International Journal of Science Education

ISSN: 0950-0693 (Print) 1464-5289 (Online) Journal homepage: http://www.tandfonline.com/loi/tsed20

Affordances and Constraints of Using the Socio-Political Debate for Authentic Summative Assessment

Jens Anker-Hansen & Maria Andrée

To cite this article: Jens Anker-Hansen & Maria Andrée (2015) Affordances and Constraints of Using the Socio-Political Debate for Authentic Summative Assessment, International Journal of Science Education, 37:15, 2577-2596, DOI: 10.1080/09500693.2015.1087068

To link to this article: http://dx.doi.org/10.1080/09500693.2015.1087068

Published online: 21 Sep 2015.

International Journal of Science Education, 2015
Vol. 37, No. 15, 2577–2596, http://dx.doi.org/10.1080/09500693.2015.1087068

Affordances and Constraints of Using the Socio-Political Debate for Authentic Summative Assessment

Jens Anker-Hansen* and Maria Andrée
Department of Mathematics and Science Education, Stockholm University, Stockholm, Sweden

This article reports from an empirical study on the affordances and constraints for using staged socio-political debates for authentic summative assessment of scientific literacy. The article focuses on conditions for student participation and what purposes emerge in student interaction in a socio-political debate. As part of the research project, a socio-political debate was designed for assessing student competences of scientific literacy in classroom practices. The debate centred on a fictive case about a lake where a decline in the yield of fish had been established. The students were assigned the task of participating in the debate from appointed roles as different stakeholders. Data were collected with video recordings of the enacted student debates. Student participation was analysed with the theoretical framework of communities of practice. The results show that multiple conflicting purposes of the socio-political debate as an assessment task emerged. The emergent purposes were (1) putting scientific knowledge on display versus staying true to one's role, (2) putting scientific knowledge on display versus expressing social responsibility, (3) putting scientific knowledge on display versus winning the debate, and (4) using sources tactically versus using sources critically. As these purposes emerged in classroom practice, tensions between different ways of enacting participation in the debates became manifest. Based on these findings, this article discusses the affordances and constraints for using a socio-political debate for classroom-based assessment of scientific literacy and argumentation in terms of validity, reliability and affordability.

Keywords: *authentic assessment; argumentation; communities of practice; scientific literacy; citizenship; validity; reliability*

Introduction

Science classroom-based assessments often emphasise propositional knowledge, rather than procedural knowledge, and one critique against this is that the assessments

*Corresponding author. Email: jens.anker-hansen@skolverket.se

become poor signifiers of how students can engage with science outside school (Aikenhead, Orpwood, & Fernsham, 2011). In addition, a specific use of science is often given beforehand – in contrast to what is required of participation in professional and civic practices outside school, where people are held more accountable for how they use science (Aikenhead et al., 2011) as well as for what science they use (Andrée, 2005). An alternative approach for classroom assessment of scientific literacy is to create assessment tasks that are similar to situations citizens are likely to encounter, and then observe how students act. If one considers knowledge as situated in social contexts, then knowledge displayed in school environments might not necessarily reflect how students relate to science in everyday life (Andrée, 2005; Jakobsson, Mäkitalo, & Säljö, 2009; Serder & Jakobsson, 2014). This perspective of scientific literacy requires assessment to acknowledge a multitude of civic practices and not only the traditions of school science (Linder et al., 2011).

Assessment tasks that aim to simulate 'real-life' situations are often referred to as performance and/or authentic assessment (Jönsson, 2011; Wiggins, 1989). Both concepts share the aspects of assessing performance directly in open-ended tasks with no correct answers given. In addition, authentic assessment tasks are designed to use criteria that are intended to represent what is valued in out-of-school contexts. Designing assessment that reflects out-of-school endeavours involves creating assessment tasks that comprise collaboration and negotiation between students on how science is made relevant for an everyday context (Sadler & Zeidler, 2009). Such aims can be found in assessment practices where students engage in societal topics concerned with a scientific content, so-called socioscientific issues (SSI) (Sadler & Zeidler, 2009), and when students are asked to construct arguments from scientific and non-scientific evidence (Erduran, Simon, & Osborne, 2004). This article aims to contribute to the discussion on affordances and constraints for assessing student competences of scientific literacy in classroom practices.

Scientific Literacy as Citizenship Education in Sweden

This study was conducted in a Swedish educational context. In the latest Swedish syllabi for science subjects (biology, chemistry and physics), aims for citizenship education are primarily formulated as developing students' ability to participate in practices of source critique, argumentation and decision-making (The Swedish National Agency of Education, 2011a). In an official commentary on the syllabi, the Swedish National Agency of Education points to new demands on teachers to evaluate students' knowledge in action (The Swedish National Agency of Education, 2011b). A survey on Swedish science teachers' experiences with the previous national science syllabi reported that they expressed insecurity and inexperience in dealing with the assessment of science as a human and social endeavour (Skolverket, 2010). This uncertainty appears to linger after the most recent reforms, where teachers find the external support, in the form of standardised national exams, to be insufficient for assessing the different aims for citizenship education (Lundqvist & Lidar, 2013).

The Socio-Political Debate as an Assessment Task for Scientific Literacy

A practice where source critique, argumentation and decision-making are central is a socio-political debate. As citizens of a democratic society, individuals are frequently exposed to socio-political debates to which they can relate in some way or another. They might not engage in debates on a regular basis, but developing an understanding of the processes of argumentation could involve participation in debates (Duschl, 2007). Previous research on how students can learn to discuss scientific issues is extensive (Driver, Newton, & Osborne, 2000; Jiménez-Aleixandre, Rodríguez, & Duschl, 2000; Newton, Driver, & Osborne, 1999). However, the transition from providing the students with tools for argumentation and developing assessments of students' discussions is not simple. In an overview of research on argumentation in science education, Jiménez-Aleixandre and Erduran (2007) show that assessment of student argumentation has primarily focussed on students' abilities to adopt the given argumentation patterns. This focus has been criticised for neglecting the dialectical aspect of how students adapt their arguments to other students' utterances (Nielsen, 2013). Thus, there is a need for research focussing on conditions for student participation in the assessment of argumentation.

