J-STEM

Journal of Research in STEM Education

Vol 1, No 2, December 2015

j-stem.net

Contact

All queries related to manuscript submissions can be directed to Dr. Aydeniz, the Editor-in-Chief, jstemeditor@gmail.com

Mehmet Aydeniz, PhD.
Associate Professor of Science Education.
Program Coordinator, Science Education.
Department of Theory and Practice in Teacher Education
The University of Tennessee, Knoxville
A 408 Jane & David Bailey Education Complex
Knoxville, TN 37996-3442
USA
Phone: +1-865-974-0885

Publisher / Founder : i-STEM / Gultekin Cakmakci

ISSN:2149-8504

Journal of Research in STEM Education
ISSN:2149-8504 (online)
Vol 1, No 2, December 2015, PP 77.

EDITORIAL

Mehmet Aydeniz, Lynn Hodge

The University of Tenesse, Knoxville, USA

We are excited to let you know that our second issue is now published. In the second issue of the Journal of Research in STEM Education, J-STEM, we publish five articles, each focusing on different aspects of STEM education and making unique contributions to the ongoing discussion on how to best make sense of and teach STEM.

Harrison (2016) in her article "Assessment for Learning in Science Classrooms" draws ideas from a number of studies about classroom talk to make a case for and propose some of the preconditions for effective feedback in science classrooms. She provides the theoretical basis for formative assessment pedagogy and the opportunities such pedagogy creates for teachers to attend and respond to student learning.

The second article, Explicit Teaching and Scaffolding to Enhance Concept Learning by Design Challenges" by Van Breukelen, Smeets and De Vries (2016) builds on two previous studies and provides insight into how using design-based challenges promotes pre-service science teachers' conceptual understanding of science ideas in addition to their performance in design and investigative skills. Moreover, they elaborate on the role of scaffolding in reported improvements in students' concept learning and performance in investigative tasks.

The third article on this issue, Developing Effective STEM Animations: Application of a Multimedia Learning Theoretical Framework" by Adetunji and Levine (2016) introduces a Multimedia Learning Environments theoretical framework for developing effective STEM animations. They present both the process employed in the development of these animations along with data on the wide-spread dissemination of Sci-Toons, its impact on viewers, and its impact on students involved in their production.

The fourth article, "A "Scientist" on the radio: Understanding the framing of STEM to the public," focuses on STEM communication, Bowen, Zurawski, and Bartley (2016) document the processes by which listener interests ultimately end up discussed in a radio broadcast and what influences the "science" that is presented on-air. They discuss the ways in which the STEM topics and content are mediated by radio station personnel, often times distorting the factual content available to the public and misrepresenting the practices of the research fields. They discuss how "the interests of the commercial radio station around "entertainment" provides a distorted view of science.

Finally, Mueller, Hall and Miro (2016), in their article entitled, "Testing an Adapted Model of Social Cognitive Career Theory: Findings and Implications for a Self-Selected, Diverse Middle-School Sample" provides the results of a research study in which they tested how SCCT predicts a sample of middle school girls' career choices. Their study has implications for cultivating students' interests in STEM-related careers.

Collectively these articles bring out interesting points on how to best promote and teach STEM in both formal and informal learning contexts. In doing so, they draw attention to resources, discourse, and instructional practices that support STEM learning and the interaction of contextual aspects that inform individual's perspectives of STEM fields and professions. These ideas contribute to ongoing efforts to further make sense of STEM as an integrated concept and as distinct content areas that share common dispositions and practices.

Journal of Research in STEM Education
ISSN:2149-8504 (online)
Vol 1, No 2, December 2015, PP 78-86

RESEARCH REPORT

Assessment for Learning in Science Classrooms

Christine Harrison[1]

King's College, London, UK

Abstract: *Classroom assessment has grown in prominence over the last few decades and particularly the formative approach to instruction through assessment for learning (AfL). This paper draws across a number of studies about classroom talk to make a case for and propose some of the preconditions for effective feedback through Assessment for Learning interactions in science classrooms. As such, it provides an underpinning for formative pedagogy that structures classroom activities to provide more feedback and so create opportunities for teachers to respond to learners' needs. Creating effective pedagogy, where feedback drives future learning, is a complex set of practices that requires both novice and experienced teachers to think carefully how they might build and evolve activities and dialogue, that help students voice and develop their conceptual understanding and an understanding of how they learn.*

Keywords: Formative assessment, dialogic classroom, feedback, responsive teaching

One of the most influential drivers for the way assessment has been conceptualised over the last three decades has been the Black and Wiliam (1998) paper which surveyed the evidence that linked assessment with improvement in learning. The goal of assessment, in the context of assessment for learning (AfL), requires a prospective view of learning in which the concern is not solely with the actual level of performance, but with anticipating future possibilities (Heritage, 2010).

Since the 1998 Black and Wiliam review, a number of researchers have proposed and debated the definition and value of AfL (Bennett, 2011; Black & Wiliam, 2009; Kingston & Nash, 2011; Klenowski, 2009; Swaffield, 2011). AfL is an approach to pedagogy that allows students to discuss and share their ideas with others (Black et al, 2002). Over time there has been an increasing emphasis with AfL on students taking an active role in their own learning and assessment processes. It has also been recognised that it is not sufficient that students merely learn how to address their immediate learning challenges. AfL needs to enable and empower students to learn how to learn and to motivate them to keep on learning. Indeed, Boud (2000) argues that assessment is sustainable only if it 'meets the needs of the present without compromising the ability of students to meet their own future learning needs' (Boud, 2000, p. 151).

Assessment reflects teachers' understanding of learning, as well as what is valued (Drummond, 2008). It is based on the premise of feedback and the aim is to strengthen and facilitate feedback through a variety of routes in the classroom from promoting discussion to providing comments on pieces of student work and supporting learners in peer and self-assessment scenarios. Sadler (1989) highlighted the 'indispensible' role of active student participation in the process and action on feedback.

Sadler also argued that formative assessment should be intrinsic to, and integrated with, effective instruction:

[1] Department of Education & Professional Studies, King's College, UK. E-mail: christine.harrison@kcl.ac.uk

Harrison, C. (2015). Assessment for learning in science classrooms . *Journal of Research in STEM Education*, *1*(2), 78-86

Formative assessment is concerned with how judgments about the quality of student responses (performances, pieces, or works) can be used to shape and improve the student's competence by short-circuiting the randomness and inefficiency of trial-and-error learning (Sadler 1989, p. 120).

This paper looks at aspects of AfL in science classrooms, focusing particularly on classroom discussion. It begins by looking at the type of work that scientists do to promote scientific thinking and uses this to conceptualize how science teachers shape their instruction to help learners improve their scientific understanding. It looks specifically at the role of questions and classroom dialogue in science classrooms and how these help students advance and reshape their ideas about science concepts.

Science and Scientific Thinking

Science is about finding explanations of the various phenomena that we encounter in the world around us. What is difficult about science is that these explanations are not simple and often require a consideration of several pieces of evidence, some of which seem contradictory (Osborne, 2015). For example, gardeners add fertilisers, which are high in nitrates, to soil to help plants grow, and yet biochemists tell us that plants are composed of mainly carbon compounds. These two facts together do not make sense until the scientist outlines the details of the processes that nitrates and carbon compounds are involved within the plant and then begins to look for links between these various processes. In fact, the carbon compounds in plants are produced in the process of photosynthesis, when atmospheric carbon dioxide is bonded with water to produce glucose. The glucose is then used as a chemical building block to build all the materials in the plant, so that new cells can be formed. One of the follow-on reactions in the plant combines the nitrates from the fertiliser to some of these glucose molecules to eventually form amino acids. These amino acids are then linked together into proteins that form the cytoplasm in the plant cells. To link the initial two facts and understand the processes involved in plant growth takes several steps of explanation and other pieces of evidence that need to be linked together in a specific way to understand the concept.

As you can see from this example, explanations in science involve a wide range of conceptual understanding, skills and ideas. Many of these ideas are quite abstract or simply difficult to comprehend. It is more than linking facts together to form an idea. It involves exploring these ideas so that they make sense of the situation and requires careful consideration of evidence and an ability to ask questions and look for solutions. In science classrooms, students develop the lifelong skills critical to thinking creatively, as they learn how to solve problems using logic and reasoning. These skills are essential for several tasks in science, including drawing conclusions from experimental data (Harrison, 2014). Once ideas are debated and shared, scientists propose relationships, laws and rules to explain the phenomena experienced. Experiments allow scientists to test these ideas out through predicting what should happen according to their theories of how the world works (Tiberghien, 2000).

These skills are essential for several tasks in science, including drawing conclusions from experimental data. While many famous scientists seem to have worked on their ideas as individuals, most scientists today work in teams. This is partly because the ideas they are testing are often crosscurricular or require some aspect of application and may require specialists with technological, mathematical or computer modelling experience as well as an understanding of biology, physics or chemistry. Ideas, within these teams of scientists, are discussed and challenged and new evidence is often generated through experimentation or alternative interpretations of previous data. The new understanding that the science teams reach through this process requires the support and challenge from others. In other words, science is socially constructed, where the aim is to make sense of the phenomena that humans encounter in their everyday lives through explanations generated though collaboration and communication with others (Kuhn, 1962). Fundamentally science is a creative process (Osborne, 2015), its difference from other subject disciplines being in the consideration and testing of evidence.

The Role of Talk in Science Classrooms

Many science teachers believe that it is important to create classroom scenarios that emulate the way that scientists work. (Brickhouse, 1990; Duschl, 1990; Gallagher, 1991; Lederman, 1986, 1991). While it would be ludicrous to expect students to always start from phenomena and work towards theoretical understanding in a discovery-type approach, creating activities in which students' ideas are considered as part of the learning sequence is important. The problem is that most teachers do not plan learning experiences where the aim is to share and work with student ideas about science. Instead teachers plan how to introduce the ideas of science to students and so the emphasis is not on interaction but on how the teacher provides explanations of scientific ideas. From the student viewpoint, science is transmitted as a series of facts and theories that is provided for them to learn by the teacher. If student's ideas are to form part of the learning scenario in science lessons, then there needs to be a change in pedagogy.

Teacher-talk dominates many UK science classrooms, where talk is a vehicle for question-and-answer recitation teaching (Alexander, 2006). In this approach, the teacher controls and dominates the classroom talk and the role of the learner is to guess the answers that the teachers hold inside their heads. Learners are encouraged to participate but only at a level of agreement or elaboration and their role is to provide the 'right answer' in the correct place in the teacher's narration as this extract from one of the science teachers on the King's-Medway-Oxfordshire Formative Assessment project indicates:

> *I'd become dissatisfied with the closed Q&A style that my unthinking teaching had fallen into, and I would frequently be lazy in my acceptance of right answers and sometimes even tacit complicity with a class to make sure none of us had to work too hard … They and I knew that if the Q&A wasn't going smoothly, I'd change the question, answer it myself or only seek answers from the 'brighter students'. There must have been times (still are?) where an outside observer would see my lessons as a small discussion group surrounded by many sleepy onlookers.* James, Two Bishops School (Black et al., 2002)

It has been widely accepted for several decades now that learners' cognitive development is driven by interactions between children, adults and society (Vygotsky, 1978, Brunner & Haste, 1987, Halliday, 1993). When students are faced with new experiences they need to make sense of them (Lindfors, 1999). Language is at the heart of this process. The learner uses talk to engage with the new knowledge and to try to understand it within their own personal frameworks through interactions with other learners and their teacher. Part of this they achieve through comparison with their previous thinking in that area but the major part of this learning is in negotiating common meaning with others that are also engaged in the learning experience. In this way, new knowledge is socially constructed (Vygotsky, 1978) and communication through dialogue is essential in achieving this. Barnes (1976) believes that talk is essential for students to reflect upon and change the ways they interpret reality and this is vital in the science classroom.

Much learning may go on while children manipulate science apparatus or during a visit, or while they are struggling to persuade someone else to do what they want. But learning of this kind may never progress beyond manipulative skills accompanied by slippery intuitions, unless learners themselves have an opportunity to go back over such experiences and represent it to themselves. There seems every reason for group practical work in science, for example, normally to be followed by discussion of the implications about what has been done and observed, since without this what has been half understood may slip away. (Barnes (1976 pp30-31)

In science lessons, students should be encouraged to explore their ideas about science and look for evidence that either supports or refutes their scientific thinking. This requires science lessons to focus more on oracy than literacy (Harrison, 2006) where questions and talk foster a culture of inquisitiveness (Chin, 2004). Science teachers need to develop classroom experiences where this can happen and create learning environments where students can learn ' with and from one another' (Heritage, 2007, p.144). This is very different from the more traditional approach to science teaching where the teacher delivers science ideas as though they are facts to be learnt and where practical work simply illustrates the relationships, laws and theories of science.

Students are introduced to scientific ideas, experience phenomena and construct meaning through discussion in the science classrooms. By sharing ideas with peers and interacting with their teacher they engage with exploratory talk (Asoko & Scott, 2006) that develops their understanding within a social context (Duit & Treagust, 1998). In this way, classroom talk is not simply about conveying what has been learned but has value in helping learning. The science teacher stimulates this discourse through the stimuli they introduce students to and the questions they ask. The types of questions that teachers ask and the ways that they use them in the classroom strongly influences the discussion and therefore influences the cognitive processes that students engage in as knowledge is co-constructed (Chin, 2006).

The importance of talk in science lessons is possibly more evident in primary classrooms and Harlen (2006) highlights the use of discussion for reflection and communication, that enables children to sort out their ideas as well as engage in science activities. Keogh and Naylor (2007) found that students freely engage in exploratory talk when presented with scientific questions. Such talk involves linking old ideas with new ones to create bigger and new ideas. It also requires the child to be prepared to consider and change ideas, if a peer or the teacher challenges their idea or produces an alternative that seems more compelling. Keogh and Naylor (2007) are keen proponents of the notion that science lessons should be environments where students are willing to consider alternative ideas, justify their opinions and base their decisions on evidence and reasoning. This flexibility of mind and encouragement to think is essential to science learning. While talk has a role in enhancing and ensuring cognitive development, it also has a role helping students regulate their learning. Talk is therefore highly valued in science as it provides the vehicle to engage in scientific thinking.

While questioning plays a regular part in most science activities in the classroom (Van Zee et al., 2001; Harlen, 2006) to be effective it should promote dialogue and progress students' scientific thinking. The skill for the teacher is in asking questions that both tap into student ideas and also anticipates the degree of detail that is needed in an answer. Khwaja and Saxton, (2001) highlight the importance of question clarity in the following example. If a student is given the question 'what is inside the body' compared to the instruction 'draw the bones inside the body', the outcomes may be very different. One version may give rise to a child showing a good understanding of, for example, the ribs but the other may give misleading information because the child omits the ribs from the diagram in order to show the organs underneath the skeleton. It would be easy for a child's understanding to be misinterpreted or for misconceptions to be missed because of the specific question being asked. It is therefore important to both bring clarity to questions and instructions in the science classrooms and also to anticipate how the learners might respond. Indeed, one of the skills that examination question writers work on is developing questions by starting with the answers they feel show competence and then working backwards to what sort of questions would produce such answers in the detail and scope they want.

Black and Harrison (2001) suggest the use of 'big questions' in science lessons as these can be used to set the scene for the whole lesson with the subsequent development of smaller questions being introduced or developed to help build up ideas towards answering the 'big question' for that lesson. So asking a big question like 'what is friction like on the Moon?' requires the students to think about whether gravity plays a role in friction and how conditions are different on the moon compared to Earth as well as explaining what friction is and does. This creates a range of assessment opportunities for the teacher as students unpack their ideas for each of the the subsidiary questions and then utilize this knowledge to make connections to explain the big question. This approach helps both the students and teacher unpack the various ideas involved and encourages a more dialogic approach in the science classroom which encourages learning to take place. Harlen (2006) also discusses the usefulness of introducing a question and then reviewing it at different points in the learning sequence, to both assess and monitor developing understanding. This interplay of assessment within a teaching sequence allows for more formative opportunities during the lesson such that errors or incomplete understanding can be identified and worked on as they become evident within the classroom talk. In this way, teachers become more responsive in their teaching and students come to recognise that, by voicing their ideas, they can be helped to improve and consolidate their understanding of science concepts.

Creating the AfL Classroom

Creating classroom talk that enables teachers to tap into student thinking is central to the classroom environment that is needed for feedback to function. Part of this will be whole class discussion, where the teacher orchestrates feedback by asking probing questions about what students currently think or by asking students to reflect on what they have just experienced in an activity. Another part of the feedback is generated from collaborative group work, where students exchange ideas with one another and begin to test out, alongside their peers' ideas, what they think is happening within a science activity. In both these situations, the student receives feedback that shapes their understanding but the teacher also is picking up considerable assessment evidence of understanding, alternative ideas and student needs. It allows formative assessment to be dynamic and teaching and learning to be responsive. However, teachers need to work on both planning for and orchestrating classroom talk so that the focus is on diagnosis and improvement of student thinking and not just accuracy and correctness of what is said.

Teachers use a range of interventions also help scaffold the emerging ideas as they arise within groups. Having knowledge of the likely problems or misconceptions that might arise within a topic is essential here. The teacher can then plan an activity that helps learners reveal their emerging ideas or that might challenge their existing ideas. Through careful eavesdropping of student conversations and whole class talk the teacher comes to understand what this specific groups of learners know, what they partly know and what they do not yet know (Black & Harrison, 2004). This helps teachers to more carefully pitch the next events in the learning, be they in the discussion that follows or in follow-on activities in same lesson or in subsequent lessons. In this way, the "partly known" ideas are sorted out and unknown ideas introduced as well as the correct ideas consolidated. This approach is a type of responsive teaching, where the teacher responds to the learners' needs. This allows learners to work at their 'leading edge of learning' as their knowledge base is continually being challenged and upgraded through this process.

There have been numerous, mostly qualitative articles, reporting on small scale case studies of AfL, often carried out by teachers. These articles tend to explore the virtues of a particular AfL strategy and provided exemplification of how these strategies worked in the classroom. (Stow, 1997, Black & Harrison, 2001; Leakey, 2001; Harrison & Howard, 2009, 2011; Keogh & Naylor, 1998, 2007: Harrison, 2006, 2014) A number of articles also reported case studies of teachers who monitored their understanding of AfL and how this changed over a period of time, for example as a result of an intervention or from doing particular activities (Black & Harrison, 2001. Sato et al, 2005; Harrison, 2005, 2006, 2013, 2014). Another key feature of the articles included discussion of the need to develop a classroom climate which fosters a constructivist approach and encourages feedback practices. So, the development of AfL pedagogy is far from simple and requires teachers to both examine and modify their activities to create opportunities for collecting evidence of student ideas and for providing feedback and action that help improve these ideas.

If students come to learn by building ideas and concepts through social construction, then a teacher's role is to facilitate the conditions in which this dialogue can take place. This involves both organizing the social setting so that dialogue is likely to be fruitful and also working with the learner to ensure that they actively engage with the learning activity. The latter is reached through diagnostic exploration of where the student is in their learning and then providing scaffolding (Bruner & Haste, 1962) to take ideas forward. The role of scaffolding is to help learners examine and reshape their emergent ideas. Sometimes scaffolding is misunderstood by teachers. Many believe it is giving hints and clues to enable learners to reach a correct answer. Instead teachers need to conceptualize scaffolding as probes and prompts that allow students to reveal their ideas so allowing the learner to examine and then re-examine the sense they are making of the shared meaning arising from the class discussion. With scaffolding, as with formative assessment, teachers provide various forms of material, social or conceptual assistance aimed at supporting students' reasoning, participation, and learning (Sawyer, 2006). Through scaffolding, the learner can enter the dialogue about shared ideas and begin to see other student's thinking and so make judgements about where they are in their own sense-making. "Fading" is critical component of scaffolding (Pea, 2004). In this teachers gradually withdraw their support and handover control to students as learners develop their understanding. This helps promote student ownership of ideas.

Alexander (2006) takes these ideas further and suggests that teachers adopt a dialogic approach towards pedagogy, which he defines as classroom interaction that include questions that provoke thoughtful answers that are used to promote further questions and comment. This means that the goal is to use questions as the building blocks of dialogue rather than as a set of disconnected answers. It is common practice for teachers to plan for 'question and answer' sessions in their lessons, but Alexander believes they need to replace this approach with questions that can seed productive discussion. Through the process of dialogic teaching, individual teacher-student and student-student exchanges are chained into coherent lines of enquiry where connections, relationships and differences help shape the understanding of all those involved. Two features of Alexander's perception of classroom dialogue are relevant here. The first is that dialogue involves 'reciprocity' and this encapsulates the idea that both students and teachers learn from engagement in classroom talk. Students begin to see how their scientific ideas fit with that of their peers and that of their teacher and to consider how these might change. Teachers begin to understand students' emergent ideas, and sometimes misconceptions that students may hold, and also how these change as the learners interact with others and have their ideas challenged or consolidated. The second is the idea that dialogue is ' cumulative (Alexander, 2004) and this represents the building capacity of learning through interactions with others and possibly explains the link between AfL and improvement.

