieminen et al. (2016). All teachers have more than 10 years of teaching experience in mixed comprehensive schools in south of England and central Finland, respectively. English school was a Girl School and Finnish schools were public as c. 97% of Finnish schools are.

Table 1

Participants

Case	Topic	Data	Students	Coding
1 Jane; England	Mechanics and electromagnetism	2x90 min	Y11; 16 years; *n* = 18	ESRU & KCL
2 Anna; Finland	Electricity	3x45 min	Y8; 14 years; *n* = 19	ESRU & KCL
3 James; Finland	Optics	1x 45 min	Y7; 13 years; *n* = 13	ESRU & JYU
4 Maria; Finland	Division (math)	1x45 min	Y3; 9 years; *n* = 23	ESRU & JYU

Data analysis

All interactions on-the-fly were identified from the audio-video (Finland) and audio recordings (England). These were episodes were teacher-student, teacher-student group or teacher-whole class discussions took place. Each episode focused on one underlying theme, such as the teacher discussing with a group of students on how to measure voltage. The end of an episode was marked by either a change in topic of discussion or starting a discussion with another group of students. The discussion focused on conceptual, procedural knowledge and/or inquiry. Further, only episodes which included the Use component were considered as interaction on-the-fly episodes.

All on-the-fly episodes were transcribed verbatim. The transcripts were analysed using three independent coding schemes for characterizing interactions on-the-fly. Two of the coding schemes (ESRU and KCL) were adapted from the theoretical frameworks presented above. Third coding system (JYU) arose from data. The coding systems are described in the subsection below.

In terms of interrater reliability the authors have extensively discussed the application of ESRU and KCL coding systems using one episode from a Finnish lesson as paradigmatic example. The code attribution for the paradigmatic example was discussed in detail until we have reached full agreement. This gave rise to the development of a coding manual that was applied by one and in some cases two independent researchers within each research institution. In the case of JYU coding system, ESRU codes were used first. After that two researchers examined the interactions on-the-fly episodes. The episodes were carefully compared to each other and searched for similarities and differences in formative assessment discussions. As a result, four different kinds of formative assessment discussions were identified.

ESRU coding system

The main categories of the ESRU model (Ruiz-Primo & Furtak, 2007) are teacher Elicits (E), Student responds (S), teacher Recognizes (R), and teacher Uses (U). Eliciting is conceptualized as all instances when the teacher creates an opportunity to collect evidence of learning. Using is conceptualized as all instances when the teacher creates an opportunity to make use of the learning evidence that was gathered to either e.g. ask a follow up question or provide feedback to students.

The ESRU main categories were applied to each speaking turn (one speaking turn is defined as the non-interrupted speech utterances of one single individual, this could be either the teacher or student). One speaking turn can include more than one ESRU code. For instance, the teacher may first recognize student answer and use that information afterwards. ESRU codes reveal eliciting and using which are important phases when the teacher looks for information about student understanding and helps them to proceed or express their thinking.

The following excerpt provides an example of the application of ESRU coding system. The excerpt was taken a physics inquiry lesson in England (Jane). The students were asked to investigate the factors that influence the strength of an electromagnet. One of the lesson's learning aims was to identify a mathematical relationship emerging from the data collected during the inquiry activity. The excerpt presented here illustrates a conversation between the teacher and one of the students in a group regarding the group's findings.

Turn	Speaker	Transcription	Code 1
	Jane	So what you get?	E
2	Student	It is all kind of awkward…	S
3	Jane	Go on	U
4	Student	So I don't know, it is as if …	S
5	Jane	So that could be a straight line or it could be a relationship that tails to that end. How would you know? What could you do to improve it? Where are you missing data?	U E
6	Student	Here (points at results)	S
7	Jane	Yes, fill those gaps in and those gaps in and then you would know.	U

KCL Coding System

This coding system was adapted from Torrance and Pryor (2001) and Alexander (2006). Based on Torrance and Pryor's description of convergent and divergent formative assessment practice,

we developed two categories to describe teachers' use of questions or follow up comments (TD-Divergent and TC-Convergent).

Teacher Divergent talk refers to all instances in the dialogue where the teacher asks questions to probe and encourage students' thinking and to promote discussion. The questions used are open-ended (typically, *how?* and *why?*), and promote higher order thinking. The focus of this type of talk is not only to unveil what the student is thinking but also create an environment where the thinking can be extended, so ultimately it is about opening up and keeping the discussion going.