One way of dealing with socio-political debates in school is through role-play or arranged discussions, where students are appointed roles or standpoints from which to argue. A socio-political debate, although staged in school practice, may offer greater possibilities for increased validity of assessment than, for example, a written test. Åberg, Mäkitalo, and Säljö (2010) found that classroom debates functioned as a creative educational setting where students had to respond to the complexities of producing knowledge relevant to a situation. Their study showed that students were familiarised with how to construct arguments by drawing on a multitude of potential relevant sources of knowledge. Thus, by means of a socio-political debate, schools can offer students an opportunity to engage in tasks that have some resemblance with participation in democratic decision-making processes. Though debates afford participation authentic to democratic decision-making processes, constraints are also involved with role-play. Previous research points to role-play as a mediator to broadening perspectives in science education; it shows that the participation of students is very likely to alter, as a discussion turns in favour of a certain viewpoint (Kolstø, 2000; Ødegaard, 2003; Simonneaux, 2001). This means not only that a student might experience problems displaying knowledge on a topic if other participants act in opposition to the role of the allotted stakeholder, but also that participation in the debate will influence students' opinions. Such constraints have consequences for students' opportunities to put knowledge on display for teacher assessment. There is thus a need to investigate the affordances and constraints for fair student participation in socio-political debates when debates are used as assessment tools.

The idea that teachers are trading off fair student participation for tasks that are believed to better represent out-of-school practices is not new to the field of authentic assessment (Champagne & Newell, 1992). One example concerns activities outside school having multiple tacit purposes that the teachers risk not considering when

assessing student participation in those activities inside the classroom (Sambell, McDowell, & Brown, 1997; Terwilliger, 1997). Consequently, when participating in authentic tasks, the students have to find ways of addressing the purposes emerging from the authentic context as well as the purposes stated in the assessment criteria. Another example is assessment tasks where students' collaboration may be constrained by social processes pertaining to group work. Consequently, what one gains in validity one risks losing in reliability, since students will not be offered equal conditions for participation (Ratcliffe & Grace, 2003). On the other hand, Moss (2003) argues that a trade-off between validity and reliability only exists in terms of standardisation where tasks are constructed for atomistic measurements and where there is a given norm for ideal performance. Similarly, Chang and Chiu (2005) argued that validity and conditions for student participation in authentic assessments cannot be evaluated through comparisons to general principles of standardisation, since student participation is expected to be versatile. They argue that studying how students deal with different task formats would be much more informative. In sum, despite extensive research on argumentation in science education, there is still a need for research concerning the specific affordances and constraints involved with authentic assessment, particularly from the perspective of student participation.

Redefining Validity, Reliability and Affordability in Authentic Assessment Design

The aim of authentic assessment is to give a wide array of information of student participation in out-of-school practices rather than standardising teachers' assessment (Wiggins, 1989). Authentic assessment is thus designed from different principles than 'traditional' assessments (Wiggins, 1989). Bell (2007) summarises authentic assessment in science education as (1) having explicit and public criteria, (2) involving collaboration, (3) being contextualised, (4) representing realistic and fair practices, (5) using complex and multifaceted scoring, (6) identifying strengths, (7) having multiple purposes, (8) enabling integration between knowledge developed from different sources, (9) enabling evidence to be added or removed, and (10) encouraging metacognition and reflection. Compliance with these requirements necessitates taking a step away from atomistic test items where a specific skill or content is evaluated, towards a more holistic classroom assessment where teachers design assessment for the needs of their students and their educational goals (Ratcliffe & Grace, 2003). Steps towards authentic assessment thus involve modifying definitions of validity, reliability and affordability, as used in standardised educational measurements (Moss, 2007).

Traditional validity of a test item can be summarised in four aspects (Waugh & Gronlund, 2013): (1) content-related (how a sample of tasks represents the domain intended to be measured), (2) criterion-related (to what extent student performances can be used for prediction on future performance or estimation on present performance on other valued measures), (3) construct-related (to what extent performances describe constructed abilities such as argumentation skills) and (4) the consequences

of having experienced the assessment tasks (for example, willingness to engage in similar issues). However, in authentic assessment, such as described by Bell (2007), where there is not one set of correct answers to relate to when assessing performance, these traditional aspects of validity become complex (Moss, 2007). When citizens can use science in so many ways, it is difficult to evaluate if students' performances in assessment tasks represent how they actually act in out-of-school contexts. It is, for instance, shown that the attempts to construct tasks that reflect everyday contexts in the Programme for International Student Assessment (PISA) did not always allow Swedish students to relate to their various experiences of everyday practices (Serder & Jakobsson, 2014). Consequently, authentic assessment needs to be validated through what consequences it has for the students (Linn, Baker, & Dunbar, 1991).

Reliability of performance assessments is typically evaluated by the use of different judges or different teachers using the same scoring rubrics (Waugh & Gronlund, 2013). In light of Bell's (2007) summary of authentic assessment, students' performances may display many aspects of scientific literacy in a task. However, to assess many different aspects of scientific literacy simultaneously *in situ* may be difficult for a teacher. Different teachers may pay attention to different aspects when assessing students' performances and thereby produce different accounts of the quality of performance (Gipps, 1995). Teacher inference is therefore a central aspect of reliability. Another central aspect of reliability, upon which this article will focus, concerns what opportunities are afforded to students for putting knowledge on display when tasks are made more authentic. For example, a student may have many ideas about how to deal with a particular socio-scientific issue but be unable to share those ideas due to interactional constraints (Kolstø & Ratcliffe, 2007). Such constraints could be turn-taking in the discussion, the composition of the discussion group, or the characteristics of the conversation, for example, the ability to use the specific language of the conversation.

In sum, striving towards greater validity by means of authenticity and performance assessment offers challenges for achieving reliability. Thus in authentic assessment, the concept of a 'fair' assessment task is very complex, and teachers have to negotiate 'fairness' in relation to the assessment task. One direction is to follow the advice of Ratcliffe and Grace (2003) to drop requirements of high reliability in single assessment tasks for the benefit of achieving high validity. Instead, reliability might be attained through the use of multiple assessment tasks to create a better holistic picture of student competences.

In the Swedish standards aimed to guide teachers, the progression of quality in student argumentation is described in rather unspecified terms, ranging from 'simple' to 'well developed in width and depth'. Consequently, teachers in Sweden are faced with the challenge of developing unspecified standards into models that are concrete and functional enough to be used for continuous formative and summative teacher assessment. Yet, the models that teachers use would preferably also reflect what is actually valued in civic life and afford student participation that is meaningful for future engagement in societal endeavours – not simply what is easy to measure. This implicates a necessity for teachers to explore the construction of different kinds of tasks and criteria to find a desired balance between validity and reliability.