Good classroom talk is not always easy to achieve and sustain with all classes and requires teachers to:

- create challenging classroom environments where it is safe to make mistakes
- produce good questions that challenge and spark off ideas
- train their learners to listen to and respond to the ideas of their peers
- group students in a way that encourages them to share ideas
- listen in to student discussion, noting strengths and weaknesses
- assess what the answers indicate about the students' degree of understanding
- facilitate and sustain discussion with well-timed interventions or probes
- make decisions about which student ideas to feedback to the whole class
- recognise individual student needs and how to respond to these
- use evidence of learning to respond to and plan later activities

Of course, classroom talk is not always dialogic in form; there are some occasions when the teacher is not interested in exploring students' ideas and taking account of them in the development of the lesson. Here the teacher is likely to focus on the science point of view and if ideas or questions, which do not contribute to the development of the school science story, are raised by students they are likely to be re-shaped or ignored by the teacher. This kind of talk is authoritative in nature (Scott & Ametller, 2007). Scott et al (2006) argue that shifts between communicative approaches are an inevitable part of teaching whose purpose is to support meaningful learning of scientific knowledge. They suggest that a necessary tension therefore exists between authoritative and dialogic approaches and that it is the interplay between these two types of dialogue that supports development of scientific understanding, So sometimes students are engaged in learning about a science topic as the teacher provides narrative and explains accepted scientific ideas, while, at other times, the students need to be engaged in exploring their own scientific thinking and how this fits with accepted science ideas.

Summary

AfL has revolutionised many science classrooms and focusing on questions and classroom talk as a vehicle for learning has enabled teachers to create classroom environments more closely aligned with the ways that science works. Developing science classrooms where dialogue drives learning is an ambitious approach for science teachers because they need to think carefully about how they use activities to help students voice and share ideas about scientific phenomena. While many teachers have attended professional development courses or engaged with on-line training about AfL, many of these opportunities for professional learning are not domain specific and so science teachers need to consider how AfL strategies fit with the classroom activities they use with students. It is clear that the STEM community needs to work on the way that talk is used in classrooms in order to help students strengthen their scientific understanding and this therefore has implications both for pre-service and in-service training.

References

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

Asoko, H. and Scott, P. (2006). 'Talk in science classrooms.' In W. Harlen (Ed.), *ASE Guide to Primary Science Education*. Hatfield: Association for Science Education.

Barnes, D. (1976). *From communication to curriculum*. Harmondsworth, England: Penguin.

Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy and Practice, 18*, 5-25.

Black, P., & Harrison, C. (2001). Feedback in questioning and marking: the science teacher's role in formative assessment. *School Science Review*, *82*(301), 55–61.

Black, P., & Harrison, C. (2004). *Science inside the black box*. London: GLAssessment

Black, P., Harrison, C., Marshall, B., Lee, C., & Wiliam, D. (2002). *Assessment for learning: Putting it into practice*. London: OUP.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles Policy and Practice*, *5*(1), 7-73.

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, *21*(1), 5-31.

Boud, D. (2000). Sustainable assessment: Rethinking assessment for the learning society. *Studies in Continuing Education*, *22*, 151–167

Brickhouse, N. W. (1990). Teachers' beliefs about the nature of science and their relation- ship to classroom practice. *Journal of Teacher Education*, *41*, 53–62.

Bruner, J. S. & Haste, H. E. (ed) (1987). *Making sense: the child's construction of the world.* London:Routledge

Chin, C. (2004). Students' questions: fostering a culture of inquisitiveness in science classrooms. *School Science Review, 86*(314), 107–112.

Chin, C. (2006). Classroom interaction in science: teacher questioning and feedback to students' responses. *International Journal of Science Education*, *28*(11), 1315– 1346.

Cowie, B. (2005). Pupil commentary on assessment for learning. Curriculum Journal, 16, pp.137–151.

Duit, R., & Treagust, D. F. (1998). Learning in science - From behaviourism towards social constructivism and beyond. In B. J. Fraser, & K. Tobin (Eds.), International handbook of Science Education, Part 1 (pp. 3-25). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Duschl, R. A. (1990). *Restructuring science education*. New York: Teachers College Press.

Gallagher, J. J. (1991). Perspective and practicing secondary school science teachers' knowl- edge and beliefs about the philosophy of science. *Science Education*, *75*, 121–134.

Harlen, W. (2006). *Teaching, learning and assessing science 5-12*. London: Sage.

Harrison, C. (2005). Teachers developing assessment for learning: Mapping teacher change. *International Journal of Teacher Development*, *9*(2), 255–263.

Harrison, C. (2006). Banishing the quiet classroom. *Education Review*, *19*(2), 67-77

Harrison, C., & Howard, S. (2009). *Inside the Primary Black Box: Assessment for Learning in the Classroom.* London: GLAssessment.

Harrison, C,. & Howard, S. (2011). Issues in primary assessment 2: Assessment for learning. *Primary Science*. 116, 5-7.

Harrison, C. (2013). Collaborative action research as a tool for generating formative feedback on teachers' classroom assessment practice: the KREST Project. *Teachers and teaching: theory and practice*, *19*(2), 202–213.

Harrison, C. (2014a). Assessment for learning. In R. Toplis (Ed.), *A formative approach in learning to teach science in the secondary school*. London: Routledge

Harrison, C. (2014b). Assessment of Inquiry Skills in the SAILS Project. *Science Education International*, *25*(1), 112-122

Halliday, M. A. K. (1993). Towards a language-based theory of learning. *Linguistics in Education*, *5*, 93-116

Heritage, M. (2007). Formative Assessment: What do Teachers Need to Know and Do? *Phi Delta Kappan*. *89*(2), 140-145

Heritage, M. (2010). *Formative assessment: making it happen in the classroom*. Thousand Oaks, CA: Corwin Press.

Keogh, B., & Naylor, S. (1998). Teaching and learning in science using Concept Cartoons. *Primary Science Review*, *51*, 14–16.

Keogh, B., & Naylor, S. (2007). Talking and thinking in science. *School Science Review*, *88*(324), 85–90.

Khwaja, C., & Saxton, J. (2001). It all depends on the question you ask. *Primary Science Review*, *68*, 13–14.

Kingston, N., & Nash, B. (2011). Formative assessment: A meta-analysis and a call for research. *Educational Measurement: Issues and Practice*, *30*(4), 28–37.

Klenowski, V. (2009). Assessment for learning revisited : an Asia- Pacific perspective. *Assessment in Education: Principles, Policy and Practice*, *16*(3), 263-268.

Kuhn, T. S.(1962). *The structure of scientific revolutions*. Chicago: University of Chicago Press.

Leakey, A. (2001). Fantastic feedback. *Primary Science Review*, *68*, 22–23.

Lederman, N. G. (1986). Relating teaching behavior and classroom climate to changes in students' conceptions of the nature of science. *Science Education*, *70*, 3–19.

Lederman, N. G. (1992). Students' and teachers' conceptions of the nature of science: A review of the research. *Journal of Research in Science Teaching*, *29*, 331–359.

Lindfors, J. W. (1999). *Children's Inquiry: using language to make sense of the world.* New York: Teacher's College Press

Pea, R. D. (2004). The social and technological dimensions of scaffolding and related theoretical concepts for learning, education, and human activity. *The Journal of the Learning Sciences*, *13*(3), 423-451

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, *18*, 119–144.

Sato, M., Coffey, J., & Moorthy, S. (2005). Two teachers making assessment for learning their own. *The Curriculum Journal*, *16*(2), 177–191.

Sawyer, R. (2014). Introduction: The New Science of Learning. In R. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences: Second Edition* (pp.1-20). Cambridge University Press: Cambridge

Scott, P. H., & Ametller, J. (2007). Teaching science in a meaningful way: striking a balance between 'opening up' and 'closing down' classroom talk. *School Science Review*, *88*(324), Hatfield: ASE.

Scott, P., Mortimer, E. & Aguiar, O. (2006). The tension between authoritative and dialogic discourse: a fundamental characteristic of meaning making interactions in high school science lessons. *Science Education*, *90*, 605-631.

Stark, R. & Gray, D. (2001). What science do they think they know?. *Primary Science Review*, *68*, 8–11.

Stow, W. (1997). Concept mapping: a tool for self–assessment. *Primary Science Review*, *49*, 12–15.

Swaffield, S. (2011). Getting to the heart of authentic Assessment for Learning. *Assessment in Education: Principles, Policy & Practice*, *18*(4), 433-449.

Tiberghien, A. (2000). Designing teaching situations in the secondary school. In R. Millar, J. Leach, & J. Osborne (Eds.), *Improving science education: The contribution of research* (pp. 27–47). Buckingham, UK: Open University Press.

Vygotsky, L. (1978). *Mind in Society*. Cambridge: Massachusetts: Harvard University Press.

Journal of Research in STEM Education
ISSN:2149-8504 (online)
Vol 1, No 2, December 2015, PP 87-105

RESEARCH REPORT

Explicit Teaching and Scaffolding to Enhance Concept Learning by Design Challenges

Dave van Breukelen[a 1], Maurice Smeets[a], Marc de Vries[b]

[a]*Fontys University of Applied Sciences for Teacher Education Sittard, Sittard, The Netherlands*

[b]*Delft University of Technology, Science Education and Communication, Delft, The Netherlands*

Abstract: *This paper presents a mixed methods study, carried out among 21 first-year student teachers, that investigated learning outcomes of a modified Learning by Design (LBD) task. The study is part of a series of studies that aims to improve learning, teaching and teacher training. Design-based science challenges are reasonably successful project-based approaches for breaking down the boundaries between traditional STEM subjects. Previous learning outcomes of the extensively studied LBD approach demonstrated a strong positive effect on students' skills. However, compared to traditional classroom settings, LBD provided little profit on (scientific) concept learning. For this, according to two preliminary studies, a lack of explicit teaching and scaffolding strategies, both strongly teacher-dependent, bears a share of responsibility. The results of the third study discussed in this paper indicate that emphasizing these strategies strengthens concept learning without reducing positive effects on skill performances.*

Keywords: Learning by Design, science, technology, skills, concepts, learning outcomes

The world around us is constantly changing and getting more complex. Partly because science and technology have grown progressively denser in our personal lives, where most of the world's issues ask for an interdisciplinary approach to meet humans' needs. We might expect that school systems respond accordingly by delivering juveniles ready to face these issues. However, many curricula are traditionally dominated by separate disciplines (Scott, 2008) where international studies, e.g. ROSE (Sjöberg & Schreiner, 2010), demonstrate a decreasing interest in and understanding of science and technology. Aikenhead (2006) states that unidisciplinary science curricula result in sterile, dehumanized science content that has little appeal to students and is often perceived by them to be irrelevant. Several studies indicate that a holistic understanding of science and technology, through interdisciplinary teaching, may improve students' motivation and understanding (Lustig et al., 2009; Osborne & Dillon, 2008; Rennie, Venville, & Wallace, 2012). If we want students to learn how to apply knowledge and skills in daily life, their education experiences must involve them in learning and applying knowledge and skills of related disciplines in recognizable contexts (Bybee, 2013). Therefore, many national governments aim for interdisciplinary science, technology, engineering, and mathematics (STEM) education (National Science and Technology Council, 2013; Office of the Chief Scientist, 2013; Parliamentary Office of Science & Technology, 2013). In this context technology should be seen as purposeful and goal-directed activities where knowledge (e.g. conceptual, procedural) and skills (e.g. design, experimentation, craft) are used to solve practical problems and meet needs (International Technology Education Association, 2007).

A quite successful approach to respond to the strong interplay of STEM subjects, called Learning by Design (LBD; Kolodner, 2002b), has been studied extensively from 1999 until 2003. Those studies showed high

[1] Corresponding author. Fontys University of Applied Sciences for Teacher Education Sittard, Sittard, The Nethertlands,
E-mail: d.vanbreukelen@fontys.nl

van Breukelen, D., Smeets, M., & de Vries, M. (2015). Explicit teaching and scaffolding to enhance concept learning by design challenges. *Journal of Research in STEM Education*, *1*(2), 87-105.

student-involvement and, compared to non-LBD classes, significantly better collaboration, metacognitive and science skills. However, concept learning lagged behind (Kolodner, 2002; Kolodner, Camp, et al., 2003; Kolodner, Gray, & Fasse, 2003) what ties in with the call to educate students with a more conceptual understanding of design technology (International Technology Education Association, 2007).

Prior to this study, Study 1 (Van Breukelen, De Vries, & Schure, 2016) and Study 2 (Van Breukelen, Van Meel, & De Vries, 2016) took place that studied LBD's conceptual learning process in detail from a student's and teacher's perspective. By unravelling the learning process it became clear to what extent students used and learned scientific concepts during design activities and how much room for improvement was left. The preliminary studies showed that students were able to manage a conceptual learning gain comparable to average learning gains of many (traditional) physics-related courses (Hake, 1998). In that way, the findings of Kolodner, Gray, et al. (2003) are confirmed. Although students learned science at an apparently acceptable level more progress should be possible because LBD, compared to traditional educational settings, provides a sound theoretical basis for a higher level of concept learning, which will be discussed later on. The preliminary studies revealed two (interrelated) causes that prevented concept learning from reaching a (much) higher level. First, the complexity and extendedness of design challenges obscured scientific content (What to learn?) and forced students to become product- and process-focused (What to do and deliver?). This resulted in the use and learning of loose facts with too little coherence. Second, explication of underlying science had too little attention during task construction and teacher intervention, resulting in addressing an incomplete and disguised framework of conceptual knowledge. This study aims to address these problems by suggesting improvements, based on literature on learning sciences, concerning explicit teaching and scaffolding strategies, which will be discussed in detail later on. The central research questions are: To what extent will application of explicit teaching and scaffolding strategies result in learning a more comprehensive, coherent and explicit framework of scientific knowledge, and how this affects the conceptual learning gains?; Does skill performance, despite the interventions, still increase and how it is (cor)related with concept learning? Finally, a comparison takes place with previous studies to discover why and how the improvements affect (concept) learning.

LBD and desing composition of the challenge used for this study

LBD contains, as shown in Figure 1, two essential components for learning skills, practices and content: (re)design and investigation. Within these components are a variety of reflective hands-on and heads-on activities concerning design technology, science practices, public presentation, collaboration and teacher-guided class discussions. Students (operating in design groups) are faced with a design challenge where they first have to explore and establish things they need to know/learn for succeeding. By information seeking and experimentation they find answers to questions raised in order to apply them in the design

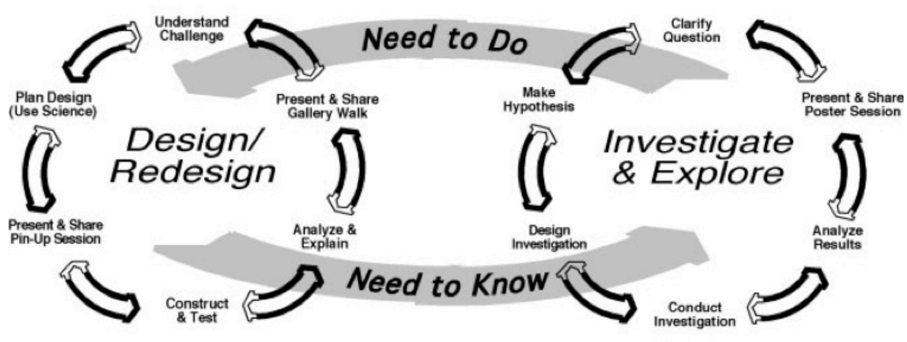

Figure 1. *Learning by Design's Cycles. Reprinted from Kolodner (2002b, p. 339).*

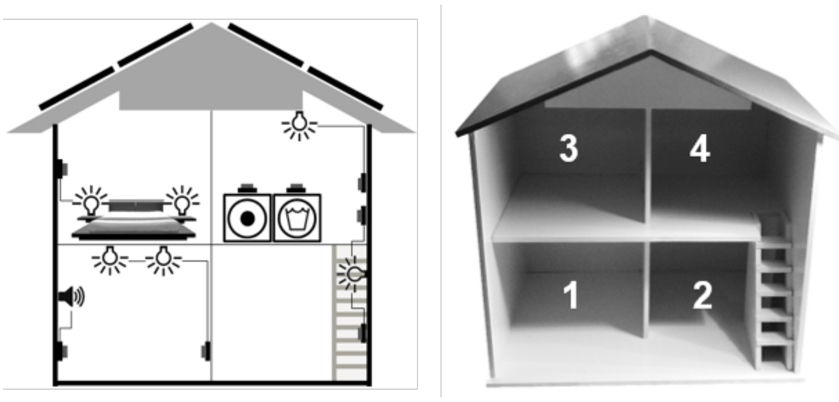

Figure 2. *Model house and layout*

Investigation of this application may lead to additional questions and reinvestigation. To incentivize the understanding of design-related principles and concepts, teacher-guided activities take place (poster and pin-up sessions, whole-class discussions and gallery walks). During these activities experiences and insights are shared among groups, feedback is being given and science is being made explicit. In general, LBD provides a constructivist learning environment where students experience the necessity to learn (Kolodner, Hmelo, & Narayanan, 1996) driven by the fact that students' pre-task conceptions are not sufficient for succeeding: design challenges deliberately address cognitive conflicts. A more scientific framework of knowledge is necessary to cope conflicts and reach conceptual change (Abdul Gafoor & Akhilesh, 2013; Cobern, 1994). Based on literature, e.g. Brandsford, Brown, Donovan, and Pellegrino (2003), LBD contains several elements that are beneficial to conceptual change: collaboration, reflection, contextual learning, applying what is learned, learning from failures and iteration, and connecting skills, practices and concepts.

For this study, an existing LBD challenge was modified for better concept learning. The challenge originated from Study 2 that also concerned first-year student teachers (science). Design groups (3 students per group, randomly chosen) were challenged to design a solar power system for a model house, shown in Figure 2, by taking into account the design specifications in Table 1 that headed for using underlying science, decision-making and creative thinking. Regarding these specifications and the scientific objectives in Table 2 the most fundamental design principles concerned proper wiring (combining series and parallel parts) and regulating current, voltage and resistance for maximum efficiency. Table 3 shows how the LBD elements in Figure 1 were applied resulting in an activity, guided through an instructive presentation and a student's and teacher's guide, that took six periods of 90 minutes.

Table 1

Design specifications and components

Design specifications			
A. [Room layout] Room 1: 2 lamps operated by 1 switch and a doorbell (SP) operated by 1 switch. Room 2: 1 lamp operated by a set of 2-way switches (staircase wiring). Room 3: 2 lamps operated by 1 switch. Room 4: 1 lamp operated by 1 switch, 1 washing machine (M2) with adjustable speed operated by 1 switch, 1 dryer (M1) operated by 1 switch.			
B. [Solar powering] The entire lightning has to be connected to a separate (combination of) solar cell(s). The same applies to the doorbell and washer-dryer combination.			
C. [Efficiency] The energy efficiency of the entire wiring has to be as high as possible and, in addition, the use of materials as less as possible. In any case, it is not allowed to use more components than available.			
Component	Quantity	Component	Quantity
Motor 1.5V DC (M1)	1	Solar cell 4V/35mA	4
Motor 3.0V DC (M2)	1	Solar cell 5V/81mA	4
Mini-speaker 800mW (SP)	1	Solar cell 0.5V/400mA	4
Set of LEDs, resistors, wires and switches	1	Solar cell 0.5V/800mA	4

Note. Reprinted from Van Breukelen, Van Meel, et al. (2016).

Table 2

Scientific objectives and interrelatedness with the challenge

DC electric circuits objectives	Interrelatedness
1. Physical aspects of electric circuits: resistance is a property of an object and hinders current flow (Ohm's Law); equivalent resistance in series increases and in parallel decreases as more elements are added; the necessity of a closed circuit to enable current flow; interpret pictures, diagrams, symbols of a variety of circuits.	• Resistors are necessary to reduce current flow and a variable resistor is necessary to adjust the washing machine's speed. Furthermore, students have to interpret and design a variety of circuit parts in order to meet the requested wiring.
2. Energy and power: apply the concepts of energy (dissipation, conversion and conservation) and power (work done per unit time) to a variety of circuits.	• Students have to establish the amount of energy supply and consumption by the designed circuit in order to reach maximum efficiency.
3. Current: understand and apply conservation of current (Kirchhoff's point rule) to a variety of circuits; explaining the behavior of an ideal current source.	• Combining series and parallel parts (solar cells and components) to meet design specification forces students to investigate and calculate current flow and potential differences. Furthermore, students have to investigate the behavior of (combined) solar cells to get informed about differences compared to (well-known) voltages sources.
4. Potential difference, voltage: the amount of current is influenced by potential difference; apply the concept of Kirchhoff's loop rule (åV = 0 around a closed loop); explaining the behavior of an ideal voltage source.	

Note. Reprinted from Van Breukelen, Van Meel, et al. (2016).