Teacher Convergent talk, on the other hand, refers to all instances in the dialogue where the teacher asks questions to check if students are on an intended path, and if they know something. The questions used in this context are typically close-ended questions that promote lower order thinking (mostly recall of factual knowledge). As the focus of this type of talk is to check if students are on an intended path, it is likely that this type of talk reveals repeating or reformulating the same questions until students give the answer the teacher is looking for.

Based on Alexander (2006) description of instances of exposition and instruction in classroom dialogue, we have developed yet two more categories to describe teacher's feedback (TL-teacher Lectures and TA- teacher affirms).

Teacher Lectures refers to all instances in the dialogue where the teacher provides information through exposition or direct instruction. According to Alexander (Chapter 3, p. 28; 2006, these are instances where the teacher either imparts information (e.g. explaining facts, principles or procedures) or simply tells students what to do. The focus of this type of talk is generally to support pupils in making sense of something, to fill in a gap of knowledge, or to give instructions on procedures. The locus on control in this case lies exclusively with the teacher.

Teacher Affirms refers to all instances in the dialogue where the teacher repeats or reformulates students' contributions. The focus of this type of talk is to signal to students the acceptance of a point, idea or fact, and/or to summarise a student's contribution. As in Teacher Lectures, the locus of control in this case lies with the teacher.

All four categories are mutually exclusive and were used to code each speaking turn. A fifth category was introduced to describe references to inquiry and the nature of science (NoS).

Given that these conversations take place in inquiry lessons, we assume that teacher would make explicit references to inquiry and NoS. These five categories are presented in Table 2. For further details, see Correia, Serret, Black, & Harrison (2016) (manuscript in preparation).

The coded episodes were analysed seeking to identify emerging patterns in teacher behavior within these episodes so as to characterise their formative assessment practice.

Table 2

Convergent/Divergent codes adapted from Torrance & Pryor (2001)

Codes	Description
TD – Divergent	Questions/comments that open up discussion (how/why?); probing thinking
TC – Convergent	Questions/comments to check if students are on track; focusing the discussion
TL – Lectures	Give information through exposition or direct instruction

TA – Affirms	Acknowledge/repeat/re-formulate student(s) contribution
TN – Inquiry	Make explicit reference to inquiry (general principles or procedures) and to NoS

The following excerpt provides an example of application of KCL coding system. It is the same excerpt as presented above.

Turn	Speaker	Transcription	Code
1	Jane	So what you get?	TD
2	Student	It is all kind of awkward…	
3	Jane	Go on	TD
4	Student	So I don't know, it is as if …	
5	Jane	So that could be a straight line or it could be a relationship that TC sort of tails to that end. How would you know?	TD
		What could you do to improve it?	TD
		Where are you missing data?	TC
6	Student	Here (points at results)	
7	Jane	Yes, fill those gaps in and those gaps in and then you would know.	TL

Figure 1 illustrates the pattern of the conversation presented above combining ESRU and KCL coding systems. ESRU codes are in the plot area in temporal sequence and KCL codes in y-axis.

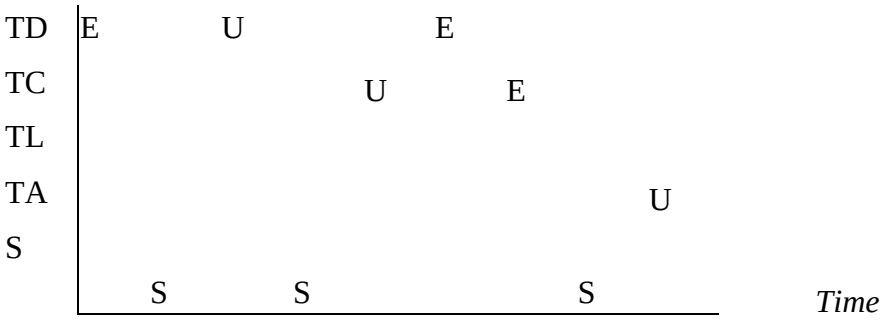

Figure 1. Pattern emerging from the analysis of one interaction on-the-fly taken from a physics inquiry lesson in England.