Participation in a Socio-Political Debate as Participation in a Community of Practice

We have argued for making assessment more authentic so that the student participation may reflect how students make meaning and coherence of science and give it purposes in out-of-school practices (Brown, Collins, & Duguid, 1989). Consequently, we focus on the aspects of assessment that concern conditions for authentic participation in the assessment task and study what purposes emerge from student interaction. In this study we draw on the work of Etienne Wenger (1998) to analyse student participation in the socio-political debates as a staged community of practice (CoP). From a CoP perspective, student participation in socio-political debates needs to be understood as a matter of engagement in a social practice involving issues of identity and belonging. In this social practice, particular experiences are negotiated and made relevant through the course of interaction. Previously, concepts of CoP have been used in science education research to describe how the introduction of new ways of acting changes students' repertoires, as a new practice becomes more meaningful to them (Andréee & Hansson, 2013; Evnitskaya & Morton, 2011; Kisiel, 2010; Munby, Taylor, Chin, & Hutchinson, 2007; Willis, 2011). With the concepts of Wenger, the work of a teacher can be described as border-brokering, which is when a teacher introduces new tools and repertoires from other communities into the classroom. In the classroom, those repertoires are negotiated into something meaningful to the classroom community. Thus, from a CoP perspective, experiences and repertoires are understood to be under negotiation by students. In the socio-political debates, students have to negotiate authenticity, and in this process, new purposes emerge.

In a socio-political debate, students negotiate meaning of personal engagement in the staged assessment task. The purposes of meaningful participation from the perspective of one stakeholder may conflict with the purposes of the assessment. For a student, it could be beneficial for the sake of the debate to avoid discussing facts that are unfavourable to the character's role, although discussing those facts could be an opportunity to put scientific knowledge on display for assessment. Consequently, students' affordances and constraints for participation in a debate are conditioned by what purposes emerge in the staged assessment. Here, we argue that these emerging (conflicting) purposes need to be observed and analysed rather than removed from the assessment. Therefore, this study investigates what conflicting purposes may emerge in a socio-political debate staged in a science classroom practice.

In this article we will specifically focus on the conflicting purposes that emerge in a debate in science class. We will further study how students negotiate participating authentically to the context of the staged debate while demonstrating their knowledge of science. This is accomplished with an empirical analysis of how students participate in a debate designed to provide the science teachers with evidence for summative assessment purposes. We then discuss the analysis of students' participation in the debates in terms of validity, reliability and affordability for assessing scientific literacy in action. The research question for the study is:

What affordances and constraints for student participation emerge when using a socio-political debate for authentic assessment of scientific literacy?

Methods

In light of the study's aim, we designed a study to collect data from students' discussions in socio-political debates. The study was set up as an intervention in collaboration with a practising science teacher in a lower secondary school. A socio-political debate was designed to allow for an investigation of socio-political debates as tools for authentic summative assessment in real science classroom practice. The intervention was designed to enable a better understanding of the processes of participation and thereby contribute to the development of classroom-based assessment of student's readiness to negotiate and resolve societal issues relating to science.

The Participants

The participating science teacher had little more than five years of experience teaching science in a lower secondary school. In the design of the assessment, she contributed with experience and knowledge of her particular teaching practice and the participating students. She also formally conducted the assessment and was held accountable for that process. The teacher used the socio-political debate as a complementary task to a written test for the purpose of summative assessment. However, her assessment of the debate will not be analysed in this article.

A total of 41 students in two science classes in school year 8 (approximately 14 years old) participated in the study. The assessment was incorporated as part of regular teaching practice, but participation in the research study was voluntary. Thus, the students were informed that they could withdraw their consent to participate in the research study at any point and be given alternative opportunities for assessment from the teacher. However, all the students and their parents did formally consent to participation.

Design of the Intervention

For the purpose of developing an authentic case, we decided to use scientific material from a real context (a lake) that presented a real issue to real people. The data chosen for the design of the case were provided from a County Administrative Board and included information about the lake concerning aspects such as acidity and toxicity. From the data, a case was developed about a decline in the yield of fish. As authentic scientific material is rarely written for an audience of school students, some simplifications of the original material were made. The simplifications were carefully made to ensure the material reflected data sources to which citizens may actually have access.

A teaching sequence of two preparatory lessons followed by a debate was designed. Since there were two classes, we observed six lessons in total. The decision of dividing

the class into pairs or groups of three and assigning them roles as PVC factory representatives (CEO and engineer), farmers, fishermen, politicians or representatives from a residential green organisation was left to the teacher. The roles were selected by the teacher and the first author from possible connections to the measured data about the lake, such as farmers possibly being responsible for eutrophication due to their use of fertilisers. Three reports describing the condition of the lake and the production of PVC were made available to the students as data material. The first report was presented as produced by the factory, the second as produced by a consultant commissioned by the green organisation, and the third as an annual report issued by the municipal authorities. None of the reports made any explicit statements concerning who was to blame for the condition of the fish in the lake. However, data about the lake's condition implicated the involvement of all parties.

The assessment criteria were taken from the chemistry curriculum standards concerning argumentation (see Table 1) (The Swedish National Agency of Education, 2011a) in combination with a model for expanding arguments (see Table 2) provided by the researchers. The proposed model for the expansion of arguments was based on the Toulmin (1958/2003) argumentation patterns (TAP) (Erduran et al., 2004). Thus, the teacher and her students used the TAP-based model when working with quality of arguments. However, the TAP-based model was not used for the analysis of this study, since our aim was to investigate emergent affordances and constraints rather than how well the students could follow a given argumentation pattern.

Table 1. Standards for communication in the Swedish national syllabi for chemistry

Minimum acceptable mark (E)	Highest mark (A)
Pupils can talk about and discuss questions concerning energy, technology, the environment and society, and differentiate facts from values, and formulate their views with **simple** reasoning, and also describe some possible consequences. In such situations, pupils can put questions and put forward and respond to views and arguments in a way which **to some extent takes the discussions forward**. Pupils can search for information on the natural sciences and use different sources and apply **simple and to some extent** informed reasoning to the credibility and relevance of their sources and information. Pupils can use information in a **basically functional** way in discussions and create **simple** texts and other communications with **some** adaptation to purpose and target group.	Pupils can talk about and discuss questions concerning energy, technology, the environment and society, and differentiate facts from values and formulate their views with **well-developed** reasoning, and also describe some possible consequences. In their discussions, pupils put questions and put forward views and respond to views and arguments in a way which **carries the discussions forward and deepens or broadens them**. Pupils can search for information about the natural sciences and use different sources and apply **well-developed and well-informed** reasoning about the credibility and relevance of their sources and information. Pupils can use the information in a **well-functioning** way in discussions and create **well-developed** texts and other communications with **good** adaptation to purpose and target group.