Table 3
Stages and activities

Stages [time]	Activities[a]	Final products[b]
Introducing the Challenge and Context [15-20 min]	Introduction of context, design challenge, activities, organization, learning sources, time schedules, materials, objectives, etc.	
Understanding the Challenge, Messing About, White boarding [50-60 min]	Exploration of the challenge, context and objectives (G) Writing down ideas, (research) questions and hypotheses (G): what to do and learn? White boarding: sharing results; feedback session (C)	Design diary stage 2 Flip chart for white boarding (G)
Investigate & Explore, Poster Session [120-180 min]	Formulate and distribute (scientific) research questions (C) Discussion "fair test rules of thumb" (C) Design and conduct experiments, collect data, conclude (G) Presentation of results: poster session; feedback session (C) Discussion about results and fair testing: redoing/adjustments (C/G)	Design diary stage 3 Final research questions (C) Fair test rules of thumb (C) Laboratory notebook (G) Experiment poster (G)
Establishing Design Rules of Thumb [20-30 min]	Determination of design rules using experiment results (C) Focus on the science content: science vocabulary and concepts (C)	Design diary stage 4 Design rules of thumb (C)
Design Planning, Pin-Up Session [80-90 min]	Devise, share and discuss design solutions: divergent thinking (G) Poster: provisional design solution (G) Pin-up session (posters): feedback session (C) Adjusting and redoing until satisfied: final design solution (C/G)	Design diary stage 5 Design posters (G) Design sketch (G)
Construct & Test, Analyze & Explain, Gallery Walk [120-180 min]	Prototyping: realization of the design solution (G) Testing the design: realization of design specifications (G) Gallery walk: determine shortcomings; feedback/reflection (C) Adjustments of the design rules and design solutions (C/G)	Design diary stage 6 Prototype design (G)
Iterative Redesign [50-60 min]	Iteration of previous steps depending on decisions made (C/G) Improving the design (G) Final discussion about design solutions and scientific concepts (C)	Design diary stage 7 Final design solution (G) Final reflection (individual)

Note. Reprinted from Van Breukelen, Van Meel, et al. (2016). C = class activity or product; G = design group activity or product.
a Available resources: electronic learning environment (ELE), smartphones, laptops, tablets, Microsoft Office software, internet access, materials and tools for design realization, materials for conducting experiments
b Design diary (ELE-archived): reflections, feedback, process descriptions and pictures/movies. Bulleted lists are stage-specific.

Preliminary studies and task adjustments

The challenge discussed in the previous section was adjusted towards better concept learning for use in this study. As discussed the adjustments listed below mainly concerned explicit teaching and scaffolding strategies.

Backward design. The pre- and post-exam outcomes of the preliminary studies revealed that high gain question were strongly task-related and crucial for succeeding. Thus, detailed task analysis is important to unravel task-exposed and -underexposed concepts and to predict learning outcomes. Additional less directive concepts, complementing the knowledge domain, should be addressed otherwise (teacher-driven) through additional teaching interventions. This approach corresponds to the idea of backward design (Wiggins & McTighe, 2006) that states that education designers must begin to think about assessment and objectives before deciding what to do and how to teach. Regarding the solar challenge, initially designed for Study 2, there were four topics of underexposed science: (1) conceptual nature of resistance, (2) nature of electrical energy and energy dissipation, (3) behavior of current in components, (4) effect of voltage changes on circuit operation. To explore 1 through 3 students used simulation software and the fourth topic was addressed by additional experimentation. All topics, addressed as interludes, were complemented by class discussions and didactic analogies for clarification. Topic 1 and 3 were addressed after experimentation (stage 3) because experimentation contained resistance-current measurements. Topic 2 was addressed after design testing (stage 6) because then efficiency calculations took place. Topic 4 was addressed after the final stage by replacing the solar cells in the final design with a traditional voltage source.

Guided discussion. For teacher guidance during class discussion the technique of guided discussion (Carpenter, Fennema, & Franke, 1996) was used to highlight and explicate underlying science. When students worked in groups they were challenged to think and make sense of what they were doing. Then, by observing students' thinking and doing it became clear what individual students understood about science. Based on this, the teacher made notes about which students should present their insights during class-discussion. This might concern insights that are incorrect but useful to initiate a discussion of common misconceptions. Eventually, more sophisticated insights are used as input to head for proper reasoning and understanding. Both inputs and class discussion provide the teacher with information about students' knowledge and (existing) cognitive gaps, whereupon better understanding can be obtained.

Informed design. Informed design aims for enhancing students' prior knowledge through preparatory activities named knowledge and skill builders (KSBs) (Burghardt & Hacker, 2004). In this way, students are better prepared to approach design challenges from a more knowledgeable base and to tackle design problems by conceptual closure. Based on Study 2 an preparatory activity was created for this study around the behavior of solar cells. Students involved in Study 2 incorrectly assumed, without testing, that collar cells behave like (ideal) voltage sources. This assumption resulted in insignificant and time-consuming experimentation and finally trial-and-error behavior during design planning. To prevent this from happening students had to do, during stage 2, some information seeking accompanied by a class discussion regarding characteristics of (combined) solar cells.

Explicit instruction and scaffolding. According to Archer and Hughes (2011) explicit instruction is characterized by a series of supports or scaffolds, where students are guided through the learning process with clear statements about the purpose of and rationale for learning activities. It embraces 16 instructional elements that aim for a systematic method of teaching with emphasis on proceeding in small steps, checking for student understanding, and achieving active and successful participation by all students. LBD takes account of most of the elements and the adjustments mentioned before also fit into explicit instruction. However, teacher guidance should also fit the educational setting. Design challenges face teachers with a new kind of classroom control (Wendell, 2008) where teachers must relinquish directive control (Burghardt & Hacker, 2004). Thus, teachers need to develop pedagogical strategies for guiding complex design-based science tasks (Bamberger & Cahill, 2013). Study 2 that investigated these strategies resulted in a framework of important teaching guidelines that were directive for teacher handling during this study.

Adjustment of the design diary. Students involving Study 1 and 2 hackled the amount of administration (design diary) mainly because little administration was necessary to learn or move on. For example, there was a lot of requested reflection but in too few occasions this affected advancement directly. As a result reflection became disturbing and abortive. Therefore, administration was reduced and many written proceedings were replaced by process pictures accompanied by short subscriptions.

Methodology

21 first-year student teachers (science) took part in this design-based mixed methods study where they faced the improved solar challenge. All participants had prior experiences on characteristic LBD elements and

sufficient prior knowledge regarding the science domain. The challenge was guided by the principal investigator (teacher trainer) because of the relatively small number of participants. According to Crouch and McKenzie (2006) this offers the possibility to establish a sustainable relationship with participants and to provide added depth to the study, all resulting in an increased validity.

Quantitative data was collected to study students' progress in concept learning and video recordings were used to generate quantitative data about skill performances. Qualitative data was used to discover how task improvements affected concept learning by comparing students' comments to previous studies.

Data collection

To study a change in conceptual understanding the pre-post-exam developed for Study 2 was used. This multiple choice test is based on the validated Determining and Interpreting Resistance Electric Circuit Test (DIRECT) specially designed for use with high school and university students (Engelhardt & Beichner, 2004). The exam contains 46 items where each objective in Table 2 is served by multiple questions. Study 2 showed, by using a control group, that there was no task learning effect from just completing the test.

Study 1 showed that students mainly learned incomplete concepts and had difficulty in making proper knowledge connections and therefore did not achieve deeper conceptual understanding. This conclusion was partially based on multiple choice tests. According to Stoddart, Abrams, Gasper, and Canaday (2010) close-ended tests like this often fail to measure conceptual understanding because students easily can make guesses and therefore knowledge structures remain invisible. Using concept maps is suggested as a more meaningful way of assessing conceptual understanding. Therefore, beside multiple choice testing, students were asked to create a concept map before and after the challenge. For this, a proposition-based concept map test was developed, based on Yin, Vanides, Ruiz-Primo, Ayala, and Shavelson (2005), where students had to create 16 fundamental propositions (a connection between two concepts by using linking words or phrases) within a set of 10 predefined task-related concepts.

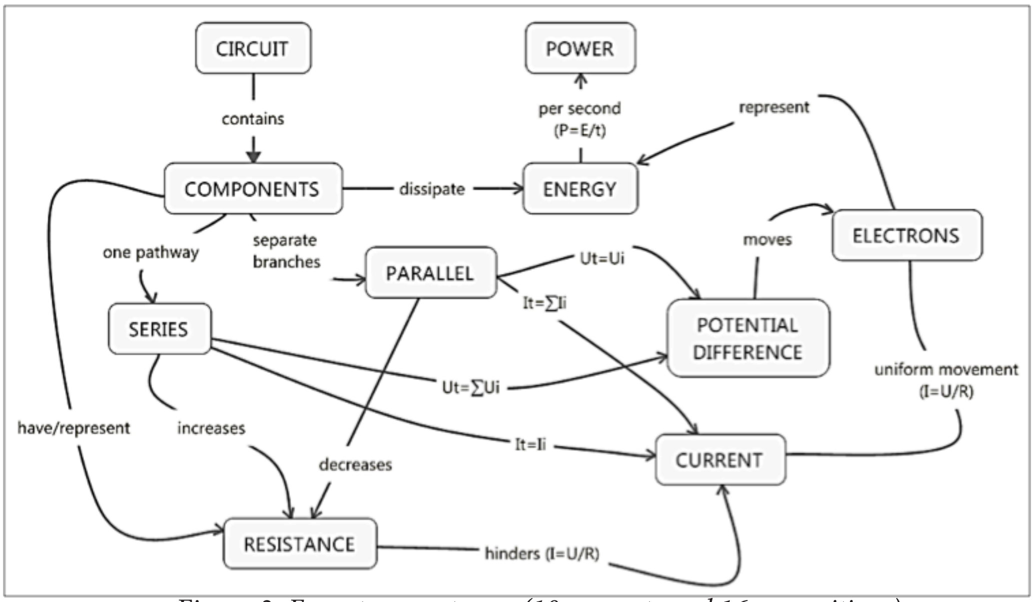

Figure 3. *Expert concept map (10 concepts and 16 propositions)*

The selection of concepts and the amount of propositions was based on a peer reviewed expert map (Figure 3) designed by two experts. According to literature, proposition-based concept map tests, based on an expert map, appear to be superior to other mapping strategies in assessing conceptual understanding (Cañas et al., 2003; Rye & Rubba, 2002). It is important to note that Yin et al. (2005) established small task learning effects in some cases due to the development of mapping skills. Those effects will be minimal for this study because students are familiar with mapping techniques.

To study an increase in students' skill performances we chose and slightly adapted the approach used in previous LBD research by Holbrook, Gray, Fasse, Camp, and Kolodner (2001) in order to make comparison possible. Students were videotaped when working, partially in groups, on similar performance tasks before and after the challenge. Tasks were taken from the Performance Assessments Links in Science Website database (SRI International Center for Technology in Learning, 1999) and were suitable for use with senior high school students (age 16-18); comparable to first-year student teachers in this study. During the pre-task students had to determine the power dissipated in a combination of two resistors connected in series to a battery. The post-task concerned the determination of how well different wires radiate heat when voltage is applied across each wire. Both tasks included three parts: (a) students designed an experiment or procedure for fair testing, (b) students ran an specified experiment and collected data, and (c) students analyzed the data to draw conclusions and make recommendations. The videotapes were analyzed, also according to Holbrook et al. (2001), on seven science-related dimensions (Table 4): negotiations during collaboration, distribution of efforts and tasks, attempted use of prior knowledge, adequacy of prior knowledge, scientific reasoning, experimentation skills and self-checks. Because the dimensions contain a mix of individual and collaboration skills each activity (a-c) started with an individual preparation, followed by a sharing session and ended with task completion by teamwork.

Afterwards an open-ended questionnaire was used to investigate students' views on which activities stimulated or impeded concept learning. Questions were based on the STARR method that provides a framework for reflection on learning outcomes (Verhagen, 2011). By interpreting students' answers, also in the light of preliminary studies, it is possible to establish whether the improvements are appreciated or room for improvement is left. Open-ended questions were used to prevent students' views from being swayed by possible answers. To confirm that the questionnaires' data reduction and interpretations were fair (respondent validation) nine students, the number that made themselves available, were interviewed simultaneously. During this session also some remarkable differences and correlations regarding learning outcomes were discussed for deeper understanding.

Analysis

The multiple choice tests were scored per student by the proportion of correct answers among 46 items. Proportions were used to calculate the gain (g): ratio of actual average gain (post – pre) to the maximum possible average gain (1 – pre) (Hake, 1998). A paired samples t test and a Wilcoxon signed-rank test were performed to investigate differences between pre- and post-scores. This combination was used because literature indicates that for relative small sample sizes using both tests increases the possibility to detect type I and II errors (Meek, Ozgur, & Dunning, 2007). Establishing Cronbach's alpha revealed the internal consistency of the exam.

The concept maps were scored per student. For this, all propositions (16 per concept map) were rated by two experts individually. Based on Yin et al. (2005) and Rye and Rubba (2002) the following scores were awarded: 3 points for a scientifically correct expert proposition (analogous to the expert map in Figure 4), 2 points for other correct propositions, 1 point for a weak or partially correct proposition and no points for incorrect propositions. Based on the experts' allocated scores the linear weighted Cohen's Kappa was calculated, which was sufficient. Then, the experts' average scores were assigned as final scores. Finally, the proportion scores (based on a 48 maximum score) were used, similar to multiple choice test analysis, to calculate gains and to investigate pre-post-score differences.

Analysis of the videotaped performance assessments took place by using a scoring rubric (See Appendix) where each performance dimension was served by a 5-point rating scale (1-5), with 5 being the highest level/score. Although the rubric's scale and dimensions are similar to that used by Holbrook et al. (2001) the level descriptors were adjusted for more validity. The original rubrics assessed skill performances by capturing the extent to which students in a group participated in practicing a skill: if more students were actively involved the group got a higher rating. According to Jonsson and Svingby (2007) this (possibly) causes validity problems because this method fails to reveal the quality of students' individual performances. Because a well-validated rubric, matching all our skill dimensions, was not available a rubric was created by combining existing rubrics. For this, we used an available qualitative framework of criteria to guaranty an acceptable level of validity, because a more sophisticated approach, achievable within this study, is still in its infancy (Baartman, Bastiaens, Kirschner, & Van der Vleuten, 2006; Moskal & Leydens, 2000). In short, rubrics compromising the following properties were selected: applicability to a 5-point scale, level descriptors based on observable behavior of individuals, univocal descriptors that actually reflect the skill dimension, some degree of validation, development based on experiences, expert involvement and suitability for the target group. Based on the final rubric two experts rated students' skill competences individually whereupon, after establishing an acceptable Cohen's Kappa, the experts' average ratings were assigned as final scores. Differences between pre- and post-ratings were also tackled by paired samples t tests and Wilcoxon signed-rank tests.

To investigate the strength of the relationships between pre- and post-assessment variables the Pearson product-moment correlation coefficient was computed for all possible combinations of variables. It is particularly interesting to find out how the multiple choice test and concept map test are correlated because they both concern conceptual knowledge. Furthermore, it reveals which skills strongly interact with concept learning.

For questionnaire analysis, at first to categorize responds, we distinguished between positive and negative opinions on the process of conceptual learning. After this, within these categories common themes were grouped and tagged by a description resulting in sub-categories of impeding or stimulating properties. Finally, all questionnaires were re-read to make sure all responses were categorized properly. During the group interview all properties were discussed and, based on students' input, slightly customized or filled up. Finally, remarkable differences and correlations regarding assessment outcomes were accompanied by a uniform group opinion on how to interpret results. For theoretical underpinning of students' opinions scientific literature was searched through.

Results

Table 4 gives a complete overview of all pre- and post-assessment results per student including mean scores and standard deviations. For the multiple choice test the average Cronbach's alpha, based on individual objectives in Table 2, is 0.72 for the pre-test and 0.69 for the post-test. The linear weighted Kappa values for the concept map and performance assessment analysis are shown in Table 5. Thus, in case of all assessments the reliability is sufficient.

The conceptual learning gains in Table 4 for the multiple choice test are significant, $t(20) = -30.87$; $p < 0.001$, just as for the concept map test, $t(20) = -24.58$; $p < 0.001$. This is confirmed by the Wilcoxon signed-rank test that gives the same p value for both tests. The mean gain for the multiple choice test (0.68) significantly increased compared to Study 2 (0.49) and Study 1 (0.35) and exceeds conceptual gains (LBD) found by Holbrook et al. (2001) that revealed mean gains up to 0.40. Compared to a large previous survey of pre-post-test multiple choice data for physics courses (Hake, 1998), that showed maximum gains between 0.60 and 0.70, our students were equally successful. Remarkably, the highest gains found by Hake (1998) resulted from interactive engagement (IE) methods designed, similar to LBD, to promote conceptual understanding through heads- and hands-on activities contributed by (peer) feedback, collaboration and intensive teacher guidance.

Table 4

Results pre- and post-assessments (n = 21): multiple choice test, concept map test, skill performances

Stud.	Multiple choice test			Concept map test			Performance assessment[b]													
	Proportion score			Proportion score			Negotiations		Distr.effort/ tasks		Prior know.use		Prior know. adeq		Scient.reason- ing		Experimenta- tion		Self-checks	
	Pre	Post	Gain[a]	Pre	Post	Gain[a]	Pre	Post	Pre	Post	Pre	Pre	Post	Pre	Post	Pre	Post	Post	Pre	Post
1	0.37	0.80	0.69	0.40	0,72	0,57	2,00	3,00	2,50	3,00	2,00	3,00	1,50	4,00	2,00	3,00	3,00	4,00	1,50	4,00
2	0.37	0.78	0.66	0.40	0,64	0,43	3,00	4,00	3,00	5,00	1,00	2,50	1,00	3,00	1,00	2,50	2,50	3,50	2,00	3,00
3	0.39	0.78	0.64	0.41	0,70	0,53	2,50	3,00	3,00	3,00	1,50	2,50	2,00	3,00	1,50	3,00	3,00	4,50	2,00	4,00
4	0.57	0.91	0.80	0.58	0,79	0,56	1,00	2,00	2,00	3,00	4,50	5,00	4,00	5,00	3,00	4,00	3,50	4,00	2,00	4,00
5	0.46	0.83	0.68	0.49	0,67	0,38	1,00	2,00	1,50	3,00	3,00	4,00	3,00	4,50	3,00	4,00	2,50	4,00	1,00	4,00
6	0.35	0.78	0.67	0.36	0,63	0,44	3,00	3,00	2,00	2,50	2,00	3,00	2,00	3,00	2,00	4,00	3,00	3,50	2,00	5,00
7	0.28	0.70	0.58	0.33	0,56	0,37	2,50	3,50	3,00	3,00	2,00	3,00	1,50	3,50	1,00	2,50	3,00	5,00	2,00	3,50
8	0.41	0.78	0.63	0.43	0,70	0,51	2,00	4,00	2,50	3,00	3,00	4,00	3,00	4,00	2,50	4,00	3,50	4,50	2,00	4,00
9	0.41	0.78	0.63	0.41	0,69	0,51	2,00	2,50	2,00	3,00	2,50	4,00	2,50	4,00	2,50	4,00	1,50	4,00	1,00	4,00
10	0.28	0.85	0.79	0.29	0,66	0,55	1,00	2,00	1,50	3,00	1,00	3,00	1,50	4,00	1,00	3,00	2,00	3,50	2,00	3,50
11	0.33	0.72	0.58	0.33	0,60	0,43	2,00	4,00	2,00	4,50	2,00	4,00	2,00	3,50	2,00	3,00	2,00	3,00	2,00	3,00
12	0.46	0.78	0.60	0.43	0,67	0,45	2,50	3,00	2,00	3,50	2,00	3,00	2,50	3,50	3,00	3,50	2,00	3,00	1,50	4,00
13	0.33	0.76	0.65	0.38	0,61	0,41	2,00	3,50	3,00	4,00	2,00	3,00	1,50	3,00	1,00	3,00	1,50	3,00	2,00	3,00
14	0.35	0.74	0.60	0.44	0,64	0,38	3,00	4,00	2,00	4,00	2,00	4,00	2,00	4,00	2,00	3,50	2,50	3,50	2,00	4,00
15	0.39	0.83	0.71	0.39	0,68	0,51	2,00	4,00	2,50	5,00	2,00	4,00	2,00	4,00	1,50	3,00	3,00	4,00	1,00	4,00
16	0.39	0.83	0.71	0.42	0,71	0,54	2,00	2,50	2,00	3,00	2,00	4,00	2,00	4,00	2,00	3,00	3,00	4,00	2,00	5,00
17	0.46	0.87	0.76	0.39	0,75	0,64	3,50	4,00	3,00	4,00	3,50	4,50	3,00	4,00	3,00	3,50	3,00	3,00	2,50	5,00
18	0.24	0.72	0.63	0.28	0,59	0,46	2,00	3,50	3,00	4,00	1,00	3,00	1,00	4,00	1,50	2,50	3,00	4,00	2,00	4,50
19	0.22	0.80	0.75	0.30	0,61	0,48	2,50	3,00	2,00	4,00	1,50	3,50	1,00	4,00	1,00	2,00	3,50	4,50	3,00	3,00
20	0.39	0.83	0.71	0.36	0,69	0,54	2,50	3,00	3,00	3,50	2,00	4,00	2,00	4,00	1,00	2,50	3,00	5,00	2,50	4,00
21	0.37	0.83	0.72	0.42	0,73	0,58	2,00	4,00	3,00	4,00	2,00	3,50	2,00	4,50	1,50	3,00	2,50	4,00	3,00	5,00
Mean (SD)	0.37 (0.08)	0.80 (0.05)	0.68 (0.07)	0.39 (0.07)	0.67 (0.06)	0.49 (0.07)	2.19 (0.66)	3.21 (0.72)	2.40 (0.54)	3.57 (0.71)	2.12 (0.84)	3.55 (0.67)	2.05 (0.76)	3.83 (0.53)	1.86 (0.74)	3.17 (0.60)	2.69 (0.60)	3.88 (0.61)	1.95 (0.55)	3.98 (0.66)

SD = standard deviation; Stud. = student

a Gain: (g) = (post%-pre%)/(100-pre%); where: high gain: (g) ≥ 0.70, medium gain: 0.70 > (g) ≥ 0.30, low gain: (g)< 0.30 (Hake, 1998)

b Mean scores awarded by two experts per skill dimension based on a 5-point scoring rubric (See Appendix), with 5 being the highest rating.