JYU Coding System

An alternative way of looking at interactions on-the-fly is to move from a micro-analysis of each speaking turn into a macro analysis of the whole on-the-fly episode (Nieminen, Hähkiöniemi, Leskinen & Viiri, 2016. We characterised on-the-fly interactions in two dimensions: collecting evidence of learning and giving feedback. Collecting evidence of learning can be split into eliciting information and making a quick interpretation, or eliciting information and probing further (typically using follow up questions). Feedback can be split into making comments or asking questions in order to help students to take the next step or to express their current thinking. A combination of these dimensions leads to a characterisation of each episode as falling under one of the four types of interactions on-the-fly (Table 3). Next, we present two examples of these categories.

Table 3

Four types of on-the-fly formative assessment interactions

	1 Quick interpretation	2 Further probing
A Help to take the next step	1A	2A
B Help to express thinking	1B	2B

Quick interpretation and helping students to proceed (1–A). In many cases, the teacher got some information of students' progress and started quickly to guide students. For example, in the physics lesson, James initiated the following discussion.

Turn	Speaker	Transcription	Code
1	James	How is it going here?	E
2	Alison	If we understood this…	S
3	James	Could you tell me how did you understand it?	E
4	Diana	So, that we have to place the object there and then look that where	S
5	James	we can see it. Yes. Okay. If we think that this is Alison's eye, so the eye can move along this line here [shows by hand]. Yes. Then you can… here it is said that: "sketch the observer in different places". You can use a ruler for example. In which line segment the eye could be for instance.	U

In this discussion James elicits, makes quick interpretation about students' difficulties and in turn 5 starts to guide students how to write down an observation.

Probing more students' thinking and helping students express their thinking (2–B). There were instances in which after the first elicitation, the teacher continued to probe the students' ideas. After that, the teacher either guided students to proceed or to express their thinking as James did in the following example from the physics lesson. He probed many times until he got information about students' ideas. After that he helped students to express their idea and prompted them to write it down.

Turn	Speaker	Transcription	Code
1	James	What is going here?	E
2	Student	Well, when… if we put that… that when…	S
3	James	Are you working with conclusions?	E
4	Students	Yeah	S
5	James	Okay. Do your prediction and your observation differ?	E
6	Students	Yeah.	S
7	James	In what way?	E
8	Student	We thought as if I sit here. Then the mirror is there, so it would be mirrored to this corner but it is in that corner.	S
9	James	H'm. So, you were here. Here is the object. You sat here. And you see the image… you supposed that it is here, but is in this side.	R
10	Student	Yes.	S
11	James	Well. So, you can answer: Do your prediction and	U

		observation differ?	
12	Student	Yeah.	S
13	James	Yes. How does the person's location influence the location of the image?	U
14	Student	It is mirrored like… If you look from side the object is mirrored to that corner where you sit or stay.	S
15	James	To that corner? So, you considered that the image is like in the different side from the object than your eye, but it is in the same side…	R
16	Student	Yeah.	S
17	James	…than your eye. Okay. Yes it is. Yes. Then just write.	R

The E-S-E-S-… -structure continues for turns 1 to 9 where James recognizes students' ideas. The turn 11 is evident point in which he starts to use information to help students to express their ideas.

DISCUSSION

Reflection on ESRU coding

The application of ESRU coding system enabled us to separate teacher's actions into Eliciting and Using. Eliciting is associated with creating opportunities to collect evidence of learning and Use is associated to creating opportunities to make use of that evidence of learning by asking a follow up question or giving feedback. This distinction is very important when charaterising teachers' formative assessment practice, because effective formative practice requires collecting evidence of learning and acting on it leading to responsive teaching (Harrison, 2015).

However, there are several limitations in ESRU coding which led us to develop KCL and JYU codes. There are difficulties which can make the coding process quite challenging. In several instances, we found it difficult to differentiate between Eliciting and Using. Often both codes could be attributed depending on an interpretation made. This interpretation could be based on what preceded a certain utterance (previous speaking turns), or an interpretation based on what come next (following speaking turns). Another limitation to the application of ESRU codes was the poor interrater agreement when coding with the ESRU sub-categories. Some sub-categories were found to be ambiguous and often did not describe the teacher actions that took place during the interaction, leaving much of our data without a code.

One problem in using ESRU main codes is that U can be either a question or a mini lecture which are very different kinds of actions of the teacher. This issue can be addressed using KCL codes to describe the nature of U as discussed later. Another problem when ESRU main codes are used is that students' role is just to answer the questions. Finnish data revealed that students asked lot of questions and they initiated significant part of on-the-fly episodes. Thus, in our definition an on-the-fly episode must include U component but E is not necessary.