Table 2. Model of expanding arguments

Lower level ----------→Higher level		
Arguments consist of claims backed by data	Arguments consist of claims backed by data and explanations to why mentioned data support the claim (warrants)	Arguments consist of claims, data, explanations (warrants) and justifications to under what conditions the statements could be held to be true (qualifier) or not (rebuttal)

Introducing the Intervention

The teacher presented the task and the reports and informed the students that she was going to use the presented model (see Table 1) as criteria for the summative assessment of the debate. Furthermore, she informed the students that she would evaluate their knowledge about (a) water as a solvent (including acidity), (b) the chemical processes involved in recycling as well as (c) source criticism. In the remaining preparatory lessons (lasting 1.5 hours each), the students had to formulate conclusions about responsibility by connecting charts and various data. The students worked with the reports and also searched for information on the Internet.

Conducting the Debate

Finally, debates about the condition of the lake and the stakeholders' responsibility were staged in the two classes. In these debates, the teacher acted as chairperson, distributing the speech and making notes for assessment purposes during the debate. She neither provided explicit feedback nor shared her assessment during the debates.

Data Collection

The staged socio-political debate was video-recorded. In total, four debates in groups of 9–11 participants (see Table 3) were studied. The students participating in the debate sat around a triangular table, with two opposing cameras recording the debate. The first author participated as an observer. The recordings of the socio-political debate were then transcribed verbatim.

Data analysis

We aimed to identify systematic categories of participation that are locally defined through emergent advantages and disadvantages in a community's endeavours, instead of regarding conventions of engagement in human interaction as predetermined by stated rules and policies (Tatli & Özbilgin, 2012; Zhu & Bargiela-Chiappini, 2013). The following analytical questions guided the analysis:

Table 3. Overview of the debates

Debate	Participants		Duration
1	Class A	11 students	27 min
2	Class A	10 students	28 min
3	Class B	11 students	27 min
4	Class B	9 students	25 min

- What conflicting purposes emerge in the socio-political debate?
- How do students negotiate how to address the conflicting emergent purposes?

The analysis was a detailed content-oriented analysis, taking into account the dialogicity at the level of utterances (Linell, 1998), in other words, how the students used and built upon each other's utterances. Conflicting emergent purposes were found in episodes where students changed (negotiated) their argumentation (purpose) after facing opposition (conflict). The data coding was conducted collectively by both authors and verified with other researchers in seminars. The identified purposes and conflicts were compared to other identified conflicts of the same and other categories. Through continuous comparison, characteristics of the categories were redefined and restructured until the authors agreed on the robustness of the categories. Thus, we performed an empirically grounded open qualitative coding of what purposes emerged in the conversations in episodes of student interaction (Corbin & Strauss, 2008).

Results

In this section we present selected episodes from the debates that illustrate students' negotiation of the tensions that emerged between different purposes of the assessment task through their participation in the debates. The negotiation concerned how to address the purposes of:

(1) putting scientific knowledge on display versus staying true to one's role,
(2) putting scientific knowledge on display versus expressing social responsibility,
(3) putting scientific knowledge on display versus 'winning' the debate, and
(4) using sources tactically versus using sources critically.

In the following episodes we will illustrate how the students created means to afford participation that addressed the conflicting purposes of the assessment task when these tensions became apparent to them.

I. *Negotiating How to Address the Purposes of Putting Scientific Knowledge on Display versus Staying True to One's Role*

The following episode, taken from Debate 1, exemplifies how students struggled to provide the teacher with evidence for marking and staying in character of their allotted role while defending the interests of their stakeholder.

Episode 1:

1. Annika (Farmer):	We have seen that the PVC factory emits heavy metals. And we irrigate with water from the lake. And we also give water to our animals. And when we humans eat the meat of the animals or the grown plants, we accumulate high levels of heavy metals. It is scientifically proven that heavy metals are harmful in large amounts and are concentrated higher up in the food chain. Many heavy metals can also cause damage to the reproductive abilities and the nervous system. What do you have to say about that?
2. Stina (CEO):	What do we have to say about that? Well, the levels are low so far, so it is not that dangerous. We can see here that it increases. Zinc and nickel. No, what am I saying? Copper, cadmium and lead and arsenic. It is very hard to stop, but we are using lead for instance to colour things, but we could make things colourless. But we have to discuss that.
3. Karl (Politician):	But is that not … is lead not … illegal to use nowadays?
4. Stina (CEO):	No. Anyway, we are not using lead, but it is being used. And, by the way, you could say something as an engineer (pokes at the Engineer with her elbow)
5. Rasmus (Engineer):	What?
6. Class	(Laughter)
7. Sven (Green Representative):	So you deny using heavy metals?
8. Stina (CEO):	No. But not lead.
9. Sven (Green Representative):	But the levels have become higher since you got here: from 0.04 to 0.10. So how do you explain that? Are you not using lead?
10. Stina (CEO):	Does it have to come from us? Is there anything proving that it's from us? Does it say that it comes from us? It (the report) is about the whole lake. It could come from the whole area.

In this episode, Stina, the CEO, displayed conclusions about heavy metals and PVC manufacture (2), but this put her character's role in a difficult position (3 and 9). The concept of role-play impaired Stina's opportunities to expose the range of conclusions that she made; some conclusions proved disadvantageous to the PVC factory (3). Stina and Rasmus, the Engineer, did not fully recognise the factory's connections to lead publicly (4). They struggled to engage in ways that were mutually meaningful both from the perspective of the characters' roles of the CEO and Engineer and bright students having drawn many conclusions from the presented material (4 and 5).

An explicit aspect of the argumentation practice was the expansion of arguments with explanations and clarifications under which circumstances their conclusions were relevant. The opportunities to demonstrate knowledge concerning a certain issue were restricted by participants' willingness to remain engaged in the topic

discussed. As exemplified in this episode, students constantly referred to responsibility for the well-being of the lake (1) to afford opportunities to address the different purposes of the assessment task. Stina reified her conclusion that none of the reports supported clear connections between increased levels of lead and the PVC factory (10). This utterance did not contradict the introduced concerns for the well-being of the lake. Thus, an opportunity to both display her scientific skills of utilising evidence and defend her character's role against the accusations was negotiated.