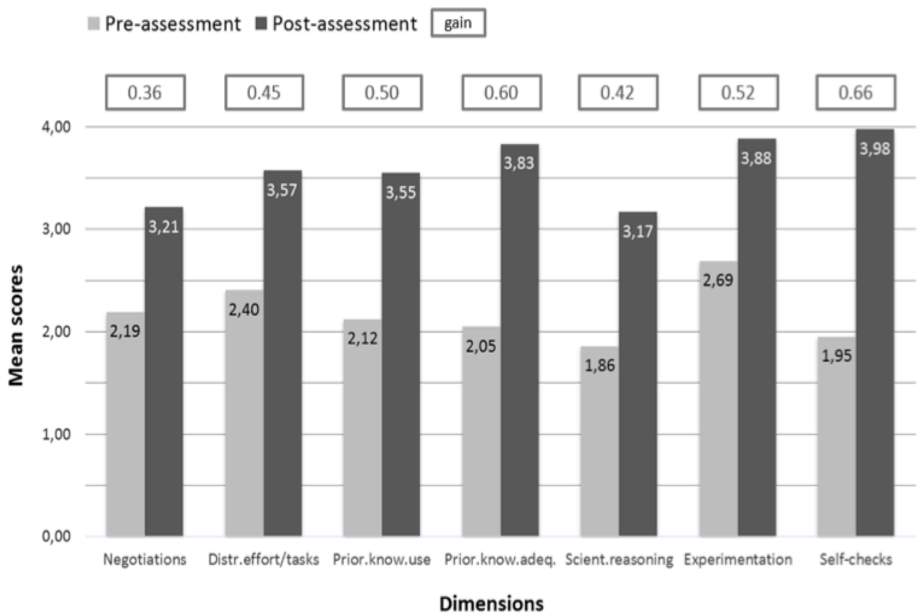

Figure 4. *Mean skill performances based on assessment results (Note: Dimensions and results are based on Table 4. Gain = (post-score – pre-score)/(100 – pre-score).*

Incidentally, a critical comment should be made because the concept map tests showed significantly lower ($p < 0.001$), but still substantial, gains (mean gain = 0.49; lowest gain = 0.37; highest gain = 0.64). This will be discussed in more detail later on.

Table 5

κw for concept map and performance assessment ratings

Test	kw pre-ratings	kw post-ratings
Concept map test	0.68 (lower limit = 0.62; upper limit = 0.74)	0.65 (lower limit = 0.58; upper limit = 0.72)
Performance assessment	0.62 (lower limit = 0.52; upper limit = 0.72)	0.66 (lower limit = 0.58; upper limit = 0.74)

Studying the performance assessment results in Table 4, shown graphically in Figure 4, it indicates that all skill dimensions show an increase in achievement level, where the highest progressions concern the adequacy of prior knowledge, experimentation skills and self-checks. However, all improvements are significant ($p < 0.001$) and fairly comparable to the performance assessment results found by Kolodner, Camp, et al. (2003). Those results showed scores between 2.00 and 3.00 for honors non-LBD students (the category befitting the students in our study at pre-testing) and scores up to 4.00 for typical LBD students (students exposed to LBD). Overall, students in this study reached, compared to previous LBD studies, much higher conceptual learning gainswhile advancement in skill performances was not hindered.

According to Table 6 there were strong (significant) positive correlations between the pre-scores of the multiple choice and concept map test, as well as for the post-scores. The gains of both tests showed a lower, but fair, correlation ($r = 0.683$, $n = 21$, $p < 0.01$) that can be explained by the fact that the mean gain for the concept map test was significant lower compared to the multiple choice test. According to Constantinou (2004) multiple choice test and concept map test results vary to a greater or lesser extent depending on which kind of learning is assessed through the multiple choice test (e.g. rote learning or meaningful learning). In general, Ruiz-Primo, Schulz, and Shavelson (1997) state that the correlation between both tests should be positive because they measure the same knowledge domain but the magnitude may differ. The interviewed students all agreed that the concept map test was more difficult because it stronger appealed to mastering well organized, relevant knowledge structures.

Furthermore, Table 6 shows moderate or strong positive correlations between the conceptual tests and three dimensions of the performance assessment (use and adequacy of prior knowledge and scientific reasoning) that also positively correlated with each other. Other positive or negative correlations between variables were not found or appeared to be weak or occasional. These findings correspond to previous findings: Schreiber, Theyßen, and Schecker (2016) found high correlations between conceptual tests and the preparation and evaluation of experiments, where prior knowledge and scientific reasoning are important, and low correlations with respect to conducting the experiment by following the rules for fair experimentation. Stone (2014) states that general skills, like collaboration and reflection, have a limited interconnectedness with more science-specific skills (practices) and the knowledge domain, where Zimmerman (2000) explicitly mentions the weak relation between conducting reception experiments and mastering conceptual knowledge and strong relations between conceptual knowledge, prior knowledge and scientific reasoning. All these insights perfectly reflect our findings where the interviewed students also emphasized the concept-free character (according to science knowledge) of collaboration, reflection and conducting a prescribed experiment. On the other hand students compared the use and adequacy of (prior) knowledge combined in combination with scientific reasoning to the mental activity important for creating a concept map, which reflects the mastering of knowledge structures.

Table 6
Positive Pearson product-moment correlations of pre- and post-assessment results (n = 21)

Variables		1	2	3	4	5
1. Multiple choice test	Pre	-				
	Post	-				
2. Concept map test	Pre	0.900**	-			
	Post	0.831**	-			
3. PA Prior knowl. use	Pre	0.790**	0.766**	-		
	Post	0.496*	0.500*	-		
4. PA Prior knowl. adequacy	Pre	0.883**	0.817**	0.920**	-	
	Post	0.568**	0.550**	0.725**	-	
5. PA Scientific reasoning	Pre	0.758**	0.689**	0.753**	0.834**	-
	Post	0.692**	0.426	0.447*	0.428*	-

Note: Based on pre-post-results in Table 4; PA = performance assessment; *p < 0.05; **p < 0.01 g

Table 7 shows the results of the questionnaire analysis where the amount of positive replies largely exceeds the negative ones. According to students, activities that directly appeal to underlying science (explicit teaching and experimentation) are invaluable for concept learning complemented by sufficient teacher and task guidance (feedback, clear instructions and transparency). These results are perfectly consistent with the results of Study 1 and 2. It is, however, surprising that in this study learning form peers is clearly more appreciated. Maybe because this study revealed less trial and error behavior and therefore students acted more like a role model or because the guided discussion approach, where the use of students' insights is directive, clarifies that peers are an important learning source. Although interviewed students seemed to confirm both statements a sloid validation failed to appear.

Table 7
Positive and negative opinions about concept learning (questionnaires; n = 21)

Influences on concept learning			
Stimulating factors (N = 106)	Perc.	Impeding factors (N = 44)	Perc.
Explication of underlying science by the teacher	28	Fragmentation of addressed science	36
(Results of) conducted experiments and simulations	20	Fragmentation of the task	29
Learning from peers: collaboration, sharing information, peer feedback	16	Uncertainty and uncomfortability (because of the new educational setting)	13
Teacher feedback during the task (no direct explication of science)	13	Too little attention to deeper assimilation of addressed science.	11
Clear instructions and transparency of tasks and objectives	12	Other (e.g. false information sharing, task duration)	11
Other (e.g. reflection, information seeking, the design context)	11		

Note: Perc. = relative distribution (%) of all replies within each category on which the corresponding sub-categories were distracted. Descriptions were redefined based on the group interview.

Taking the impeding factors into account fragmentation is an issue. First, students experienced too little coherence in addressed science and, second, students described the number of stages and accompanying administration as disruptive to the ongoing learning process. Also some students missed assimilation of addressed science for anchoring. Compared to the preliminary studies, the initial problems of addressing an incomplete science domain and a lack of science explication seem to be tackled. However, despite the improvements, coherence still is an issue and the amount of administration is still disruptive.

Discussion and implications

The adjustments, deduced from two preliminary studies, to enhance concept learning by design challenges appear to be successful because this study reveals a solid improvement of conceptual learning gains without reducing a positive effect on skill performances. Especially when the multiple choice test results (the assessment form used in all studies) are taken into account: gain-indices increased from the lower limit of medium (> 0.30) up to the very margin of high (0.70) where the latter is more or less reserved for the most successful physics-related courses (Hake, 1998). Students' responses show, which can be considered as an important reason for improved concept learning, that in contrast with the preliminary studies little comments were made about a lack of (explicit) science teaching. It seems that a combination of backward design, guided discussion and informed design is an appropriate remedy to enhance concept learning by extending strongly task-driven concepts and further deepening of all concepts. This happens, first, by introducing additional teacher-driven concepts (weakly task-driven) to complement the knowledge domain; important for understanding individual concepts. Second, by explicating and deepening all addressed science (explicit teaching). Figure 5 illustrates how contributions to conceptual learning gains may possibly collude where, of course, nearly always room for further improvement is left.

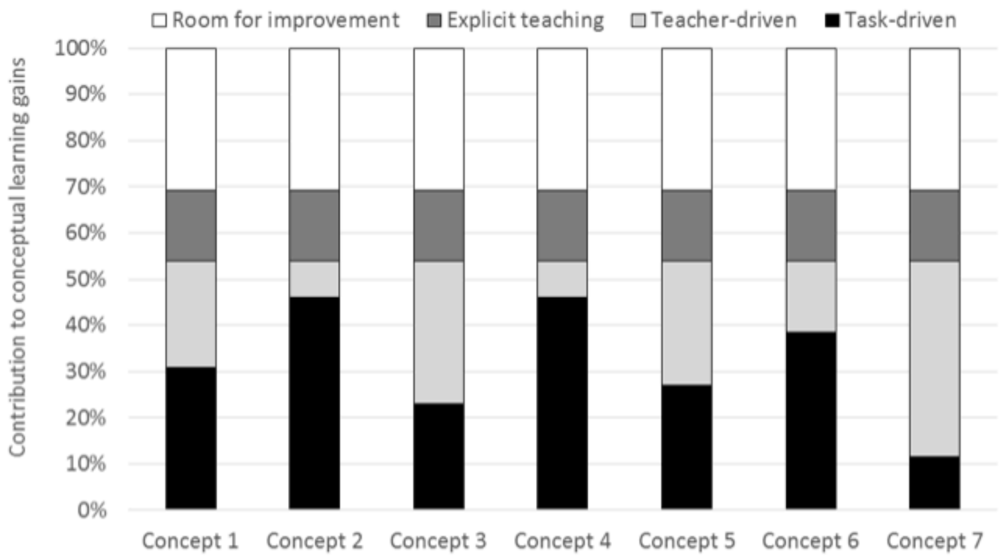

Figure 5. *Qualitative model for contribution to conceptual learning gains (Note: The (distribution of) percentages and amount of concepts are fictional and just used for illustration purposes where there is some resemblance to percentages found in the series of LBD studies. Concepts 1 through 7 represent the addressed knowledge domain.)*

This study provides some interesting clues where to search for further improvement. First, this study reveals significant positive correlations between students' conceptual performances, the use and adequacy of (prior) knowledge and scientific reasoning. Second, although students reached substantial conceptual learning gains the concept map test gains were significantly lower than multiple choice test gains. Third, students compared the use and adequacy of (prior) knowledge in combination with scientific reasoning to the mental activity important for creating concept maps. Fourth, students mentioned the fragmentation of addressed science and a lack of deeper assimilation of addressed science as important shortcomings. Thus, combining all four, more coherence of addressed science may be valuable because mastering explicit interrelationships between domain concepts enhances learning (Brandsford et al., 2003; Wiggins & McTighe, 2006). This may also improve the adequate use of knowledge and scientific reasoning and, with this, meaningful learning (important for concept mapping).

Beside fragmentation of addressed science, according to students, the same comment applies to the task itself. Although students experienced sufficient guidance and task transparency they described the number of stages and accompanying administration as disruptive to the ongoing learning process. Maybe a reduction of the number of (separate) stages and activities, through amalgamation, offers more coherence and less administration where guidance and scaffolding is shifted towards the ongoing process itself rather than breaking down into parts.

To conclude, both aspects of fragmentation, as discussed before, will be a main topic for further research. However, in general this study revealed some more interesting research themes. First, it is interesting to study the interaction between skill and concept learning in more detail because both types of learning are (partly) correlated and may strengthen each other. All the more because learning (STEM) skills is regarded as an important goal for modern education driven by a complex world economy that demands for those skills (ICF and Cedefop for the European Commission, 2015). Second, more insight is needed into the creation, use and validation of rubrics to assess skills. Third, correlations between multiple choice tests and concept map tests are often significant but widely spaced (Constantinou, 2004). Therefore it is necessary investigate this correlation in more detail and to find out how conceptual knowledge can be assessed properly depending upon the learning objectives.

Acknowledgement

This work was supported by the Netherlands Organisation for Scientific Research (NWO) under Grant number 023.001.030.

References

Abdul Gafoor, K., & Akhilesh, P. T. (2013). Strategies for facilitating conceptual change in school physics. *Researches and Innovations in Education*, *3*(1), 34-42.

Aikenhead, G. (2006). *Science education for everyday life. Evidence-based practice*. New York: Teachers College Press.

Archer, A. L., & Hughes, C. A. (2011). Exploring the foundations of explicit instruction. In K. R. Harris & S. Graham (Eds.), *Explicit Instruction: Effective and Efficient Teaching* (pp. 1-21). New York: The Guilford Press.

Baartman, L. K., Bastiaens, T. J., Kirschner, P. A., & Van der Vleuten, C. P. M. (2006). The wheel of competency assessment: Presenting quality criteria for Competency Assessment Programmes. *Studies in Educational Evaluation*, 32, 153-170.

Bamberger, Y. M., & Cahill, C. S. (2013). Teaching design in middle-school: Instructors' concerns and Scaffolding strategies. *Journal of Science Education and Technology*, *22*(2), 171-185.

Brandsford, J. D., Brown, A. L., Donovan, M. S., & Pellegrino, J. W. (2003). *How people learn: Brain, mind, experience, and school*. Washington, D.C.: National Academy Press.

Burghardt, M., & Hacker, M. (2004). Informed Design: A Contemporary Approach to Design Pedagogy as the Core Process in Technology. *Technology Teacher*, *64*(1), 6-8.

Bybee, R. W. (2013). *The case for stem education: Challenges and opportunities*. United States of America: National Science Teachers Association.

Cañas, A. J., Coffey, J. W., Carnot, M. J., Feltovich, P., Hoffman, R. R., Feltovich, J., & Novak, J. D. (2003). *A Summary of Literature Pertaining to the Use of Concept Mapping Techniques and Technologies for Education and Performance Support,* 1-108. Retrieved from http://cmap.bradercomm.net:8001/rid=1186721000045_1491728413_5664/IHMC%20Literature%20Review%20on%20Concept%20Mapping.pdf

Carpenter, T. P., Fennema, E., & Franke, M. L. (1996). Cognitive guided instruction: A knowledge base for reform in primary mathematics instruction. *The Elementery School Journal*, *97*(1), 3-20.

Chan, C. (2009). *Assessement: Practical experiment*. Retrieved April 2014, from University of Hong Kong http://ar.cetl.hku.hk/am_pe.htm

Cobern, W. W. (1994). *Worldview theory and conceptual change in science education*. Paper presented at the National Association for Research in Science Teaching, Anaheim.

Constantinou, C. (2004). Concept mapping for performance assessment in physics. In P. A. M. Kommers (Ed.), *Cognitive support for learning: Imagining the unknown* (pp. 155-166). The Netherlands: IOS Press.

Crouch, M., & McKenzie, H. (2006). The logic of small samples in interview-based qualitative research. *Social Science Information*, *45*(4), 483-499.

Engelhardt, P. V., & Beichner, R. J. (2004). Students' understanding of direct current resistive electrical circuits. *American Journal of Physics*, *72*(1), 98-115.

Franker, K. (2015). *Elementary teamwork rubric*. Retrieved April 2014, from University of Wisonsin - Stout https://www2.uwstout.edu/content/profdev/rubrics/elemteamworkrubric.html

Hake, R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics*, *66*(1), 64-74.

Holbrook, J. K., Gray, J., Fasse, B. B., Camp, P. J., & Kolodner, J. L. (2001). *Assessment and evaluation of the learning by design physical science unit, 1999-2000: A document in progress,* Retrieved from http://www.cc.gatech.edu/projects/lbd/htmlpubs/progress.html

ICF and Cedefop for the European Commission. (2015). EU Skills Panorama (2014) STEM skills Analytical Highlight, 1-5. Retrieved from http://skillspanorama.cedefop.europa.eu/sites/default/files/EUSP_AH_STEM_0.pdf

International Technology Education Association. (2007). *Standards for Technological Literacy: Content for the Study of Technology (3 ed.)*. Reston, Virginia.

Jonsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review*, *2*(2), 130-144.

Kolodner, J. L. (2002). Learning by design: Iterations of design challenges for better learning of science Skills. *Bulletin of the Japanese Cognitive Science Society*, *9*(3), 338-350.

Kolodner, J. L., Camp, P. J., Crismond, D., Fasse, B., Gray, J., Holbrook, J., . . . Ryan, M. (2003). Problem-based learning meets case-based reasoning in the middle-school science classroom: Putting learning by design into practice. *The Journal of the Learning Sciences*, *12*(4), 495-547.

Kolodner, J. L., Gray, J. T., & Fasse, B. B. (2003). Promoting transfer through case-based reasoning: Rituals and practices in learning by design classrooms. *Cognitive Science Quarterly*, *3*(2), 1-28.

Kolodner, J. L., Hmelo, C., & Narayanan, N. (1996). Problem-based learning meets case-based reasoning. *The Journal of the Learning Sciences*, *12*(4), 495-547.

Lustig, F., West, E., Martinez, B., Staszel, M., Borgato, M. T., Iosub, I., & Weber-Hüttenhoff, U. (2009). *Experiences and results from the European project 'Integrated subject science understanding in Europe'*. Paper presented at the ESERA conference, Istanbul.

Meek, G. E., Ozgur, C., & Dunning, K. (2007). Comparison of the t vs. Wilcoxon Signed-Rank test for likert scale data and small samples. *Journal of Modern Applied Statistical Methods*, *6*(1), 91-106.

Moskal, B. M., & Leydens, J. A. (2000). Scoring rubric development: validity and reliability. *Practical Assessment, Research & Evaluation*, *7*(10). Retrieved from http://PAREonline.net/getvn.asp?v=7&n=10

National Science and Technology Council. (2013). *Federal Science, Technology, Engineering, and Mathematics (STEM) education: 5-Year strategic plan*. Executive Office of the President, Washington, D.C.

Office of the Chief Scientist. (2013). *Science, Technology, Engineering and Mathematics in the national interest: A strategic approach*. Australian Government, Canberra.

Osborne, J., & Dillon, J. (2008). *Science education in Europe: Critical reflections*. Retrieved from http://www.nuffieldfoundation.org/science-education-europe.

Parliamentary Office of Science & Technology. (2013). *STEM education for 14-19 year olds*. Houses of Parliament, London.

Rennie, L., Venville, G., & Wallace, J. (2012). *Integrating Science, Technology, Engineering and Mathematics*. New York: Routledge.

Rhodes, T. L. (2010). *Scientific reasoning rubric.* Retrieved April 2014, from the Association of American Colleges and Universities https://www.aacu.org/sites/default/files/files/VALUE/VALUE_ScientificReasoningRubric.pdf

Rhodes, T. L., & Finley, A. (2013). *Using the VALUE rubrics for improvement of learning and authentic assessment*. Washington DC: Association of American Colleges and Universities.

Ruiz-Primo, M. A., Schulz, S. E., & Shavelson, R. J. (1997). *Concept map-based assessment in science: Two exploratory studies*. Retrieved from http://web.stanford.edu/dept/SUSE/SEAL/Reports_Papers/TECH436.pdf.

Rye, J. A., & Rubba, P. A. (2002). Scoring concept maps: An expert map-based scheme weighted for relationships. *School Science and Mathematics*, *102*(1), 33-44.

Schreiber, N., Theyßen, H., & Schecker, H. (2016). Process-oriented and product-oriented assessment of experimental skills in physics: A comparison. In N. Papadouris, A. Hadjigeorgiou, & P. C. Constantinou (Eds.), *Insights from research in science teaching and learning: Selected papers from the ESERA 2013 Conference* (pp. 29-43). Cham: Springer International Publishing.

Scott, D. (2008). *Critical essays on major curriculum theorists*. London: Routledge.

Sjöberg, S., & Schreiner, C. (2010). *The ROSE project: An overview and key findings*. Retrieved from http://roseproject.no/network/countries/norway/eng/nor-Sjoberg-Schreiner-overview-2010.pdf

SRI International Center for Technology in Learning. (1999). *Performance assessment links in science* (PALS). Retrieved Spring, 2015 http://pals.sri.com/tasks/tasks9-12.html

Stoddart, T., Abrams, R., Gasper, E., & Canaday, D. (2010). Concept maps as assessment in science inquiry learning – A report of methodology. *The International Journal of Science Education*, *22*(12), 1221-1246.

Stone, E. M. (2014). Guiding students to develop an understanding of scientific inquiry: A science skills approach to instruction and assessment. *CBE Life Sciences Education*, *13*(1), 90-101.