Finally, Ruiz-Primo and Furtak (2007) reported a relation between complete ESRU cycles and student learning. This relation is plausible, for example, if a dominant structure in a class is ESR, i.e. the U component is only rarely used by the teacher. However, this does not mean that incomplete cycle, such as ESR, would be an inevitably poor move in terms of formative assessment. Occasionally, it can be the teacher's pedagogical choice to stand aside without giving feedback or probing. In addition, E is often missing when students are doing inquiry in groups as noted before. In these situations, the teacher gets enough information from students

questions or statements, and the teacher's first move in the episode is U or R (recognising). We have also found that often an instance of Use appears after several instances of teacher Eliciting and the Student responding.

Reflection on KCL and JYU coding

The application of the KCL codes, in tandem with the ESRU coding, allowed us to characterize patterns of interaction at a micro level. Data analysis shows a recurrent pattern in which teacher typically initiate an interaction by using combination of open-ended and closeended questions and comments to keep the discussion going, before closing down the discussion with instructions. According to Torrance and Pryor (2001) the use of open-ended questions and comments is associated with divergent formative assessment practice, while the use of close-ended questions and comments is associated with a convergent formative assessment practice. Torrance and Pryor highlight that both types of practice contribute to learning. Moving from divergent to convergent questions is somewhat similar as opening up discussion through dialogic talk and closing it down through authoritative talk (Lehesvuori et al., 2013). Besides the differentiation between divergent and convergent questions (in both instances of E and U), KCL codes helps to see the difference between U as a question (TD or
TC) and as a lecturing move (TL). This gives much better view about the teacher's action in these situations.

Unit of analysis in ESRU or KCL coding is speaking turn or part of it. JYU coding gives a description for the whole episode. We characterized four kinds of on-the-fly formative assessment discussions. Getting data of student learning and giving feedback are two dimensions in this characterization. These two dimensions are essential in any formative assessment (Black & William, 2009) and in on-the-fly discussions (Ruiz-Primo & Furtak,
2007). We found that teachers get information about students' learning by quick interpretation or by further probing. Feedback was given in order or to help students to take the next step or to express their current thinking.

When a teacher guides students to take the next step this can be done after a quick interpretation or after further probing. In the former, the guidance is based on correspondence of the students' and the teacher's ideas. In the latter, guidance is based more on the students' ideas. Thus, the former is more authoritative guidance and the latter more dialogic guidance in the sense of Mortimer and Scott (2003) and Lehesvuori et al. (2013).

The feedback can focus also on getting the students to express their current thinking. This can be based on a quick interpretation of where the students are in their learning or it can include further probing. Through further probing the teacher gets more information about the students' learning and at the same time the students get feedback to explain more.

Similarly, as KCL coding, JYU coding gives richer information about teacher actions compared to ESRU coding. In addition, JYU codes characterize teacher's ways to get the information of student learning and to advance learning. Sometimes JYU coding types quick interpretation and further probing can be seen in ESRU structure. For example, the pattern Student initiation – Recognition – Use implies that there was no further probing. On the hand, the pattern Elicit – Student response – Elicit – Student response implies that the episode most likely contain further probing. However, ESRU coding alone cannot be used to identify whether further probing exists because often further probing happens through follow-up questions that use the information of the previous student response and are thus coded as Use.

In addition, interpreting whether an episode is of type "help to take the next step" or "help to express thinking", includes analyzing the contents of Use.

One interesting issue is the relation between JYU and KCL codes. We have some ideas about that but we have not analysed our data yet from that perspective. For example, one of our

hypotheses is that divergent questions would be more common in further probing episodes than quick interpretation episodes. In other words, it is probable that teachers should use more open-ended questions (TD-E and TD-U), even series of such questions, when they try to probe deeply students' ideas. James's second example in the JYU coding system –section seems to be this kind of episode, but more analysis must be done that we have proper understanding about the relation between JYU and KCL codes.