II. *Negotiating How to Address the Purposes of Putting Scientific Knowledge on Display versus Expressing Social Responsibility*

The following episode, from Debate 4, is an example of a student discussion on how to prioritise between conflicting needs of animals and people in and around the lake. The episode shows how students negotiated the importance of the scientific aspect in resolving the issue.

Episode 2:

1. Maja (Fisherman):	It is a lot like … It is very hard to replace, or … you have to replace PVC anyway, because it is particularly that material that is dangerous. If you bury it, it will evidently seep into the underground water and then into the lake. And when it is incinerated, dioxins are released too. So it is almost impossible to get rid of the dioxins.
2. Tommy (Politician):	Well, how do you propose that they should change their production? There are people who are educated for those jobs, who suddenly are supposed to change professions.
3. Maja (Fisherman):	We have to think about the city, our environment and our animals, among others, the fish. Yes, we have to consider the future. We can't exterminate the fish to provide people with jobs. You can always educate yourself into something new. Or a shorter education, perhaps in something different. I consider this … life-threatening to our fish.
4. Bo (Green Representative):	What suggestions do you have for them decreasing? Because dioxins do not disappear, and it is just bad for the environment. Do you have a suggestion on how to stop this on your own, then? You know that this is not good, and in the end, if your children, or our children, for instance, swim in the lake, they will consume water, and it will be used as drinking water. When the water is deox … becomes less oxygen-deficit, it becomes acidic and cannot be drank. And it is also dangerous to us? So what are your suggestions?
5. Tommy (Politician):	We absolutely agree on the dioxin levels. I think the factory could improve that. But as they said, the dioxin levels are very low in the lake.

In the presented episode Tommy, the Politician, reified concerns regarding employment of the town population (12). Maja, the Fisherman, argued that her reified concerns for dioxin pollution had a higher significance (13). Priority regarding the livelihood of PVC factory workers would be relevant from the Politician's viewpoint, since fish cannot vote whereas factory employees can. This is another example similar to the previously mentioned conflict between putting scientific knowledge on

display versus staying true to one's role. However, a conflict also exists between putting scientific knowledge versus social responsibility on display. Maja's argumentation with scientific concepts provided the science teacher with more evidence for marking than Tommy's reference to social responsibility.

Bo, the Green Representative, used the introduction (11) and feedback (13) from Maja to negotiate a merger between social responsibility and natural science. Bo was able to engage in mutually valued ways when supporting the priority of the environment by referring to dioxins remaining in the water used for inhabitants swimming and drinking (14). This introduced aspect afforded Tommy the opportunity to agree that the scientific aspect concerning dioxins was indeed important (15). This utterance thus confirmed both the scientific aspect of the topic and Tommy's participation as a politician socially responsible for the community. Bo confused the scientific concepts, possibly because the Swedish words for *oxygen [syre]* and *acid [syra]* sound almost the same. This was discussed later in the debate, but was not recognised by Tommy in the excerpt above. In summary, the participants observed the interests of many different parties such as plants, animals and people. Thus, the debate entailed negotiations of how to discuss those interests in ways that were mutually meaningful for the students' engagement in the discussions, and valued contributions to the debate by the community. These negotiations provided students with opportunities for participation as both scientifically knowledgeable and socially responsible citizens.

III. *Negotiating How to Address the Purposes of Putting Scientific Knowledge on Display versus 'Winning' the Debate*

In the following episode, from Debate 4, we show an example of how students negotiated participation in the debate in terms of 'winning' the debate rather than producing warranted arguments.[1]

Episode 3:

1. Teacher:	First the Green Representatives, then the politicians
2. Lars (Green Representative):	You don't emit very dangerous substances at the moment. But dioxins are still appearing and are increasing with every year. We can see that in our rapport, simply enough, which we ordered from the Water Agency.
3. Bo (Green Representative):	Krister Helvén
4. Lars (Green Representative):	Krister Helvén … And it shows that the pH value is much higher by your water outlet. In your own … I want to bring to your attention, your own rapport shows that, even if we have, that is what is allowed, the total stress, is low at the moment, the stress is increasing all the time, right?
5. Anders (Politician):	Well, what you said about pH is incorrect because=
6. Lars (Green Representative):	=The water outlet is higher than it was ten years ago!
7. Anders (Politician):	No. It is not=

[1]In the transcript, we have used the '=' sign as applied in Conversation Analysis (Schegloff, 2007), to describe interruptions.

8. Lars (Green Representative):	=Yes. It is higher by the water outlet of Formplastic Ltd. than in the middle of the lake and in the residential area. Is it not?
9. Anders (Politician):	Well=
10. Bo and Lars (Green Representatives):	=(Give each other a high-five)
11. Bo (Green Representative):	You nailed it there!
12. Group	(Laughter)
13. Teacher	Maja
14. Maja (Fisherman)	I want to add that the lake is acidic. Because the pH is below seven
15. Lotta (CEO):	How can the acid pH value be the fault of the factory, if the emission … if our outlet has a value of eight point five or eight point three, which means alkaline? Thus, it is not our fault.

The teacher gave the speaking turn to the Green Representatives and stated that the Politicians could follow (16). Even though the teacher specified 'the Green Representatives *then* the Politicians' (emphasis added), by assigning the speaking order to the Green Representatives and the Politicians in the same sentence, she opened up the possibility to interpret that the speaking turn was given to both parties as a shared possession. The Politicians could speak when the Greens were finished without teacher interference.

The claim of Lars, the Green Representative, was not warranted with an explanation. Instead, Lars and Bo strengthened the statements by referring to the author of the report (17, 18 and 19). Furthermore, Lars's utterance could be interpreted as establishing a connection between dioxins and acidity and holding the factory responsible for an alleged alkaline condition of the lake (17 and 19). In contrast, the provided scientific reports showed that the lake was slowly becoming more acidic. Anders, the Politician, objected Lars's statements (20). However, Anders was interrupted and not allowed to explain why he considered the utterances to be incorrect. Lars and Bo defended their claims by stating that they had read the data correctly in the report (21 and 23). They did this aggressively by interrupting the politician (21 and 23), sealing their victory with a 'high five' and declaring that they had 'nailed it' (25 and 26). Lars and Bo's mistake initially gave Anders an exigent opportunity to put scientific knowledge on display by rebutting a prior statement. However, this opportunity was lost when Lars and Bo interrupted. The declaration of victory (25 and 26) was followed by collective student laughter (27), and Anders did not express any further objections. Thus, Lars and Bo succeeded in discouraging the Politician from expanding his argumentation. Maja, the Fisherman, explicitly pointed out what the reports actually showed (29). Together with Lotta, the CEO, she helped clarify the confusion (30).