Van Breukelen, D. H. J., De Vries, M. J., & Schure, F. A. (2016). Concept learning by direct current design challenges in secondary education. *International Journal of Technology and Design Education*, 1-24.

Van Breukelen, D. H. J., Van Meel, A. M. D. M., & De Vries, M. J. (2016). Teaching strategies to promote concept learning by design challenges. Manuscript submitted for publication.

Verhagen, P. (2011). *Reflectie met de STARR-methode Kwaliteit met beleid*. Bussum: Coutinho.

Wendell, K. B. (2008). *The theoretical and empirical basis for design-based science instruction for children*. Unpublished Qualifying Paper. Tufts University.

Wiggins, G., & McTighe, J. (2006). *Understanding by design. Upper Saddle River*. New Jersey: Pearson Education Inc.

Yin, Y., Vanides, J., Ruiz-Primo, M. A., Ayala, C. C., & Shavelson, R. J. (2005). Comparison of two concept-mapping techniques: Implications for scoring, interpretation, and use. *Journal of Research in Science Teaching*, *42*(2), 166-184.

Zimmerman, C. (2000). *The development of scientific reasoning skills. Developmental Review*, *20*(1), 99-149.

Appendix: Scoring rubric for assessing skill dimensions

[1] Negotiations[a]				
1	2	3	4	5
Not at all	Rarely makes compromises to accomplish a common goal and has difficulty getting along with other group members.	Occasionally makes compromises to accomplish a common goal, and sometimes helps keep the group working well together.	Usually makes necessary compromises to accomplish a common goal.	Consistently makes necessary compromises to accomplish a common goal.
[2] Distributed efforts and tasks[a]				
1	2	3	4	5
Not at all	Performed little duties of assigned team role and contributed a little amount of knowledge, opinions, and skills to share with the team. Relied on others to do the work.	Performed a few duties of assigned team role and contributed a small amount of knowledge, opinions, and skills to share with the team. Completed some of the assigned work.	Performed nearly all duties of assigned team role and contributed knowledge, opinions, and skills to share with the team. Completed most of the assigned work.	Performed all duties of assigned team role and contributed knowledge, opinions, and skills to share with the team. Always did the assigned work.
[3] Using prior knowledge[b]				
1	2	3	4	5
Not at all	Identifies connections between prior events and experiences or prior concepts that relate to the context.	Identifies multiple prior experiences and events or prior concepts that relate to the context.	Prior events, experiences and concepts are identified and applied to the problem.	Prior events, experiences and concepts are routinely recalled that assist in problem solving.
[4] Adequacy of prior knowledge[b]				
1	2	3	4	5
Not at all	At least one of the mentions of prior knowledge is followed up on and is useful	A few of the mentions of prior knowledge were appropriate to the problem and used during the process.	Multiple mentions of prior knowledge were effectively selected and followed up on and were important for succeeding.	Nearly all mentions of prior knowledge are routinely followed up on and were important for succeeding and understanding.
[5] Use of science terms, scientific reasoning[c]				
1	2	3	4	5
Not at all	States ambiguous, illogical, or unsupported conclusions without proper use of science concepts. Demonstrates no ability to distinguish between causal and correlational relationships.	States general conclusions by using science terms, but beyond the scope of the inquiry findings limitations and implications. Demonstrates limited ability to distinguish between causal and correlational relationships.	States scientifically formulated conclusions focused solely on the inquiry findings. The conclusion arises specifically from and responds to the inquiry findings. Demonstrates appropriate ability to distinguish between causal and correlational relationships.	States a scientifically formulated conclusion that is a logical extrapolation from the inquiry findings limitations and implications. Demonstrates advanced ability to distinguish between causal and correlational relationships.

[6] Experimentation and fair testing[d]				
1	2	3	4	5
Not at all	Experimentation was done chaotically, without (proper) knowledge of how to control variables. At least one variable was tested.	Only some of the experimental tasks were done satisfactorily; some understanding of fair testing was present by controlling at least two (dependent) variables.	Most of the experimental tasks were done neatly and satisfactorily; an acceptable amount of understanding of fair testing and controlling for variables was present.	Experimental tasks were done in an organized and effective way by understanding for fair testing and controlling all variables.
[7] Self-checks[d]				
1	2	3	4	5
Not at all	A little (serious) self-checks took place without serious consideration for improvements.	A few self-checks took place by questioning several aspects of the procedure but were not properly addressed (without serious consideration)	(Some) self-checks took place by questioning several aspects of the procedure but were not optimally addressed.	Self-checks took place by questioning several aspects of the procedure and affected in improvements while performing the task

a Adapted from Franker (2015).
b Adapted from Rhodes and Finley (2013).
c Adapted from Rhodes (2010).
d Adapted from Chan (2009).

Journal of Research in STEM Education
ISSN:2149-8504 (online)
Vol 1, No 2, December 2015, PP 106-124

RESEARCH REPORT

Developing Effective STEM Animations: Application of a Multimedia Learning Theoretical Framework

Oludurotimi Adetunji [a1], Roger Levine[b]

[a]*The Science Center and Department of Physics, Brown University, USA*

[b]*Independent Consultant, Redwood City, USA*

Abstract: *Although most people believe animations can be very effective for STEM instruction and engagement, research often leads to findings that they are not superior to static graphics or other information presentations. The reason for these failures is not inherent in animations. Rather, the failures often reflect non-adherence to principles of good graphics design and a lack of understanding of principles of STEM learning. Sci-Toons, a series of animations dealing with diverse STEM topics, were developed based on a theoretical learning framework and the use of teams comprised of individuals with scientific expertise and individuals with visual design expertise. The process employed in the development of these animations is presented along with data on the wide-spread dissemination of Sci-Toons, its impacts on viewers, and its impacts on students involved in their production.*

Keywords: STEM Education, informal education, animations, videos, narratives, framework.

The need to produce more STEM graduates to maintain the national security and economic future of United States is well established (U.S. Office of Science and Technology Policy, 2012; U.S. Department of Commerce, 2012), as is the need to develop comprehensive strategies to increase the numbers of students in STEM pipelines and retain interested STEM students until graduation (Cohoon & Chen, 2003; Seymour & Hewitt, 1997; Zweben & Bizot, 2012). Women and underrepresented minorities, who now make up approximately 70% of all college undergraduates in the U.S., are historically underrepresented in STEM programs (National Science Foundation, 2013). A more diverse STEM workforce makes better use of our country's human resources and leads to more diverse points of view and more effective problem solving (Margolis, Estrella, Goode, Holme, & Nao, 2008; Melguizo & Wolniak, 2012; Page, 2007; Perna et al., 2009).

Approaches for diversifying and increasing the number of individuals entering universities to obtain STEM degrees include successful combinations of in-school and out-of-school STEM intervention programs that engage students from all backgrounds (Adetunji et al., 2012; Falk & Dierking, 2010; U.S. Office of Science and Technology Policy, 2010). Numerous reports also suggest that informal learning environments can strengthen science interest and aptitude on a national scale (National Governors Association, 2012; National Science Board, 2007; U.S. Department of Education, 2007).

One approach that most people believe can be very effective in engaging students from all backgrounds and stimulating interest in an area is the use of animations. Animations should be effective in portraying changes over time; they should be effective and engaging learning and teaching tools. However, research often leads to the findings that animations are not superior to static graphics or other presentations of information (Karlsson

[1] Corresponding author. The Science Center and Department of Physics, Brown University, E-mail: oludurotimi_adetunji@brown.edu

Adetunji, O., & Levine, R. (2015). Developing effective STEM animations: Application of a multimedia learning theoretical framework. *Journal of Research in STEM Education*, *1*(2), 106-124.

& Ivarsson, 2008; Kim, Yoon, Whang, Tversky, & Morrsion, 2007; Tversky, Heiser, & Morrison, 2013; Tversky, Morrison, & Betrancourt, 2002). The reasons for these failures is not inherent in animations. Rather, failures often reflect violations of cognitive principles of good graphics design.

According to the Congruence Principle, there should be a correspondence between the concepts portrayed and the information to be conveyed (Tversky, Heiser, & Morrison, 2013). For example, building a LEGO object is a continuous process. However, people think of the process as a series of discrete steps. Accordingly, LEGO instructions portray these steps -- and also change scale, size, and even perspective to more clearly map the task to the way the user thinks (Tversky et al., 2006). Similarly, people think of geographic travel in terms of turns at specific points. The classic map of the London Underground, designed in 1933, reflected this and emphasized the important information (the sequence of subway stops) and de-emphasized the unimportant information (distance and deviations from a straight line) (Agrawala, Li, & Berthouzoz, 2011).

Another cognitive design principle, the Apprehension Principle, states the importance of accurately perceiving and conceiving animations. The animations must be slow enough and clear enough to allow viewers to see and understand the concept being conveyed (Tversky, Morrison, & Betrancourt, 2002).

The Cognitive Theory of Multimedia Learning (CTML) provides guidelines for designing animations, based on cognitive principles (Mayer, 2005). Words and images are processed differently in the limited space comprising each channel's working memory. Each provides its own, separate contribution towards the creation of a mental model. An effective animation must recognize the cognitive load it creates and optimize the balance induced by different types of cognitive demand (called essential processing, extraneous processing, and generative processing).

Essential processing represents provision of the information necessary for understanding the constructs being demonstrated. Operationally, this can be accomplished through application of the pretraining principle (providing information before the animation, as through a glossary of terms) and the modality principle (presenting words aurally rather than visually). Extraneous processing can be reduced through adherence to the coherence principle (presenting only relevant aural and visual information; eliminating unnecessary distraction), the redundancy principle (avoiding narration that merely repeats prose presented visually); the signaling principle (providing a "road map" to delineate the organization of the material prior to presenting the material); the temporal contiguity principle (synchronizing narration with visual presentations); and the spatial contiguity principle (placing labels proximal to the elements they represent). Generative processing can be facilitated through application of the interactivity principle (providing opportunities for viewers to control elements of the animation) (Yue, Kim, Ogawa, Stark, & Kim, 2013).

Using the above principle, Yue et al., (2013) rated 860 randomly selected medical animations. Most employed at least one of these principles, but certain principles were rarely employed. Only 8% provided key terms prior to the animation; 17% used narration to enhance visual presentation, and only 19% employed the signaling principle. Conversely, nearly all (92%) placed labels proximal to images. The authors also pointed out the substantial influence of experts on the animations and the likelihood that things which are "intuitive" to the expert may not be knowledge possessed by novices, leading to the omission of steps crucial for a novice's understanding.

The process of creating animations can have strong and positive impacts on the animators as well as the audience. DiBlas, Paolini, & Sabiscu (2010) surveyed 153 kindergarten and primary school teachers who employed digital storytelling as a classroom activity about this practice's educational benefits. In comparison with "regular teaching", over 95% of these teachers felt student achievement was either "better" or "much better" with respect to interest in a subject matter, engagement, and deep understanding. Between 90 - 95% of the teachers noted improvement with respect to content organization skills, retention, technical abilities, communication abilities, and teamwork capacities. Prakash et al. (2009), working with a group of predominantly Native American and Alaska native elementary school children, demonstrated digital storytelling as an effective tool for engaging underrepresented minority students in science. And, Xu, Park, & Baek (2011) showed digital storytelling can be employed in virtual learning environments to enhance college students' writing self-efficacy.

Building on these principles and guidelines, the Multimedia Learning Theoretical Framework (MLTF) was developed. It begins with a model of the audience (Audience Pipeline Structure) to identify the ways to present information (stories) for the desired audience.

The Multimedia Learning Theoretical Framework (MLTF)

The MLTF is composed of three elements: (I) Audience Pipeline Structure, (II) Multimedia Learning Medium, and (III) Multimedia Learning Products. These are discussed below.

The Audience Pipeline Structure: STEM Literacy

The audience pipeline structure is an element of the MLTF that focuses on understanding the background of problem solvers, their preferred approach to problem solving and the development of a mechanism for engaging them in the learning process. Cognitive and neuroscientists have argued that there are different modes of thinking (Bruner, 1986; Gazzaniga, 2005; Kahneman 2003; Kahneman, 2011; Sperry, 1961). Bruner, for example, argues that there are two ways of thinking: a paradigmatic (logico-scientific) mode and a narrative mode. Figure 1 is a schematic diagram of the Audience Pipeline Structure, incorporating these different problem-solving approaches.

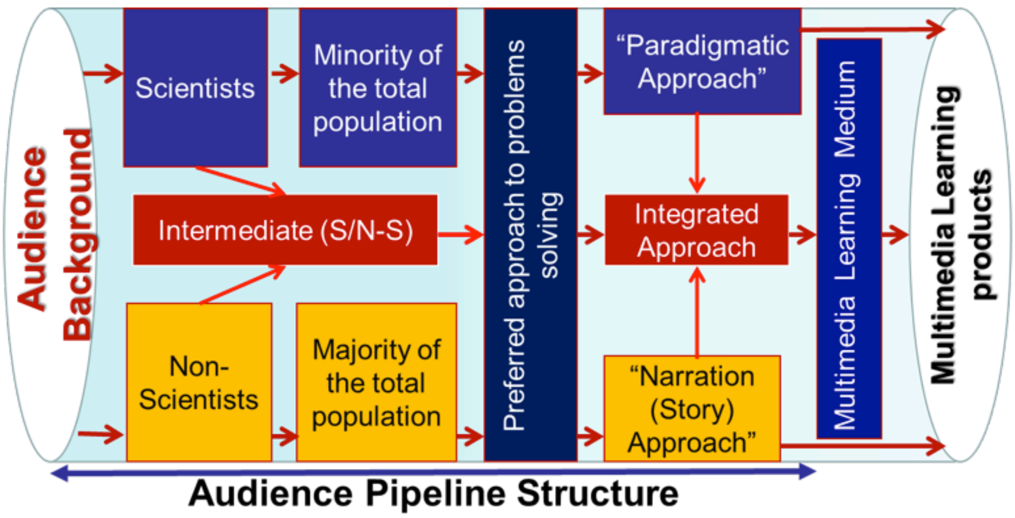

Figure 1. *Multimedia Learning Theoretical Framework (Model)*

A simple division of the audiences into scientists and non-scientists does not take into account the potential diversity of thinking styles within disciplines or the fact that a single individual can have varying roles, approaches, and levels of expertise in different contexts. This Audience Pipeline Structure makes room for an intermediate audience between scientists and non-scientists, explicitly recognizing a variety of problem-solving styles and experiences. An intermediate audience approach to problem solving is dubbed an integrated mode or approach. The integrated approach combines both the narrative and paradigmatic modes to accommodate broad thinking styles, accessible by a wide range of audiences from diverse backgrounds.

For example, a narrative approach video on "Climate Change" would be comprised of stories and pictures of the impacts of climate change along with stories and pictures of fossil fuel power-generating plants and automobiles spewing greenhouse gases into the atmosphere. A similar video, using the paradigmatic approach, would provide graphic data on climate characteristics (such as temperature and precipitation) and potential climate forcing mechanisms (such as atmospheric carbon dioxide concentrations, solar radiation, volcanism, orbital variations, etc.) over time. An intermediate approach would tell a story (or set of stories), presenting the underlying hypotheses and data (as appropriate), using pictures, videos, and other graphics.

Multimedia Learning Medium

Multimedia allows both learners and experts to construct a coherent integrated representation of a domain content's narrative, visual and verbal representations. This medium also allows smooth integration of the Audience Pipeline Structure with research on multimedia, such as instructional methods in which visuals are added to a verbal explanation so as to foster deeper understanding when learners mentally connect the verbal and pictorial representations (Mayer, 2005; Yue, Kim, Ogawa, Stark, & Kim, 2013). Animations are developed and produced by Sci-Toon Creation Group (SCG) members who learn about (and employ) design principles and issues important in the design of multimedia (dual channels, limited capacity and active learning) and evidence-based principles for the design of multimedia learning environments (Mayer, 2008; Tversky, Heiser & Morrison, 2013).

Multimedia Learning Products

This component of the Multimedia Learning Theoretical Framework is designed to convert the developed domain content narratives to storyboard and final multimedia products. Media such as animation, artistic renderings and/or video creations are used to develop multimedia learning products for engaging broad audiences in science.

Application of Multimedia Learning Theoretical Framework

Sci-Toons targets an intermediate audience. Accordingly, the Sci-Toon Creation Group (SCG) is also "intermediate", meaning it is composed of experts and novices in STEM fields. In other words, the composition of a SCG includes at least one STEM major, one faculty member and two non-STEM majors. The SCG used the MLTF's integrated approach for problem solving through interactive engagement by SCG members and included iterative feedback on Sci-Toon scripts and storyboards. The iterative feedback mechanism allows domain STEM content experts and novices who are members of the SCG to provide constant feedback during each step of the Sci-Toon development (script, storyboards and final animations). As the script is being developed, the STEM domain expert(s) contributes to the script by overseeing the overall accuracy of the scientific content. As the script undergoes iterative reviews, the narrative aspect of the integrated approach is in focus. Thus, the final script and storyboards incorporate the paradigmatic and narrative contributions from STEM experts and non-STEM expert/novices. See Figure 2 below.

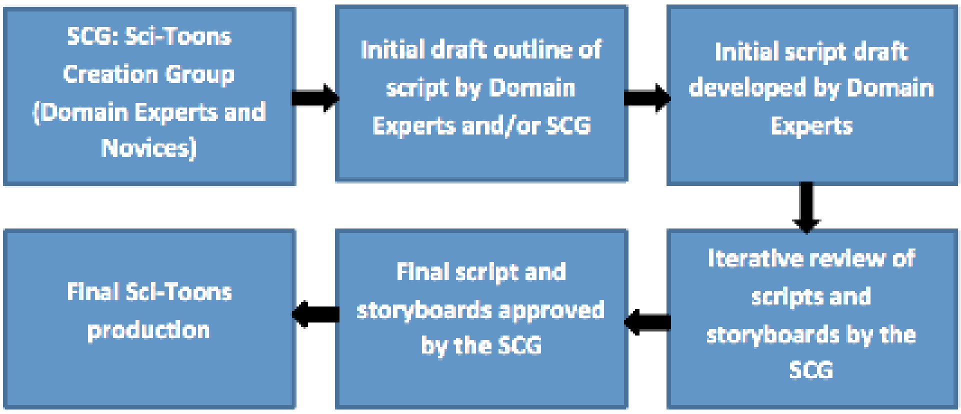

Figure 2. *Diagram of the Sci-Toon Creation Process*

This integrated approach is unique among existing online resources because it harnesses the contribution of individuals from a wide range of backgrounds in development of a final product. Table 1 shows how Sci-Toons differs from the ways other online resources are developed.

Table 1:
Sci-Toons versus Other Online Resources

Other Resources: EdX, Khan Academy, One-Minute Physics and MOOCs	Sci-Toons
Developed using paradigmatic approach	Developed using narrative or integration approach
Developed by domain content experts	Developed by domain content experts and novices
Might contain significant technical jargon	Technical jargon is limited
Emphasis is on getting the domain content right	Emphasis is on getting both domain content right and language at the level of the target audience
Content scripts are reviewed only by domain experts	Content scripts are reviewed by both domain experts and novices

The Sci-Toons model for learning (Adetunji, 2015) starts with concepts familiar and important to the SCG members and then gradually introduces more advanced scientific concepts through script development, storyboard creation and visual media. SCG members expand their understanding of scientific concepts or scientific research by examining content domain and creating stories that they animate. They learn as they explore, script, storyboard, and animate their chosen topic: non-STEM experts learn the science while STEM experts learn how to explain. Viewers (the target audience) expand their understanding of scientific concepts by following the explanations of new learners – that is, the new learners that make up the SCG.

Science Cartoon (Sci-Toon) Projects

The application of MLTF to each of the three key steps (scripting, storyboarding, and producing the animation) to Sci-Toons video development begins with selection of the SCG members. There is diversity in their academic backgrounds: A typical SCG includes STEM and non-STEM students and faculty. The SCG members begin by deciding on an area of scientific research or on a concept that the group will explore. Then, a subset of the group is charged with the development of the initial script based on this topic. The initial draft is then reviewed by the entire group. Each SCG member brings his or her unique perspective and provides feedback on the script, which goes through several iterations of reviews and edits before it is finalized. A finalized script usually has a theme and a hook (that is, something to capture or maintain the audience's interest) for the story. The script is written in language that a broad audience can understand while keeping the core of the scientific research or concept accurate.

In all of the key steps of Sci-Toon development, the SCG members focus on developing materials for an intermediate audience. This is facilitated by the SCG's diverse problem-solving styles and experiences. In addition, Sci-Toon development employs the intermediate audience's preferred approach to problem solving by incorporating both paradigmatic and narrative approaches. A description of how MLTF was applied for the production of three different Sci-Toon videos is described below.

Weather and Climate (https://www.youtube.com/watch?v=UC38Rf70px8)

During the initial meeting, when the SCG members decided to pick Weather and Climate as a topic, it was clear that time scale was an important distinguishing factor. Meteorologists focus on a very short time scale compared with the much longer time scales of interest to climatologists. The SCG developed a hook for the body of the story: When going on vacation, check the climate to decide what to pack and check the weather to decide what you should be wearing when you arrive.