From a socio-constructivist perspective on learning (Vygotsky, 1978), students learn through a negotiation on meaning making. From a formative assessment perspective, the use of open-ended questions in a genuine dialogic environment is more likely to support this meaning making process. This is because, in these contexts, teachers create opportunities for students to engage in deeper thinking, to articulate and reveal their thinking to others and to be open in considering next steps in learning. Research shows that open-ended questions tend to elicit more elaborate answers that rely on an articulation of ideas, and therefore involve deeper thinking (Chin, 2007). Similarly, further probing students' ideas benefits both the students and the teacher but because the teacher has to orchestrate the whole classroom further probing is not always possible. For example, it is sometimes better to guide many students quickly than to only guide one student after further probing. Indeed, Mortimer and Scott (2003) argue that there is place for both authoritative and dialogic teacher talk.

The use of open-ended questions during interactions on-the-fly may be perceived by teachers as more time consuming, as students' answers may require follow up questions that cannot be predicted before-hand. It is likely that due to curriculum, time-pressures, and classroom sizes that teachers feel compelled to move the learning forward at a quick pace, and this is more likely to be characterized by use of close-ended questions and comments as well as instructions.

Future research

In this paper we have described three coding schemes for studying on-the-fly interaction. For this we have used only a small part of our data. We will continue our analysis to find how well the coding schemes fit with a larger body of data from primary, lower and upper secondary physics and mathematics classes. Naturally, the coding schemes will be specified or extended if needed. Further, implementing codes with larger data will give better understanding about formative assessment across different classroom contexts, phases of lessons, educational levels, teachers and subjects.

REFERENCES

Anderson, R. (2002). Reforming Science Teaching: What Research says about Inquiry. *Journal of Science Teacher Education, 13*(1), 1–12.

Alexander, R. (2006). *Towards dialogic teaching: Rethinking classroom talk*. York: Dialogos.

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability* (formerly: Journal of Personnel Evaluation in Education), 21(1), 5–31.

Chin, C. (2007). Teacher questioning in science classrooms: Approaches that stimulate productive thinking. *Journal of Research in Science Teaching, 44*(6), 815–843.

Correia, C. F., Nieminen, P., Serret, N., Hähkiöniemi, M., Viiri, J., & Harrison, C. (in press). Informal formative assessment in inquiry-based science lessons. Proceedings of European Science Education Research Association Conference 2015.

Correia, C. F., Serret, N., Black, P., & Harrison, C. (2016). Manuscript in preparation.

Harlen, W. (2013). *Assessment & inquiry-based science education: issues in policy and practice*. Global Network of Science Academies.

Harrison, C. (2015). Assessment for learning in science classrooms. Manuscript in press.

Lehesvuori, S., Viiri, J., Rasku-Puttonen, H., Moate, J., & Helaakoski, J. (2013). Visualizing communication structures in science classrooms: tracing cumulativity in teacher-led whole class discussions. *Journal of Research in Science Education, 50(*8), 912–939.

Mercer, N., & Dawes, L. (2014). The study of talk between teachers and students, from the 1970s until the 2010s. *Oxford Review of Education*, *40*(4), 430-445.

Minner, D. D., Levy, A. J., & Century, J. (2010). Inquiry‐based science instruction—what is it and does it matter? Results from a research synthesis years 1984 to 2002. *Journal of research in science teaching*, *47*(4), 474-496.

Nieminen, P., Hähkiöniemi, M., Leskinen, J., & Viiri, J. (2016). Four kinds of formative assessment discussions in inquiry-based physics and mathematics teaching. Manuscript submitted for publication.

Scientix website: http://www.scientix.eu/web/guest/home - retrieved December 9th, 2015.

Shavelson, R.L., Young, D.B., Ayala, C.C, Brandon, P.R., Furtak, E.M., Ruiz-Primo, M.A, Tomita, M.K., Yin, Y. (2008). On the impact of curriculum-embedded formative assessment on learning: a collaboration between curriculum and assessment developers. *Applied Measurement in Education, 21*, 295–314.

Rocard, M., Csermely, P., Jorde, D., Lenzen, D., Walberg-Henriksson, H., & Hemmo, V. (2007). *Science education now: a new pedagogy for the future of Europe*. European Commission.

Ruiz-Primo, M.A., Furtak, E.M (2007). Exploring teachers' informal formative assessment practices and students' understanding in the context of scientific inquiry. *Journal of Research in Science Teaching, 44*(1), 57–84.

Torrance, H., & Pryor, J. (2001). Developing formative assessment in the classroom: Using action research to explore and modify theory. *British Educational Research Journal, 27*(5), 615–631.

Vygotsky, L. S. (1978). *Mind in society: The development of higher mental process*. Cambridge, Mass.: Harvard University Press.