IV. *Negotiating How to Address the Purposes of Using Sources Tactically versus Using Sources Critically*

The following excerpt is from Debate 1 and illustrates how the students negotiated what sources to use and how sources could be used in the debate.

Episode 4:

31. Stina (CEO):	I want to know a little more about this source. Whether there are children who get cancer.
32. Emil (Fisherman):	It also says in your report that it is carcinogenic.
33. Stina (CEO):	In our report? No. Is it report one or two? Because ours is number one.
34. Emil (Fisherman):	Two.
35. Stina (CEO):	Because ours is number one. That is, like your report. This doesn't really feel like so 'believe it' (in English). I don't know whose. It is the Green organisation's. So, it is not our report. So we haven't written that.
36. Emil (Fisherman):	So, you haven't used all the facts in your report? So, you deny things?

In this episode Emil, the Fisherman, suggests a connection between cancer and PVC. This introduction of data and conclusions functioned tactically for both furthering the interests of the character's role and putting knowledge on display for teacher assessment. The inherent conflict between the interests of the different characters' roles, in this case the CEO and the Fisherman, contributed to making questioning of the other parties' use of sources a part of the discussion. By referring to information as coming from the PVC factory, Emil created a connection between PVC and cancer stated by the factory (32). This could have been a tactical move, had he not mixed up the report from the factory with the report from the Green Organisation (34). When Stina, the CEO, discovered that data came from another stakeholder, she questioned its trustworthiness (35). Stina's rebuttal of Emil's use of information functioned both as a defence in the debate and a display of source-critical capabilities. On the other hand, with the rebuttal Stina framed the report from the Green Organisation as an unreliable source and thus limited what data to use. Emil countered with an accusation of the PVC factory's report being incomplete due to the denial of facts (36). This negotiation showed how sources could be made into less usable resources. Consequently, the participants were held accountable for their enterprise of source selection.

Discussion and Conclusions

In the socio-political debates, the students' negotiation of meaning regarding the conditions of the lake and consideration of different stakeholders became visible to both the teacher and fellow students (Wenger, 1998). In this sense, the debate offered valid evidence of students' emergent scientific literacy. However, tensions arose in the debate between the purposes of authenticity and the summative assessment task. These tensions in turn restricted the teacher's possibility to assess individual students, since the students' negotiation of purposes conditioned their possibilities to engage in the debate.

Validity as Accuracy of Assessing Students Meaning-Making about the Lake

The most apparent affordance for using the socio-political debate as an assessment format was that the students' participation provided information about how they could negotiate the relevance and functionality of the scientific reports for determining responsibility for the condition of the lake. In the debates, it became apparent which

students needed support on how to interpret the scientific facts and concepts from the reports to be able to use them correctly and productively in the discussions. This would constitute important information for classroom-based assessment on students' abilities to engage in socio-scientific argumentation. In line with prior research of argumentation in science education, the debates afforded special conditions for the instant responses from other students that are difficult to offer in written argumentation examinations (Säljö, Mäkitalo, & Jakobsson, 2011). The quality of students' arguments became clear as students' discussions made apparent what scientific conclusions became advantageous and disadvantageous to the different stakeholders. Consequently, socio-political debates offer opportunities for assessing the quality of students' argument in action and how students respond to one another and the arguments proposed in the debate.

From a 'traditional' definition of validity (Waugh & Gronlund, 2013), the teacher could not ensure that all the domains she wanted to assess – (a) water as a solvent (including acidity), (b) chemical processes involved in recycling, as well as (c) source criticism – were covered in depth in the discussions. The presented transcripts illustrate how the different debate groups focussed on different aspects of the chemistry content. Thus, they provided the teacher with different evidence in relation to the assessment domains.

The practices of engaging in socio-political debates proved to be complex; they involved purposes of expressing social responsibility and 'winning the debate', which were not stated as issues the teacher would assess. Thus, in light of the purpose of the teacher's summative assessment, these emerging purposes, although negotiated by the students, potentially interfered with the predefined criteria of the assessment task. Consequently, students who had prepared arguments from the predefined criteria might come up short in the debate due to being unable to follow the negotiation and adjust their arguments to the discussion.

If one evaluates validity in terms of what consequences the assessment has for future engagement in argumentative practices (Linn et al., 1991; Waugh & Gronlund, 2013), the teacher cannot assess how each individual student will actually engage in future discussions with other societal issues. However, the socio-political debate offers a complex practice for participation; the students are not only held accountable as students, but also in their roles as both stakeholders and socially responsible citizens (Åberg et al., 2010). The emergent purposes of 'staying true to one's role', 'expressing social responsibility' and 'winning the debate' indicate that the students authenticated their engagement to represent the interests of stakeholders such as fishermen and farmers beyond what was required from the criteria.

Reliability as 'Fair' Conditions for Participation

As expected, the results showed that sustaining reliability was constricted as the participating students were afforded different opportunities to put scientific knowledge on display. This finding is in line with previous research, pointing to the importance of taking into account the social dynamics of role-play and how students' opinions are

influenced by the debate setting (Kolstø, 2000; Simonneaux, 2001; Ødegaard, 2003). This was observed when Tommy, the Politician, abandoned his argumentation about community employment in favour of a discussion about dioxin levels. A salient purpose of a debate is to persuade other participants. Thus, there is always a risk that stakeholders might resign from the discussion if their arguments lose influence in the debate.