The main points of the story evolve around what Meteorology and Climatology are and how they are different. For Meteorology, tools such as thermometers, barometers and anemometers are used to collect temperature, air pressure and wind speed data. These atmospheric data conditions are gathered into supercomputers that use forecasting models to predict conditions in the near future. The representations use “numerical forecast equations” to create models of the atmosphere. Meteorologists take into consideration previous weather patterns and knowledge of how the different aspects of the atmosphere interact with each other and can be used to predict weather patterns, though only for a short period into the future. The farther into the future, the greater chance of error in the prediction. In contrast, climatologists study weather systems long term, over years or millennia, tracking trends and changes over time and predicting averages and changes in atmospheric conditions over years. Both Meteorology and Climatology use radar, satellite data, and sophisticated computer models for weather forecasting or climate models predictions.

In order to demonstrate the iterative nature of script development, we used text visualization software (McNaught & Lam, 2010) to enable comparisons of initial and final versions of scripts. Word visualization software displays the frequency with which different words appear in a document (ignoring common words, such as “the”, “an”, “on”, etc.) by increasing the size of a word in proportion to its occurrence. The following figure, showing word clouds for the initial and final scripts used for Sci-Toons were produced using TagCrowd (TagCrowd, 2015).

Figure 3. *Visualization of the top 50 words of the initial Weather and Climate script (left) and of the top 50 words of the final script (right)*

In the initial Climate and Weather Sci-Toon, the most frequent word was “change”, appearing 23 times. Four of these uses were in the context of “climate change”, reflecting the content experts’ desire to incorporate this important, but extraneous, concept in the video. However, in the final script, the word “change” was not among the 50 most common words and the phrase “climate change” did not appear. Terms such as models, planning, predictions, atmosphere, and forecasting increased dramatically in their prevalence, reflecting the increased focus on these terms and concepts for telling a story in the final version. (See Figure 3.)

How Do We See Color? (https://www.youtube.com/watch?v=pvC9MQvqHMQ)

The initial topic for this Sci-Toon was color. It quickly became clear that color is a broad topic that cuts across multiple disciplines including physics, chemistry, neuroscience, and the visual arts. Each of these fields has contributed to the understanding of color from different perspectives. The SCG finally decided to focus on how color is perceived. Their hook focused on the different ways in which color is used: Artists use different colors to express emotion, marketers attach colors to their brands to make them more recognizable and animals use colors to try to avoid predation.

Color was defined as the limited range of light that humans can see, which, in turn, introduced the concept that color was based on the wavelength of light reflected by an object. The Sci-Toon then focused on how visible color or light is perceived by our eyes and the role of different parts of the eyes including the cornea, pupil, retina, ganglion cells and the light sensitive cells known as rods and cones. Rods are responsible for our perception of light and dark and our peripheral vision. Cones allow us to see color. They are found in the center of the retina, where light is focused. In humans, cones come in three varieties, each sensitive to a different range of light.

The final part of the story focused on other animals' perception of color. Animals have different kinds and number of cones, which allow them to see the world differently from humans. Dogs are more limited in the wavelengths they can see because they only have two types of cones, leaving them "colorblind" to differences between red and green. Butterflies have four types of cones in their eyes, letting them see ultraviolet light. The animal with the best color vision might be the mantis shrimp, which has 12 different kinds of cones.

The evolution of the Sci-Toon, from initial script to final script, is presented as a word cloud (Figure 4). One can see the decreased prevalence of terms relating to the physical properties and characteristics of light (reflectance, newton, prism) that characterized content experts' classical presentations of color and the increased prevalence of certain words, such as cells, cones, retina, and rods in the final How Do We See Color? Sci-Toon, reflecting the increased importance of biological concepts in the story being told.

Figure 4. *Visualization of the top 50 words of the initial How do we see color? script (left) and of the top 50 words of the final script (right)*

The Conductive Polymer (https://www.youtube.com/watch?v=UjMbwS0LOkU)

The goal of this Sci-Toon was very clear from the onset. The SCG wanted to develop a script about the discovery of conductive polymers that would be very engaging to any audience. The initial script started with a description of electric current and its usage in everyday gadgets such as mobile phones, tablets and laptop computers. The script continued with the history of the accidental discovery of a polymer, polyacetylene, which exhibited metallic properties and the mechanism by which this polymer conducts electricity. As the initial script was being revised by the SCG members, there was a general consensus that it lacked a compelling hook. It was decided that the fact that certain non-metallic objects, such as some plastics, could be good conductors of electricity would be surprising and engaging to the audience and would serve as the hook.

Figure 5 shows the evolution of the Conductive Polymer script. In contrast to the other word clouds presented, the basic theme and presentation in the initial and final scripts were quite similar. Unlike the production of the other Sci-Toons, the SCG members working on the project discussed the project extensively, clearly articulating the goal and concepts to be covered before beginning to write the initial script. The Conductive Polymer script was finalized after the second iteration of reviews. A slightly different process was employed for the other two Sci-Toons, for which the SCG members initially chose the topic and then designated

1-2 members to write an initial draft. The drafts were then reviewed and edited by the all participating SCG member and went through at least five iterations of reviews before the script was finalized.

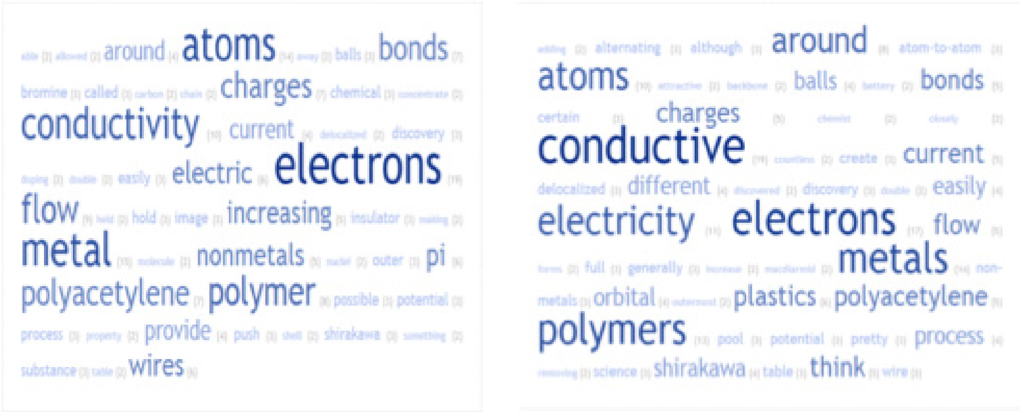

Figure 5. *Visualization of the top 50 words of the initial Conductive Polymer script (left) and of the top 50 words of the final script (right)*

Methodology

Viewers' demographics were measured through use of Google Analytics (www.google.com/ analytics/), based on information gathered from logged-in users on all electronic devices.

In order to assess the impacts of participation in a SCG on college students, a brief, five-minute electronic survey was developed and emailed in December, 2014, to all 23 students who had been involved with Sci-Toons prior to the Fall 2014 semester. A copy of the survey is included as Appendix 1. Two reminder emails were sent to non-respondents, two and four weeks after initial survey distribution. Completed surveys were received from 10 of the respondents, for a response rate of 43%. The survey and data collection procedures were determined to be exempt from a need for review by Brown University's Institutional Review Board. Informed consent was obtained from all individual participants included in the study.

Impacts of viewing the Conductive Polymer video were assessed through a brief survey of students participating in the Brown Science Prep (BSP) Program. BSP is designed to engage public high school students in science. Students come to Brown on a weekly basis to do science experiments and demonstration-based lessons in an informal environment. Thirty-four students viewed the video and completed a voluntary electronic survey after viewing the Sci-Toon.

Results

We begin by providing demographic data about individuals downloading Sci-Toons videos and information about average view times. These are followed by data provided by college students who were part of the Sci-Toons creation groups (SCG) and who responded to our survey. They were involved in the production of between one to four different Sci-Toons. Finally, data provided by high schools students who viewed the Conductive Polymer Sci-Toon are presented.

Viewer demographic and download data

All the three Sci-Toons described in this paper have been viewed multiple times and downloaded in many different countries. Table 2 shows the view counts and the number of countries where the videos have been downloaded since they were uploaded (publication date) to the Sci-Toon YouTube Channel. To our surprise, in a 13-day period, from May 19, 2015 to May 31, 2015, the Color Sci-Toon video went viral: the number of downloads increased by 45,888 (from 8,761 to 54,649). Over the same period, the number of downloads of the Conductive Polymer Sci-Toon only increased by 430 (from 7,253 to 7,683).

Table 2.
View counts and Number of Countries where the Conductive Polymer, Color, and Climate and Weather Sci-Toons have been viewed (as of May 31, 2015)

Sci-Toon	Publication Date	Number of Views	Number of Countries
Conductive Polymer	October, 2013	7,683	109
Color	April, 2014	54,649	179
Climate and Weather	January, 2015	347	27

Demographics of viewers of the Conductive Polymers, the Color, and the Weather and Climate Sci-Toons are presented in Figures 6 -8. Overall, the majority of the Conductive Polymer Sci-Toon were male (75.5% male vs. 24.5% female); nearly half (47.1%) were between 13 -24 years of age). The Color Sci-Toon was more likely to be viewed by females than males (38.1% vs 61.9%, respectively). The Color Sci-Toon viewers were also younger, with 62.2% between 13 - 24 years old and 37.8% aged 25 years or older. The Weather and Climate viewers were equally likely to be female (51.6%) or male (48.4%). Like the Conductive Polymers viewers, about half (51.6%) were between 13 - 24 years of age.

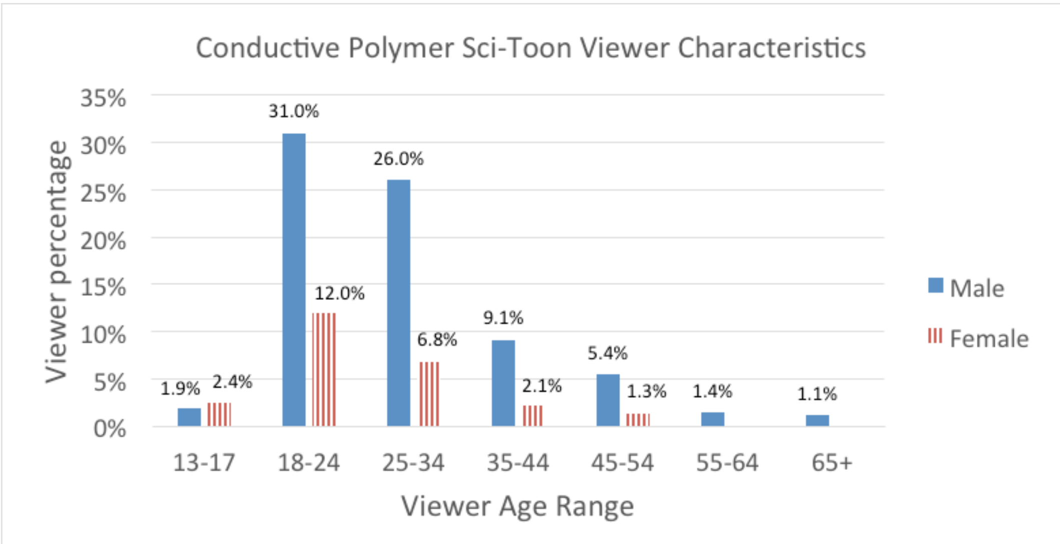

Figure 6. *Age Range and Gender of Conductive Polymer Sci-Toon viewers*

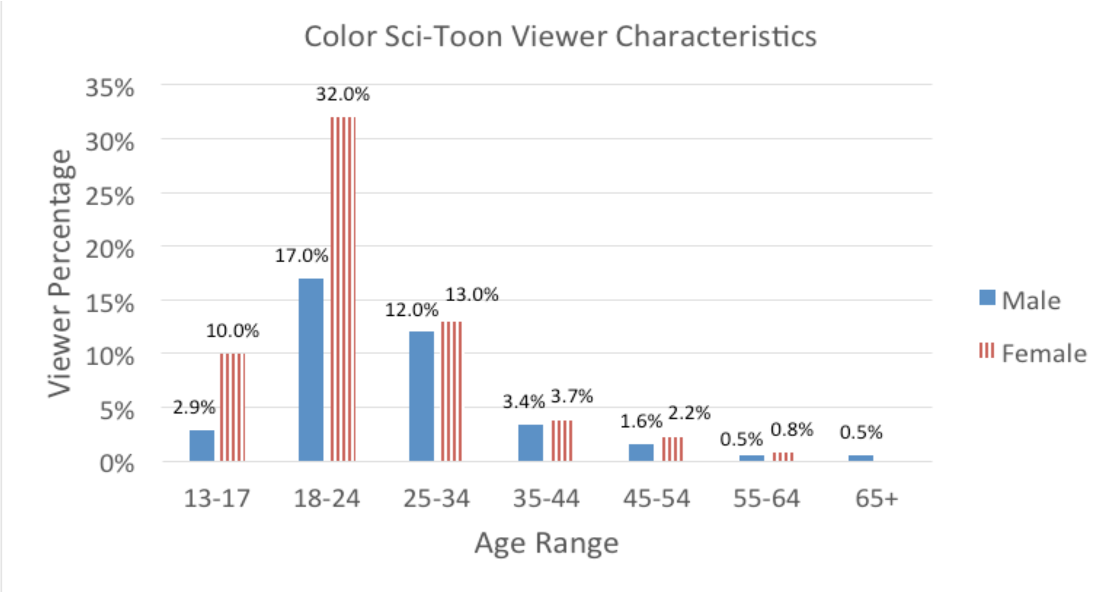

Figure 7. *Age Range and Gender of Color Sci-Toon viewers*

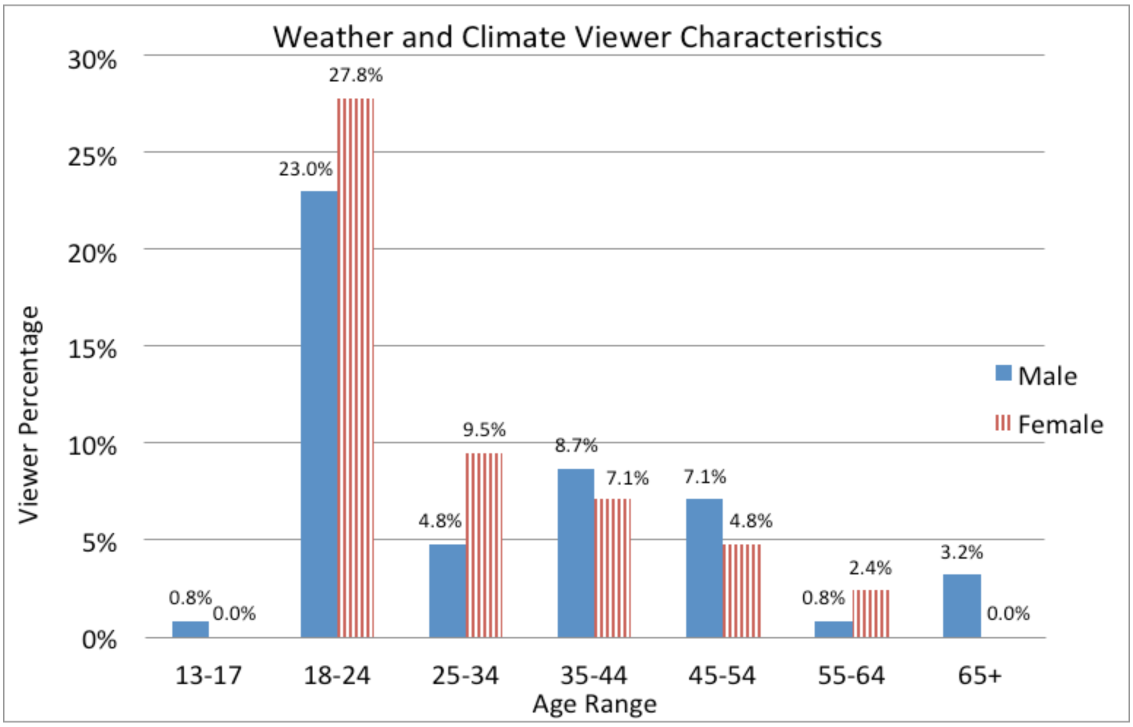

Figure 8. *Age Range and Gender of Weather and Climate Sci-Toon viewers*

The viewer demographic data do not necessarily reflect the numbers of individuals viewing a complete Sci-Toon. Some viewers may download the video and view later; others will watch and log off before the end of the video. For example, the Conductive Polymer Sci-Toon had an average viewing time was 3.49 minutes -- or about 57% of the length of the Sci-Toon.

Viewer demographic and download data

The survey completed by students who were SCG members included items asking them to evaluate how their participation in the Sci-Toons project affected their abilities to understand contemporary science issues and to explain scientific principles and issues to others. All of the students indicated positive changes, with the majority of respondents indicating these abilities were moderately improved. (See Figure 9.)

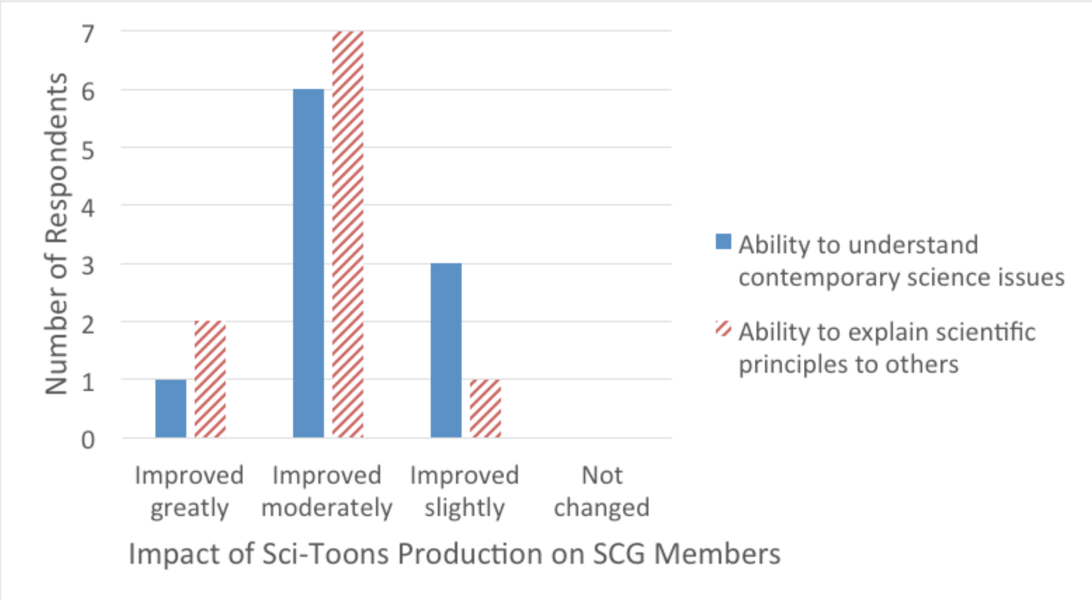

Figure 9. *Self-assessed Impacts of Participation in SCGs on College Students*

Respondents were also asked how important they felt it was for the general public to understand science and how difficult it was for the general public to understand science, both before and after their Sci-Toons experience. Overall, respondents reported non-statistically significant increases in the importance of the public understanding science and decreases in the public's perceived difficulty in understanding science. See Table 3.

Table 3.
Self-reported Changes in SCG College Students' Attitudes about Public Science Understanding

Item	Before	After	Change
How important is it for the general public to be able to understand science*	3.80	4.30	+0.50
How difficult is it for the general public to understand science**	2.56	3.00	+0.44

*: Not important=1; Slightly important=2; Moderately important=3; Very important=4; Extremely important=5.
**: Extremely difficult=1; Very difficult=2; Moderately difficult=3; Moderately easy=4; Very easy=5; Extremely easy=6.

Half of the respondents were science majors; half were not. All of the non-science major respondents reported that their Sci-Toons experience either slightly or moderately increased their likelihood of pursuing a STEM-related career. All of the science majors reported no change in the likelihood of their pursuit of a STEM-related career.

Impacts on Viewers

Thirty-four students participating in the Brown Science Prep Program viewed the Conductive Polymers Sci-Toon and completed a brief, voluntary survey afterwards. This survey included items asking students to indicate their level of agreement with various statements about the Sci-Toon.

Results are presented in Table 4. A large majority of the students indicated they felt the Sci-Toon was interesting (94.1%) and easy to understand (88.2%). Most also felt that they learned a lot (88.2%) and that watching the Sci-Toon was a good use of their time (88.2%). Nearly all would recommend it to friends who like science (94.1%). However, less than half would recommend it to friends who do not like science (44.1%).

Table 4.
Agreement with Statements about the Conductive Polymer Sci-Toon by High School Students

	Rating*	% Agreement**
I would recommend watching Sci-Toons to friends who like science	3.21	94.1%
The Sci-Toon was interesting	3.12	94.1%
I learned a lot from the Sci-Toon	3.06	88.2%
The Sci-Toon was easy to understand	3.00	85.3%
Watching the Sci-Toon was a good use of my time	2.97	88.2%
I would recommend watching Sci-Toons to friends who do not like science	2.35	44.1%

*: 1 = Strongly disagree; 2 = Disagree; 3 = Agree; 4 = Strongly agree
**: Percentage of students either agreeing or strongly agreeing with the statement.

Discussion

The Multimedia Learning Theoretical Framework (MLTF) model can easily be adopted for use in both formal and informal academic settings. As an experimental approach to teaching and learning, it allows both STEM and non-STEM learners along with experts to work together as domain concepts are being presented. The process is intended to expand understanding of the domain concept for both the learners and the experts as they engage in the iterative process of script development. The experts are exposed to a new way of thinking about the domain concepts as they work in collaboration with the learners within the integrated approach of the MLTF's Audience pipeline structure.