Possible Ways forward

Teachers in Hong Kong and Singapore have used alternative assessments like performance assessment and project work (Cheng, 2006; Koh & Luke, 2009). Project work often produces written products for the teacher and performance assessment of, for example, laboratory work, which can be marked individually by separating students. Consequently, alternative assessment formats may be easier to administer for individual assessment. So, a major question is if the affordances of assessing students in interactions outweigh the constraints that are involved with activities where students' affordability is very dependent on the actions of other students. One expressed purpose for using alternative assessments to standardised tests is to collect evidence of how well communication of conclusions, source critique and construction of arguments is negotiated in groups (Sadler & Zeidler, 2009). If students are separated in performance assessment, it would reduce these processes. Group project work reports will provide the result of a negotiation. However, to assess the process, the teacher has to observe how students work and discuss (Boud, Cohen, & Sampson, 1999). The individual marking of such group discussions will require the same teacher attention as the assessment of debates (with the exception of project work conducted as collaboration rather than competition). Using peer assessment to complement what the teacher can observe has been suggested in the assessment of collaborative work (Cheng & Warren, 2000). However, the unfairness of group dynamics cannot be completely avoided if the negotiation process is to be included in the assessment.

A possible critique of how the teacher directed the debate might be that she should have ensured that all students were given equal opportunities (e.g. in time slots) to develop their standpoints and interpretations of the scientific reports, and that she should have ensured that students stayed on the scientific topic rather than focusing on 'winning the debate'. However, if the socio-political debates are to become authentic, students need to be given the creative freedom and accountability to decide on the focus of the discussion (Åberg et al., 2010). As Jönsson (2011) explains, making assessment authentic is not about investigating how facts and theories learned in one situation are applied in another situation. Rather, it involves how students deal with situations they might face in society. Students are likely to find themselves in situations where they somehow represent the interest of a stakeholder, where they have to challenge rhetorical strategies or evaluate the importance of science in a context and the usefulness of sources. Consequently, for future creation of assessment criteria in debates, students should be informed that they are going to be assessed on

how they mutually engage, negotiate a repertoire and deal with their accountability for these challenges. By studying conflicts between emergent purposes of the debate, students and teachers can better prepare for such conflicts.

Finally, to increase 'fairness' in the assessment, our conclusion concurs with previously formulated conclusions on authentic assessment, stating that students should be allowed to work with a variety of contexts and engage in a multitude of debates with diverse stakeholders (Ratcliffe & Grace, 2003). This study underscores the importance of varying contexts and roles and shows how the different characters afford different kinds of participation. Engaging more than one assessing teacher would also ensure that more evidence was gathered from all the students' participation in the debate.

The results of this study raise new questions regarding the challenges for teachers to infer from and score students' participation in socio-political debates. For example, what information can and do teachers actually use from socio-political debates to assess students' argumentation as scientific literacy? How do teachers value the qualities of argumentation when this is vaguely described in the national standards? Is there a tacit professional norm upon which teachers can draw? These questions concerning teacher 'judgement' (Cooksey, Freebody, & Wyatt-Smith, 2007), would be valuable to investigate to further advance one's understanding of affordances and constraints for using various forms of authentic assessment tools for classroom-based teacher assessment.

References

Åberg, M., Mäkitalo, Å., & Säljö, R. (2010). Knowing and arguing in a panel debate: Speaker roles and responsibilities to others. In K. Littelton & C. Howe (Eds.), *Educational dialogues. Understanding and promoting productive interaction* (pp. 13–30). London: Routledge.

Aikenhead, G., Orpwood, G., & Fernsham, P. (2011). Scientific literacy for a knowledge society. In C. Linder, L. Östman, D. A. Roberts, P. O. Wickman, G. Erickson, & A. MacKinnon (Eds.), *Exploring the landscape of scientific literacy* (pp. 28–44). New York, NY: Routledge.

Andrée, M. (2005). Ways of using 'Everyday Life' in the science classroom. In K. Boersma, M. Goedhart, O. De Jong, & H. Eijkelhof (Eds.), *Research and the quality of science education* (pp. 107–116). Dordrecht: Springer.

Andrée, M., & Hansson, L. (2013). Marketing the 'Broad Line': Invitations to STEM education in a Swedish recruitment campaign. *International Journal of Science Education*, *35*(1), 147–166.

Bell, B. (2007). Classroom assessment of science learning. In S. Abell & N. Lederman (Eds.), *Handbook of research on science education* (pp. 965–1006). New York, NY: Routledge.

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher*, *18*(1), 32–42.

Boud, D., Cohen, R., & Sampson, J. (1999). Peer learning and assessment. *Assessment & Evaluation in Higher Education*, *24*(4), 413–426.

Champagne, A. B., & Newell, S. T. (1992). Directions for research and development: Alternative methods of assessing scientific literacy. *Journal of Research in Science Teaching*, *29*(8), 841–860.

Chang, S. N., & Chiu, M. H. (2005). The development of authentic assessments to investigate ninth graders' scientific literacy: In the case of scientific cognition concerning the concepts of chemistry and physics. *International Journal of Science and Mathematics Education*, *3*(1), 117–140.

Cheng, M. H. (2006). Junior secondary science teachers' understanding and practice of alternative assessment in Hong Kong: Implications for teacher professional development. *Canadian Journal of Math, Science & Technology Education*, *6*(3), 227–243.

Cheng, W., & Warren, M. (2000). Making a difference: Using peers to assess individual students' contributions to a group project. *Teaching in Higher Education*, *5*(2), 243–255.

Cooksey, R. W., Freebody, P., & Wyatt-Smith, C. (2007). Assessment as judgment-in-context: Analysing how teachers evaluate students' writing 1. *Educational Research and Evaluation*, *13* (5), 401–434.

Corbin, J., & Strauss, A. (2008). *Basics of qualitative research* (3rd ed.). Thousand Oaks, CA: Sage.

Driver, R., Newton, P., & Osborne, J. (2000). Establishing the norms of scientific argumentation in classrooms. *Science Education*, *84*(3), 287–312.

Duschl, R. (2007). Quality argumentation and epistemic criteria. In S. Erduran & M. P. Jiménez-Aleixandre (Eds.), *Argumentation in science education: Perspectives from classroom-based research* (pp. 159–175). New York, NY: Springer.

Erduran, S., Simon, S., & Osborne, J. (2004). Tapping into argumentation: Developments in the application of Toulmin's Argument Pattern for studying science discourse. *Science Education*, *88*(6), 915–933.

Evnitskaya, N., & Morton, T. (2011). Knowledge construction, meaning-making and interaction in CLIL science classroom communities of practice. *Language and Education*, *25*(2), 109–127.

Gipps, C. (1995). Reliability, validity and manageability. In H. Torrance (Ed.), *Evaluating authentic assessment: Problems and possibilities in new approaches to assessment* (pp. 105–123). Buckingham: Open University Press.