Tversky et al. (2006) noted the need for collaboration between graphic designers and domain experts in order to create effective visualizations. They also noted the practical difficulties of such collaborations. The MLTF model was applied as the basis of Sci-Toons to help overcome these difficulties and the problems that have prevented animations from reaching their potential as tools for STEM education and approaches for engaging students and the general public with STEM. This model accommodates broad thinking styles, accessible by a wide range of audiences from diverse backgrounds. It is applicable to secondary, intermediate school, and even primary school students, for whom digital story-telling has been shown to be an effective means of engaging children with science (DiBlas, Paolin, & Sabiescu, 2010; Prakash et al., 2009).

The prevalence of downloads is suggestive of wide use. In addition, the demographics of downloaders suggest different audiences are being reached by the different Sci-Toons. The Color Sci-Toon was more likely to be viewed by females; the Conductive Polymers Sci-Toons, more likely to be downloaded by males. Assuming that the 13 - 17 years olds are primarily secondary school students, about 4% of the Conductive Polymer Sci-Toon viewers and 13% of the Color Sci-Toon viewers were secondary school students. Similarly, it is reasonable to assume that post-secondary students make up a substantial proportion of the 18 - 24 year olds, who comprised 43% of the Conductive Polymer Sci-Toon, 49% of the Color Sci-Toon, and 51% of the Weather and Climate Sci-Toon viewers.

We can only speculate on reasons that the Color Sci-Toon's popularity skyrocketed in the last half of May, 2015. These increases occurred across all age categories but were most pronounced among the younger viewers (an eight-fold increase for 13 - 17 year olds and nearly a ten-fold increase for 18 - 24 year olds). Overall, male viewers increased by a factor of 4.5; female viewers increased by a factor of 7.6. The Sci-Toons were roughly of comparable length. Viewing time for the Climate Sci-Toon is 3 minutes, 56 seconds; for the Color Sci-Toon, 4 minutes, 3 seconds; and for the Conductive Polymer Sci-Toon, 6 minutes, 4 seconds. We noticed that the Color Sci-Toon, which is formally listed as How Do We See Color?, is the only Sci-Toon of the eight that have been published whose title is a question. Perhaps individuals who choose questions as search terms would be more likely to find and download a video with a matching title -- but this does not necessarily explain the sudden increase in popularity.

Impacts were assessed through a survey of students directly involved with the production of Sci-Toons. We feel post-participation surveys provided suggestive evidence of impacts, particularly on attitudes towards STEM and STEM careers for non-STEM majors. We anticipate administering pre- and post-participation surveys to provide stronger evidence of positive impacts. For assessing impacts on viewers, data from a survey of high school students were positive. Their agreement with a statement that they learned a lot from the Sci-Toon is suggestive of the animation's positive impact on learning. However, metacognitive judgments of learning are weak indicators of learning (Jaeger & Wiley, 2014; Paik & Schraw, 2013; Thiede et al., 2010; Yue, Bjork, & Bjork, 2013). These students, by virtue of their choice to participate in a voluntary science program, were almost certainly positively predisposed to the sciences before viewing the video and represent one of the important target audiences for Sci-Toons. In addition, a small number of viewers of the on-line videos provided optional comments which were nearly all positive, with several posing intriguing technical questions and requesting further information.

The underlying goals of Sci-Toons are to improve STEM education and stimulate interest in STEM. The Committee on STEM Education of the National Science and Technology Council recently issued a five-year STEM Education Strategic Plan, recommending five high-priority investment areas: (1) Improving STEM instruction, (2) Increasing and sustaining youth and public engagement in STEM, (3) Enhancing the STEM experience of undergraduate students, (4) Better serving groups historically underrepresented in STEM fields, and (5) Designing graduate education for tomorrow's STEM workforce (National Science and Technology Council, 2013). Application of the MLTF for the production and distribution of Sci-Toons can potentially improve STEM instruction, by providing comprehensible and easily accessible materials for educators to use in the pedagogy; increase and sustain youth and public engagement in STEM, through viewing Sci-Toons, and enhance the STEM experience of undergraduates, as indicated by favorable attitude changes toward science reported by college students involved with the production of Sci-Toons. These materials, although not explicitly targeted at underrepresented minorities, were designed to be accessible for all learners.

References

Adetunji, O. O. (2015). *UMRI Alumni: From Characterization of High Temperature Solution Growth Cr 2+: CdSe to Developing Videos to Enhance Metacognition and Diversity in Materials Science and STEM (Science, Technology, Engineering, and Mathematics)*. In MRS Proceedings (Vol. 1762, pp. mrsf14-1762). Cambridge University Press. doi:10.1557/opl.2015.188

Adetunji, O. O., Ba, J. C. M., Ghebreab, W., Joseph, J. F., Mayer, L. P., & Levine, R. (2012). Geosciences awareness program: A program for broadening participation of students in geosciences. *Journal of Geoscience Education*, *60*(3), 234-240. doi:10.5408/10-208.1

Agarwala, M., Li, W., & Berthouzoz, F. (2011). Design principles for visual communication. *Communications of the ACM*, *54*(4), 60-69. doi:10.1145/1924421.1924439

Bruner, J. (1986). *Actual minds, possible worlds*. Cambridge, MA: Harvard University Press.

Cohoon, J., & Chen, L-Y. (2003). Migrating out of computer science. *Computer Research News, 15*(2), 2-3.

DiBlas, N., Paolini, P., & Sabiescu, A. (2010). Collective digital storytelling at school as a whole-class interaction. *International Journal of Arts and Technology*, *5*, 271-292.

Falk, J. H., & Dierking, L. D. (2010). The 95% solution: School is not where most Americans learn most of their science. *American Scientist*, *98*, 486-493.

Gazzaniga, M. S. (2005). Forty-five years of split-brain research and still going strong. *Nature Reviews Neuroscience*, *6*(8), 653-659. doi:10.1038/nrn1723

Jaeger, A., & Wiley, J. (2014). Do illustrations help or harm metacomprehension accuracy? *Learning and Instruction*, *34*, 58-73. doi:10.1016/j.learninstruc.2014.08.002

Kahneman, D. (2003). Maps of bounded rationality: Psychology for behavioral economics. *The American Economic Review*, *93*(5), 1449-1475. doi:10.1257/000282803322655392

Kahneman, D. (2011). *Thinking, fast and slow*. New York, NY: Farrar, Straus & Giroux.

Karlsson, G.,& Ivarsson, J. (2008). In T. Hansson (Ed.), *Handbook of research on digital information technologies: Innovations methods, and ethical issues*, (pp. 68-82). Hershey, PA: Information Science Reference.

Kim, S., Yoon, M., Whang, S.-M., Tversky, B., & Morrison, J. B. (2007). The effect of animation on comprehension and interest. *Journal of Computer Assisted Learning*, *23*, 260-270. doi:10.1111/j.1365-2729.2006.00219.x

Margolis, J., Estrella, R., Goode, J., Holme, J. J., & Nao, K. (2008). *Stuck in the shallow end: Education, race and computing*. Cambridge, MA: MIT Press.

Mayer, R. E. (Ed.). (2005). *Cambridge handbook of multimedia learning*. New York, NY: Cambridge University Press.

Mayer, R. E. (2008). Applying the science of learning: Evidence-based principles for the design of multimedia instruction. *American Psychologist*, *63*(8), 760-769. doi:10.1037/0003-066X.63.8.760

McNaught, C., & Lam, P. (2010). Using Wordle as a Supplementary Research Tool. *The Qualitative Report*, *15*(3), 630-643.

Melguizo, T., & Wolniak, G. C. (2012). The earnings benefits of majoring in STEM fields among high achieving minority students. *Research in Higher Education*, *53*, 383–405. doi:10.1007/s11162-011-9238-z

National Governors Association. (2012). *The role of informal science in the state education agenda (Issue Brief)*. http://www.nga.org/files/live/sites/NGA/files/pdf/1203INFORMALSCIENCEBRIEF.PDF. Accessed 1 May 2015.

National Science and Technology Council. (May, 2013). *Federal Science, Technology, Engineering, and Mathematics (STEM) Education Five-Year Strategic Plan.* Washington, D.C.: Executive Office of the President of the United States. Available at https://www.whitehouse.gov/sites/default/files/microsites/ostp/stem_stratplan_2013.pdf. Accessed 1 May 2015.

National Science Board. (2007). Science, technology, engineering, and mathematics (STEM) education issues and legislative options. In R. Nata (Ed.), *Progress in Education* (pp. 161-189) Washington, DC: Author.

National Science Foundation, National Center for Science and Engineering Statistics. (2013). *Women, minorities, and persons with disabilities in science and engineering: 2013*. Special Report NSF, 13-304. Arlington, VA. http://www.nsf.gov/statistics/wmpd/. Accessed 1 May 2015.

Page, S. (2007). *The difference: How the power of diversity creates better groups, firms, schools and societies*. Princeton, NJ: Princeton University Press.

Paik, E. S. & Schraw, G. (2013). Learning with animation and illusions of understanding. *Journal of Educational Psychology*, *105*(2), 278-289. doi:10.1037/a0030281

Perna, L., Lundy-Wagner, V., Drezner, N. D., Gasman, M., Yoon, S., Bose, E., & Gary, S. (2009). The contribution of HBCUS to the preparation of African American women for Stem careers: A case study. *Research in Higher Education*, *50*, 1–23. doi:10.1007/s11162-008-9110-y

Prakash, A., Ohler, J., Cooper, C., McDermott, M., Heinrich, J., Johnson, R., Leeper, L., Polk, N. & Wimer, T., (2009). AMIDST: Attracting Minorities to Geosciences Through Involved Digital Story Telling. *Eos Trans. AGU*, *90*(52), Fall Meet. Suppl., Abstract ED52A-03.

Seymour, E., & Hewitt, N. M. (1997). *Talking about leaving: Why undergraduates leave the sciences*. Boulder, CO: Westview Press.

Sperry, R. W. (1961). Cerebral organization and behavior: The split brain behaves in many respects like two separate brains, providing new research possibilities. *Science, 133*(3466), 1749–1757. doi:10.1126/science.133.3466.1749

TagCrowd. (2015). Tagcrowd.com, Accessed 1 May 2015.

Thiede, K., Griffin, T., Wiley, J., & Anderson, M. (2010). Poor metacomprehension accuracy as a result of inappropriate cue use. *Discourse Processes*, *47*, 331-362.

Tversky, B., Agrawala, M., Heiser, J., Lee, P., Hanrahan, P., Phan, D., Stolte, C., & Daniel, M.-P. (2006). Cognitive design principles: from cognitive models to computer models. In: L. Magnani (Ed.), *Model-based reasoning in science and engineering* (pp. 1–20). London: College Publications

Tversky, B., Heiser, J., & Morrison, J. (2013). Space, time, and story. In: B.H. Ross (Ed.), *The psychology of learning and motivation* (pp. 47-76). Waltham, MA: Academic Press, Elsevier.

U.S. Department of Commerce. (2012, January). *The competitiveness and innovative capacity of the United States*. http://www.commerce.gov/sites/default/files/documents/2012/january/ competes_010511_0.pdf. Accessed 1 May 2015.

U.S. Department of Education. (2007). *Report of the academic competitiveness council*. Washington, DC: Author.

U.S. Office of Science and Technology Policy. (2010, September). *Prepare and inspire: K-12 science, technology, engineering, and math (STEM) education for America's future*. http://www.whitehouse.gov/sites/default/files/microsites/ostp/pcast-stem-ed-final.pdf. Accessed 1 May 2015.

U.S. Office of Science and Technology Policy. (2012, February). *Engage to excel: Producing one million additional college graduates with degrees in science, technology, engineering and mathematics*. http://www.whitehouse.gov/sites/default/files/microsites/ostp/fact_sheet_final.pdf. Accessed 1 May 2015.

Xu, Y., Park, H., & Baek, Y. (2013). A New Approach Toward Digital Storytelling: An Activity Focused on Writing Self-efficacy in a Virtual Learning Environment. *Educational Technology & Society*, *14*(4), 181-191.

Yue, C., Bjork, E., & Bjork, R. (2013). Reducing Verbal Redundancy in Multimedia Learning: An Undesired Desirable Difficulty? *Journal of Educational Psychology*, *105*(22), 266-277. doi:10.1037/a0031971

Yue, C., Kim, J., Ogawa, R., Stark, E., & Kim, S. (2013). Applying the cognitive theory of multimedia learning: an analysis of medical animations. *Medical Education*, *47*, 375-387. doi:10.1111/medu.12090

Zweben, S., & Bizot, B. (2012). Continued increase in undergraduate CS degree production; slight rise in doctoral production. *Computer Research News*, *24*(3), 7-24.

Appendix-Survey of SCG Students

Demographics

What is your gender?

Male [ ] Female [ ]

What year student are you?

Freshman [ ] Sophomore [ ] Junior [ ] Senior [ ] Other [ ] (Please Specify) ___________

What was your major when you started working with Sci-Toons?

____

When did you start working with Sci-Toons?

_________ _________

Month Year

Did you change your major?

Yes [ ] IF YES, What major did you switch to and why?

____

____

____

No [ ]

Since you began working with Sci-Toons, how many science, technology, engineering, or math (STEM) classes have you taken? If none, enter "0".

_____________ STEM classes

How many of these classes were electives rather than required classes? If none, enter "0".

_____________ STEM classes

IMPACTS ON ATTITUDES TOWARD STEM AND STEM CAREERS

We are interested in learning how participation in Sci-Toons has affected you.

As a result of your Sci-Toons experience, how has your ability to understand contemporary science issues (such as climate change, genetically modified foods, and the impacts of fracking) changed?

It has improved greatly [ ]
It has improved moderately [ ]
It has improved slightly [ ]
It has not changed [ ]

As a result of your Sci-Toons experience, how has your ability to explain scientific principles and issues to others changed?

It has improved greatly [ ]
It has improved moderately [ ]
It has improved slightly [ ]
It has not changed [ ]

Before Sci-Toons, how important did you feel it was for the general public to be able to understand science?

Extremely important [ ]
Very important [ ]
Moderately important [ ]
Slightly important [ ]
Not important [ ]

After Sci-Toons, how important did you feel it is for the general public to be able to understand science?

Extremely important [ ]
Very important [ ]
Moderately important [ ]
Slightly important [ ]
Not important [ ]

Before Sci-Toons, how difficult did you feel it was for the general public to understand science?

Extremely difficult [ ]
Very difficult [ ]
Moderately difficult [ ]
Moderately easy [ ]
Very easy [ ]
Extremely easy [ ]

After Sci-Toons, how difficult did you feel it is for the general public to understand science?

Extremely difficult [ ]
Very difficult [ ]
Moderately difficult [ ]
Moderately easy [ ]
Very easy [ ]
Extremely easy [ ]

As a result of your Sci-Toons experiences, how has the likelihood of your pursuing a STEM-related career changed?

It has increased greatly [ ]
It has increased moderately [ ]
It has increased slightly [ ]
It has not changed [ ]
It has decreased slightly [ ]
It has decreased moderately [ ]
It has decreased greatly [ ]

Formative Evaluation Questions

About how much time did you spend on the Sci-Toons project?

____

How did this compare with the amount of time you expected to spend on Sci-Toons?

It was much more time than I expected [ ]
It was more time than I expected [ ]
It was about as much time as I expected [ ]
It was less time than I expected [ ]
It was much less time than I expected [ ]

How did the amount of effort you put into Sci-Toons compare with the effort of the other members of your team?

It was much greater than the other team members [ ]
It was greater than the other team members [ ]
It was about the same as other team members [ ]
It was less time than other team members [ ]
It was much less than other team members [ ]

How much of an impact did you have on the final version(s) of the Sci-Toon compared to other members of the team?

It was much greater than the other team members [ ]
It was greater than the other team members [ ]
It was about the same as other team members [ ]
It was less time than other team members [ ]
It was much less than other team members [ ]

Did you feel that you understood the scientific principles underlying the Sci-Toon(s) you worked on?

Definitely yes [ ]
Mostly yes [ ]
Most no [ ]
Definitely no [ ]

Did you feel that you understood the process of making a Sci-Toon video?

Definitely yes [ ]
Mostly yes [ ]
Most no [ ]
Definitely no [ ]

In what ways, if any, has your Sci-Toons experience influenced your education or career plans?

What did you like best about your Sci-Toons experience?

What did you dislike about your Sci-Toons experience?

Journal of Research in STEM Education
ISSN:2149-8504 (online)
Vol 1, No 2, December 2015, PP 125-141

RESEARCH REPORT

A "Scientist" on the Radio: Understanding the Framing of STEM to the Public

G. Michael Bowen [a1], Richard Zurawski[a], Anthony Bartley[b]

[a]*Mount Saint Vincent University, Canada*

[b]*Lakehead University, Canada*

Abstract: *News media presentations of STEM (and particularly science) in various formats have been critiqued for the many ways by which they misrepresent both the facts of the discipline and the practices of the discipline and the researchers in them. Another issue is that the material is presented in a format – basically a one-way transmission – with usually little opportunity for questions by the recipients (i.e., readers, listeners, viewers, etc.) to be addressed when they don't understand something. One news media format which might allow this dialogic activity is the radio call-in show format which is structured so that the public can ask questions of a "scientist" with the opportunity for follow-up questions to address what are discontinuities in the listener's understanding. In this paper we document the processes by which listener interests ultimately end up discussed in the radio broadcast and what influences the "science" that is presented on-air. Our analysis reports the ways in which the STEM topics and content are mediated by radio station personnel, often times distorting the factual content available to the public and misrepresenting the practices of the research fields, as they engage in information management practices which are typical of opinion-driven shows (such as those on the topics of politics or sports) which are designed to create controversy and drama to increase ratings.*

Keywords: News, Media, Scientist, Journalism

Critiques of the misrepresentations of researchers and their findings in the news media are long-standing. In 1899 James H. Hyslop, a professor of logic and ethics at Columbia University in New York, wrote a letter to the editor of Science magazine in which he stated:

> "More than one-half of the interviews alleged to have been held with me were the fabrications of reporters who never saw me, and the other half omitted what I did say and published what I did not say." (p. 696).

Since then a voluminous body of research about how STEM fields, and particularly science and scientists, are presented in the news media have been published focusing both on large grain issues (such as incorrect facts) and on smaller grain ones such as writing style and the effect of specific genres of writing. Despite that research, according to Nelkin (1995) "…the style of reporting [science] has been remarkably consistent over time" (p. 1).

That we rarely hear stories critiquing how STEM fields are presented in the news media is relatively unsurprising given that the public tends to take writing about STEM topics (individually or collectively) in the news media at face value often attributing complete neutrality and lack of bias to reporters consistent with

[1] Corresponding author. G. Michael Bowen, Mount Saint Vincent University, 166 Bedford Highway, Halifax, Nova Scotia, B3M 2J6. E-mail: gmbowen@yahoo.com

Bowen, G. M., Zurawski, R., & Bartley, A. (2015). A "Scientist" on the radio: Understanding the framing of STEM to the public. *Journal of Research in STEM Education, 1*(2), 125-141.

how journalists themselves would like to be seen (Deuze, 2001). This may be because the implicit training the broader public has received in STEM topics, through schooling, itself offers a very narrow, restricted view of each of those subjects (most often, in school, being presented as completely different and unconnected topics).

The practices of STEM disciplines and the production of facts within those can be parsed in many different ways. Hodson (1998) argued for the use of three broad domains for thinking about the teaching and learning of science (which, we'd argue, is equally applicable to technology, engineering and mathematics in varying degrees):

i) 'Science'; i.e., products of science, including various laws and theories;

ii) 'About Science'; e.g., characteristics of processes and products of science including, for example, that developing knowledge in science is non-linear, theory-dependent and often influenced by investigators' idiosyncrasies, that certain cognitive 'skills', such as variable control, are used, and that there are various positive and negative effects of technological products on individuals, societies and environments and ways to rectify such problems; and,

iii) To 'Do Science'; i.e., expertise, confidence and motivation required to generate and communicate knowledge using methods of science and technology in unique problem-solving contexts.

Much like its presentation in the news media, school STEM subjects that the public experienced as students was most often about (i), sometimes a bit about (ii), and rarely about (iii). Hodson (2003) himself later modified his schema to include "Technology" with science, and to add a fourth dimension he considered important – "Engaging in sociopolitical action". It is this latter dimension which a critical analysis of news media provides us particular insight into, as it is the socio-political aspects of STEM disciplines – often presented in the form of "scientific" controversy – which constitute a large part of news media representations of STEM issues (consider such topics as nuclear energy, immunization, global warming, acid rain, pollution, and so forth). Elsewhere we have argued that the traditional approach to the presentation of STEM topics in the classrooms influences how the public engages in these sociopolitical issues because their expectations about how science is done and communicated, as portrayed in the classroom when they were students, is not reflected in the "science in the making" (Shapin, 1992) or "technology in the making" presentations of issues such as climate change (Bowen & Rodger, 2008).

For most of the public the news media is a significant source of information about STEM issues (and particularly science issues; Boyes & Stanisstreet, 1992; Dispensa & Brulle, 2003; Lewenstein, 2001; Schibeci, 1990) where it plays a significant role in shaping the discourse on socio-political STEM topics such as climate change (Boykoff, 2008; Wilson, 1995) influencing "both the public's knowledge and attitudes towards science." (Pellechia, 1997; p. 63) News media are the most frequent source for learning information about science/STEM topics (Miller et al, 2006; Nelkin, 1995), and 80% of adults obtain their science information and science learning from local television broadcasts.