Jakobsson, A., Mäkitalo, Å., & Säljö, R. (2009). Conceptions of knowledge in research on students' understanding of the greenhouse effect: Methodological positions and their consequences for representations of knowing. *Science Education*, *93*(6), 978–995.

Jiménez-Aleixandre, M. P., & Erduran, S. (2007). Argumentation in science education: An overview. In M. P. Jiménez-Aleixandre & S. Erduran (Eds.), *Argumentation in science education* (3–27). New York, NY: Springer.

Jiménez-Aleixandre, M. P., Rodríguez, A. B., & Duschl, R. A. (2000). "Doing the Lesson" or "Doing Science": Argumentation in high school science. *Science Education*, *84*(6), 757–792.

Jönsson, A. (2011). *Lärande bedömning* (2nd ed.) [Learning assessment]. Malmö: Gleerups.

Kisiel, J. F. (2010). Exploring a school-aquarium collaboration: An intersection of communities of practice. *Science Education*, *94*(1), 95–121.

Koh, K., & Luke, A. (2009). Authentic and conventional assessment in Singapore schools: An empirical study of teacher assignments and student work. *Assessment in Education: Principles, Policy & Practice*, *16*(3), 291–318.

Kolstø, S. D. (2000). Consensus projects: Teaching science for citizenship. *International Journal of Science Education*, *22*(6), 645–664.

Kolstø, S. D., & Ratcliffe, M. (2007). Social aspects of argumentation. In M. P. Jiménez-Aleixandre & S. Erduran (Eds.), *Argumentation in science education* (pp. 117–135). New York, NY: Springer.

Linder, C., Östman, L., Roberts, D. A., Wickman, P. O., Erickson, G., & MacKinnon, A. (2011). *Exploring the landscape of scientific literacy*. New York, NY: Routledge.

Linn, R. L., Baker, E. L., & Dunbar, S. B. (1991). Complex, performance-based assessment: Expectations and validation criteria. *Educational Researcher*, *20*(8), 15–21.

Linell, P. (1998). *Approaching dialogue: Talk, interaction and contexts in dialogical perspectives* (Vol. 3). Amsterdam: John Benjamins Publishing.

Lundqvist, E., & Lidar, M. (2013). Nationella prov i NO och lärares val av undervisningsinnehåll [National assessments in science and teachers' choices of content]. *Utbildning & Demokrati*, *22*(3), 85–106.

Moss, P. A. (2003). Reconceptualizing validity for classroom assessment. *Educational Measurement: Issues and Practice*, *22*(4), 13–25.

Moss, P. (2007). Reconstructing validity. *Educational Researcher*, *36*(8), 470–476.

Munby, H., Taylor, J., Chin, P., & Hutchinson, N. L. (2007). Co-op students' access to shared knowledge in science-rich workplaces. *Science Education*, *91*(1), 115–132.

Newton, P., Driver, R., & Osborne, J. (1999). The place of argument in the pedagogy of school science. *International Journal of Science Education*, *21*(5), 553–576.

Nielsen, J. (2013). Dialectical features of students' argumentation: A critical review of argumentation studies in science education. *Research in Science Education*, *43*(1), 371–393.

Ødegaard, M. (2003). Dramatic science. *A Critical Review of Drama in Science Education. Studies in Science Education*, *39*(1), 75–101.

Ratcliffe, M., & Grace, M. (2003). *Science education for citizenship*. Philadelphia, PA: Open University Press.

Sadler, T. D., & Zeidler, D. L. (2009). Scientific literacy, PISA, and socioscientific discourse: Assessment for progressive aims of science education. *Journal of Research in Science Teaching*, *46*(8), 909–921.

Sambell, K., McDowell, L., & Brown, S. (1997). "But is it fair?": An exploratory study of student perceptions of the consequential validity of assessment. *Studies in Educational Evaluation*, *23* (4), 349–371.

Schegloff, E. A. (2007). *Sequence organization in interaction: A primer in conversation analysis, Vol. 1*. Cambridge: Cambridge University Press.

Serder, M., & Jakobsson, A. (2014). "Why bother so incredibly much?": Student perspectives on PISA science assignments. *Cultural Studies of Science Education*. doi:10.1007/s11422-013-9550-3

Simonneaux, L. (2001). Role-play or debate to promote students' argumentation and justification on an issue in animal transgenesis. *International Journal of Science Education*, *23* (9), 903–927.

Skolverket. (2010). *Ämnesproven i biologi, En redovisning av utprövningsomgången 2009* [The National assessments in biology, a report from the trial 2009]. Stockholm: Skolverket.

The Swedish National Agency of Education. (2011a). *Curriculum for the compulsory school, preeschool class and leisure-time centre 2011*. Stockholm: Skolverket.

The Swedish National Agency of Education. (2011b). *Kommentarmaterial till kursplanen i kemi* [Comments to the chemistry syllabus]. Stockholm: Skolverket.

Säljö, R., Mäkitalo, Å., & Jakobsson, A. (2011). Appropriering genom argumentation: Kontroverser, mångtydighet och redskap för tänkande [Appropriation through argumentation: Controversies, ambiguity and tools for thinking]. In R.Säljö (Ed.), *Lärande och Minnande* (pp. 128–157). Stockholm: Nordstedts.

Tatli, A., & Özbilgin, M. F. (2012). An Emic approach to intersectional study of diversity at work: A Bourdieuan framing. *International Journal of Management Reviews*, *14*(2), 180–200.

Toulmin, S. E. (1958/2003). *The uses of argument*. Cambridge: Cambridge University Press.

Terwilliger, J. (1997). Research news and comment: Semantics, psychometrics, and assessment reform: A close look at "authentic" assessments. *Educational Researcher*, *26*(8), 24–27.

Waugh, C. K., & Gronlund, N. E. (2013). *Assessment of student achievement* (10 ed.). Upper Saddle River, NJ: Pearson.

Wenger, E. (1998). *Communities of practice: Learning, meaning and identity*. New York, NY: Cambridge University Press.

Wiggins, G. (1989). A true test. *Phi Delta Kappan*, *70*(9), 703–713.

Willis, J. (2011). Affiliation, autonomy and assessment for learning. *Assessment in Education: Principles, Policy & Practice*, *18*(4), 399–415.

Zhu, Y., & Bargiela-Chiappini, F. (2013). Balancing Emic and Etic: Situated learning and ethnography of communication in cross-cultural management education. *Academy of Management Learning & Education*, *12*(3), 380–395.