The stage against which this news media critique can be held is the view that journalism holds of itself. In a review of the literature a decade ago, Deuze (2001) identified the ideals to which journalism holds itself:

- Journalists provide a public service;
- Journalists are neutral, objective, fair, and, thus, credible;
- Journalists must enjoy editorial autonomy, freedom, and independence;
- Journalists have a sense of immediacy, validity, and factuality;
- Journalists have a sense of ethics and legitimacy.

Even if these ideals were truly held to in their entirety there would still be reason to be concerned about the use of news media in the classroom to teach about STEM, if they are not being held to then there would be even more reason to be concerned. Fairclough (1992) warns that the media wields enormous power in its discourse and that understanding what it disseminates to the public as the ideals of the news media deteriorate

(Deuze, 2001, Zurawski, 2011) is to be ignored at our peril. It is clear from this that a critical examination of STEM topics in the news media and how they are presented to the public is warranted.

STEM Programs in the United States

There is a broad and detailed academic literature on the news media and issues with their presentation of science and scientists. We will start with a description of the writing practices of journalists, how they differ from academic writing in STEM research, and what the implications are for that with respect to the representation of research.

The following discussion often focuses on "science" (in contrast with the various STEM disciplines) and how it is presented in the news media because the idea of STEM is a more recent creation and past literature (a) past academic literature doesn't necessarily distinguish between science, technology, engineering and mathematics in how it defines "science" in its analyses and, (b) news media typically focuses on more traditional "science" in its discussion (i.e., physics, astronomy, health sciences, climate change, pollution) and less on the other aspects of STEM.

How Journalists Tell Stories

No matter which type of news media format is being used – radio, television, print, web page – reporters generally structure their story in the same way. The standard approach in journalism for telling a story is to use the opposite approach to that which we learned in our language arts schooling, and which is similar to that in science, where one starts with the context for the story and builds, through various stages, towards a climax at the end of the story. In science/STEM academic writing for journals that progression generally involves – in order – a literature review, a statement of a research question, a described methodology, a description of the data that was collected, and then a discussion of why that data was important and what was learned from that. In STEM academic writing this sequence is important as it helps build and reinforce the credibility of the findings and conclusions.

Journalists approach their writing in a completely different way than we were taught to write in language arts classes in school. They use an approach referred to as an "inverted pyramid" where the story starts with the main information and then context/evidence is laid out in a narrative below proceeding from the most to the least significant information (see Scanlan, 2003 for more details). In addition to that, journalists also provide little coverage of some important aspects of STEM inquiry such as the methods used in the investigation (thereby misrepresenting important aspects of the nature of science) and they use a narrative style adopting a "sentimental language" with metaphors and analogies to describe the world. This narrative approach has been found to have a high appeal to a readership (Halkia & Mantzouridis, 2005) but along with the missing descriptors misrepresents the practice of researchers as they work against "summative or comparative analysis that is necessary for evaluating the merits" of that or competing academic research work (Charney, 2003; p.237).

Part of establishing the credibility of a "science" fact claim lies in the detailed description of the methodology used to establish the piece of evidence. Yet, news media articles about STEM investigations "frequently omitted methodological and contextual information [of the science studies], features most often mentioned as critical for a complete journalistic account of science." (Pellechia, 1997; p. 49) and this lack of a methodological description has been also noted by others (Dimopolous & Koulaidis, 2003; Bucchi & Mazzolini, 2003; Einsiedel, 1992).

Apart from the structure of the piece itself, journalists also construct their story in a way that is the opposite of the approach taken in STEM subjects in a different way. Figure 1 illustrates the two different story approaches. Journalists tend to write a science story by referring to an "expert" who is talking about an idea and even if they are talking about a collaborative research paper they will often only refer to a single researcher when discussing the findings (Figure 1 top half). Journalists are encouraged to write about "A person, doing something, for a reason" (see Bowen, 2011) and when writing about science this means that the story often focuses on one individual researcher despite the number of other researchers who may have participated in various ways – directly or indirectly – in creating the knowledge claim (in fact, journalists will deliberately ignore the other researchers in the writing, even against the specific wishes of the person being interviewed (unpublished data; Kim P. Good, pers. comm.)).

STEM researchers, on the other hand, write papers which are about an idea or ideas which they subsequently support by referring (a) to the evidence that they collected and (b) to the collection of authorities who contributed to developing the knowledge claim(s) that the research article is about (Figure 1 bottom half). Thus, the writing approach engaged in by journalists often misrepresents the communal aspect of how STEM research is often done (see Latour & Woolgar, 2013) as it is presenting research as if it was done by individuals as opposed to collectives of researchers over a long period of time.

The use of only a single source to discuss (and more importantly, support) a STEM idea/knowledge claim/finding (as portrayed in Figure 1) lends itself towards a very common journalistic practice, and that is the way in which journalists provide "balance" to a story by first presenting the story as having two "sides" and then choosing one individual to speak to each side of a story. For instance, between 1988 and 2002, 88% of newspaper articles on global warming science had a "denier" perspective included in them (Boykoff & Boykoff, 2004; also see McBean and Hengeveld, 2000; Curtis, 2007). In general this approach does not work well for a great number of STEM topics, particularly those involving the environment, animals, new technologies, etc (Myers, 1996)…in other words socio-scientific issues. The practice of providing "balance" is intended to "display objectivity" and originates in British parliamentary democracy for which political issues were seen as "having two, and only two, sides" (Myers, 1996; p. 34/35). Yet, what it accomplishes is a misrepresentation of some of the very core ideas of STEM/science reporting issues by reporting them as black or white without any room for nuance or grey areas (Maille et al, 2010).

One core idea that this for/against presentation misrepresents is fundamental to the very nature of how STEM knowledge claims are constructed. STEM research is not a process whereby a definitive "answer" is produced but is, rather, a probabilistic endeavour where researchers produce understandings of "likely" answers. News media, however, portrays STEM (and particularly science) investigations as producing "certainty" (see Collins, 1987) ignoring that at its core scientific research is a probabilistic enterprise. Journalistic "balance" comes about from first dichotomizing scientists into "for" and "against" communities and then treating those as if they lie at extreme poles of absolute statements as opposed to the grey areas involving nuance and probability that characterize science. This artificial creation of balance exaggerates the debate and underplays the consensus (Wilson, 2000; p. 11). Given the desire of journalists to create dramatic tension in their work for narrative purposes (as would happen in fictional novels and stories) they would have little tendency to frame their representation of scientists' work in other ways. Even from this basis alone teaching students "about [STEM] science" using news media distorts their understanding of STEM science as a practice and how it establishes factual claims. An examination of discussion threads appended to news articles in science makes it clear that this distortion of the practice of science and how it derives its fact claims by the news media is not understood by a large number of the reading public (Bowen & Rodger, 2008). This misunderstanding leads to an "all-too common popular view [amongst journalists], if scientists cannot produce definitive results, then they are not doing their job properly" (Charney, 2003; p. 237).

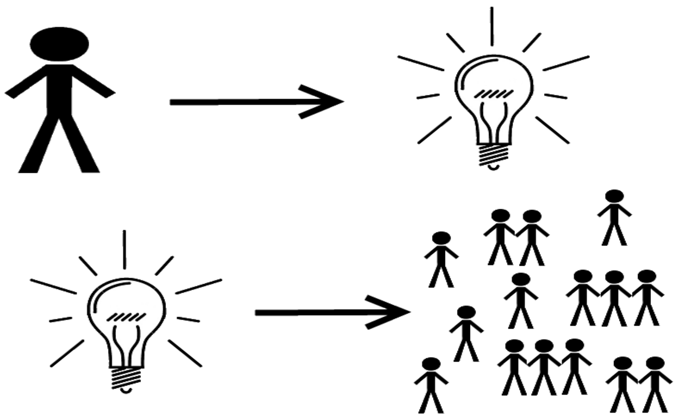

Figure 1. *The top half of the diagram illustrates how journalism portrays science research ideas/facts/discoveries as if they extend from a single individual. The bottom half of the diagram illustrates how writing in STEM science is about an idea which is buttressed using various individual and collections of authorities (Donald, 2012). This "community" aspect of STEM discovery is often ignored in journalistic writing.*

Who/What Influences Which Stories are Chosen

At one level, narrative factors are what drive journalistic attention (McComas & Shanahan, 1999) such that "When journalists see the possibility of narrative interest, their attention goes up." (Shanahan & Good, 2000; p. 294). In other words, the possibility of being able to tell a good story drives their interest in a topic. "The central assumption of the popular press journalists is that unless the science which they cover can be seen to be relevant to the reader's daily life and concerns, it is not suitable for coverage." (Hansen, 1994; p. 129) Yet, one needs to recognize that despite their claim that the news is a "mirror held up to society, it is actually a highly selective account of events. News is a version of reality shaped in significant part by journalistic norms and conventions. In addition, journalists are shaped by pressures by those who have a vested interest in the topic or the newspaper/magazine" (Dispensa and Brulle, 2003; p. 81).

Irrespective of the type of journalism – print or broadcast journalism – there is an "editor" or "producer" of some sort whose job it is to control what stories journalists are doing and how they are going about structuring them to be presented to the public. It is uncommon for journalists to be able to make their own choices without considerable input from their editor(s) or producer(s). In those instances they might receive direction to cover one story or another without any insights into why the choice was made, or it might even run against their instincts about what the best story would be (see Bowen (2014) for detailed examples of this).

Overall, ideology determines what is considered STEM news by journalists and producers/editors and the interpretation of the information is strongly entangled with ideological viewpoints (Carvalho, 2007). This results in a narrowing of STEM science topics discussed in the news media, with a general focus on issues that affect the daily lives of individuals (Hansen, 1994) such as biomedicine (Bucchi & Mazzolini, 2003), health issues (Almeida, 2013; Pettersen, 2005; Hansen, 1994), behavioural studies (Clark & Illman, 2006), and environmental stories (Einsiedel, 1992).

That these issues arise are unsurprising given that the science (and other STEM area) backgrounds and understandings of journalists (and their editors/producers) writing about science are weak. Generally news journalists do not have training in science (Hansen, 1994; nor do some consider it to be necessary, see Bowen (2014)) and generally they have notable gaps in their knowledge in the science areas they write about (Pettersen, 2005; Wilson, 2000; Maille et al, 2010) as well as having a "restricted" knowledge of statistics and scientific discourse and the very "nature of science" (Pettersen, 2005).

The touchstones of STEM research – “uncertainty, probability, an uncertain time frame, and the frequent pursuit of knowledge for its own sake” – run against the journalistic preference for “certainty, fewer caveats, shorter time lines, and the entertainment of audiences” (Einsiedel, 1992; p. 100). It is hardly surprising, given the above, that popular media accounts of STEM research “may breed cynicism about the scientific community itself” (Charney, 2003; p. 238).

The Public Interest in the Communication of STEM Issues

In this review we've identified many areas of considerable problem regarding communicating science to the public. How might we redress issues communicating STEM to the public while at the same time better understanding the public interest in STEM research itself?

In the following section we examine a radio call-in show (called “Science Files”, even though the questions from the public that it addresses range broadly across STEM topics) where an expert “guest host” (who has a quite broad and encyclopedic background in science, technology and mathematics issues) joins a regular morning show host taking STEM questions from the general public and providing answers to those questions. Our interests in this project include (a) documenting production practices within a broadcast facility to understand how these mediate and influence the “science” available to the public in the broadcast itself (having documented elsewhere how practices apparent “outside of the studio” during the public broadcast/presentation of science influence the view of science/science claims available to the public; Bowen (2014)), and (b) analyzing how the interaction between the hosts and those calling into the station frame and present science to the listening public. In the sections below we first describe the process by which STEM “topics” enter into the discussion on the radio call-in show. Then we briefly analyze three of the radio broadcasts to examine the types of questions asked by the callers and how those arise.

Data Sources and Methods

In our examination of the call-in show on a local commercial radio station (serving a market of about 1 million potential listeners) we have used two information sources (one triangulated) for the analysis conducted herein. The first author (Bowen) obtained a journalism degree as part of an academic sabbatical five years ago. This provided him the opportunity to participate in radio news broadcasts, including on science topics, as a student journalist which helped shape his perceptual lens examining the call-in show in this study as he observed half-a-dozen of the radio broadcasts from the producer's booth (and later subsequently acted as a co-host himself with related science radio call-in show). The second author (Zurawski) is a journalist with a broad science background (particularly in physics, meteorology and weather forecasting) and considerable knowledge about other STEM issues. Amongst other accomplishments he has produced three broadcast television series (one science, one for math and one in weather) for children, dozens of TV science documentaries (broadcast on Discovery Channel, ZDF, VisionTV, etc), written popular press books on both climate change and the news media presentations of science, and is in ongoing demand to give public talks on science, STEM and related socio-political issues. This author is the “scientist/guest” host in the radio broadcasts described in this paper, who addresses the public STEM questions and is the resource used to provide detailed descriptions of the practices engaged in to produce the call-in show. The involvement of both authors in news media journalism, the area of this paper, and the second author's involvement as the science “guest” means that this work falls within the realm of “participatory research” (Bergold & Thomas, 2012). Conversations and email exchanges between these two authors framed the description of the processes involved in the radio call-in show production (and the third author acted as a “critical friend” in our final analysis of our documents and conversations and in the production of the final paper).

In addition, we downloaded publicly-available recordings of the on-air broadcast and transcribed and analyzed three of the broadcast radio call-in shows. . The analysis was conducted to provide insights into the construction of STEM understanding and research practices through the interaction of the caller, the morning-show host, and the "scientist/guest" host. Topics and questions (asked by those calling/writing in to the show) were then discussed to provide insights into the public interest in and understanding of STEM issues.

It is important to note that comments discussing the host or the producer are made in full recognition that they are serving particular roles within the broadcast of a radio show. That we may report that these roles and how they are enacted may not well represent STEM disciplines and the knowledge developed by them or STEM issues is not a criticism in any way of the individuals enacting these roles because they are acting (as we report) according to common journalistic practices, instead we are highlighting issues with how those common journalistic practices, that may well suit sports, politics or regular news reporting, seem to conflict with the presentation of STEM disciplines and issues.

Data and Discussion

Background of the Station and Call-in Show

The call-in show was broadcast (as well as streamed online) on a commercial station to the local market twice each week during regular hours having a consistent host on Monday to Friday. This regular schedule allowed listeners to plan to listen to the science call-in show or to plan to call, email or tweet in a question for the show. The call-in show was promoted in "blurbs" the day before and morning the morning of the show before it occurred. Once the show started and the science host was introduced the phone number, email address and twitter handle for the show were provided and listeners were encouraged to call in with their questions.

According to the Bureau of Broadcast Management (which conducts surveys of listenership twice a year) the station had a typical listenership (when the study was conducted) of around forty thousand listeners and the science call-in show was at or near the top of all of the local station programming with regards to number of listeners, caller volume, internet responses, and web listenership.

The show itself was staffed by four people (there was also a "news" broadcast person who reads the news in a short broadcast on the "halves" (i.e., at the top and bottom of the hour), with brief commercial breaks on the "quarters" (i.e., at fifteen minutes before and after the hour). The other staff were: the show "producer" who runs the technical equipment supporting the show (i.e., microphones, controlling the information and timing screens, running advertisements, promotional audio, weather forecasts, pre-screening choosing which callers get on air, and other blurbs) from the control booth, the show "radio personality" (i.e., regular morning show host) who ran a 3 hour news and information discussion show (which includes interviews and call-in components) on weekdays, and the "scientist/guest" host who answered the questions raised during the radio call-in show segments during the week. Each had responsibilities specific to the running of the show (Table 1). Both the regular show host and the producer spend about half their work day preparing for the show, including lining up callers and topics to be discussed as part of the show (if there were no calls the host and "scientist/guest" host discuss topical science issues researched in advance). The producer works in conjunction with the show host to "create" topics deemed to be of interest to the station, timely topics breaking in the news, and local issues for the local market.

Table 1:
Roles of Participants in Radio Call-in show.

"Caller"	Show Producer	Regular Show Host	"Scientist/Guest" host
Asks initial question Elaborates question if show guest requests it, or responds to the answer May participate in conversations with the two hosts if the producer and regular show host allow it.	Takes calls Screens calls (removing callers who don't have a question, who want to "rant" on a topic, or who have non-science based questions (these were typically "known" callers with "issues")) Finds out topic(s) of accepted call(s) and passes name of caller(s) and topic to show host electronically* May conduct web searches on topic to forward to host* Make sure show sticks to timings	Be entertaining and engaging Stick to timings Make sure there is no "dead air" Act as a foil for the "scientist/guest" host Repeat and repeat things (radio broadcast is unlike other media in this regard) Moderate the phone-ins Raise follow-up questions to get the "scientist/guest" host to elaborate on his explanations and provide more detail Have own topics of discussion for "scientist/guest" host in case there are no call-ins.	Answers questions as accurately as possible Reframes questions if necessary Be entertaining and engaging Have his own topics of discussion in case there are no call-ins. Creates dialogue and banter with host

*Usually by email message, sometimes by voice over earphones.

How "Topics" Reach the Airwaves

Topics discussed on-air emerged from interactions between the caller (who could contribute questions using a phone call, by email, or by tweeting), the producer, the host and the "scientist/guest" host. Not every question asked by someone calling-in made it to air, and there was often a considerable influence enacted by the show producer and host on what was discussed by the "scientist/guest" host on-air (at other times there was none, as the topic was raised and discussed by the "scientist/guest" host himself).

The following describes how the show began. The call-in show was preceded by a news break on-the-hour and then the "scientists/guest" host was introduced. There was then often "banter" between the host and the "scientist/guest" host and during that banter, varying in where it was introduced, was a request for listeners to call-in with questions. Despite the other ways listeners could contribute questions (i.e., email or tweets) phone calls which allow interaction between the caller and the "host" and "scientist/guest" host appeared to be the preferred method of having questions asked. Therefore, the preferred hierarchy for discussion appeared to be (a) voice/listener call-ins, (b) discussion topics selected by the host or the "scientist/guest" host, (c) then emails or tweets submitted by listeners. Occasionally (c) reflects questions submitted during the week between broadcasts of the science call-in show.

When a phone call was received it was first "screened" by the producer – who was at a control console in the control booth – who evaluated whether the caller should be allowed on-air or not. In the few circumstances observed where a caller was not allowed on-air it appeared to be because the question was not really a science question or because the caller was a repeat caller who had caused issues when calling in the past. Based on the reactions of the producer (and later, the host) it was clear that some callers were repeat callers.

After the producer determined that a question was suitable for broadcast, the question topic was then communicated to the host (usually by electronic message (using a messaging tool like email which was part of the production management system) or by voice communication to the hosts headset (which was not audible to that of the "scientist/guest" host)). Depending on call volume or other show management responsibilities, if they had time the producer would then look up websites on the topic(s) (which had been forwarded to the regular host) and if any relevant URL's were found they were emailed to the host (but not the "scientist/guest" host, who was expected to respond to the call-in question extemporaneously).

Often at the start of the show there was no listener phone calls received by the show producer. Emails and tweets containing questions submitted by listeners went directly to the call-in show host. Thus, for discussion the host then could choose from any emails or tweets he'd received between show broadcasts, he could raise a question which he'd discussed with the "scientist/guest" host before the show was on-air (in these instances it may be a topic proposed by the "scientist/guest" host himself), he could raise a topic suggested by the producer (for which online resources/information may have been provided to him), or he could choose from topics he'd selected n himself prior to the show. These latter topics were important in laying the foundation for an engaging show. The host would have a reserve of what he considered to be "hot button" STEM news topics culled from the print (and other) news and other sources designed to kick-start the callers, tweets and emails. Often these topics would cover typical denier/contrarian STEM controversies designed to cue audience response and participation. Examples of the line of questions/topics to discuss with the "scientist/guest" host might include recently published articles by "anti-vaxxers", or climate change deniers or creationism tenets. It was clear, watching the main host, that he viewed his role as one which involved introducing, encouraging or increasing "drama" (a common role of news media).

It was notable that questions which were chosen by the regular host were ones which he had sufficient time to prepare for in advance by researching them online (or having websites provided by the producer) allowing him to engage in meaningful and seemingly knowledgeable dialogue with the "scientist/guest" host when the questions were asked. Even as the dialogue was taking place between the caller and the "scientist/guest" host, the main show host would search Google or Wikipedia on his computer to garner additional information on the question or topic and then chime in as the conversation progressed, to bring himself into the question discussion (between the caller and the "scientist/guest" host), often with the intent of continuing the journalistic pro/con adversarial polarization to "spice up" the dialogue and raise the drama level with the goal of raising listener engagement (strategies for doing this, such as choosing particular "verbs of saying", are described in Bowen (2014)). This was the case especially if the show host thought the question asked by the caller was too "low key" (i.e., having low listener interest or not being dramatic enough to be engaging).

Questions and topics raised by the show host derived from the following sources:

- provided by the "scientist/guest" science expert (in advance of the show).
- provided by the producer (who would email him links to science topics the producer thought were interesting).
- derived from emails/twitters from listeners (sometimes identified as coming from a tweet/email by the host, sometimes not).
- from the host's own research about science topics before the show reflecting his own interest as well as being topics he thought would be engaging and interesting.

In contrast with the host, the "scientist/guest" host was often unaware of what question/topic was going to be introduced and was expected to engage it without any opportunity to prepare. In addition, unlike the host who had access to a computer which was actively used by him throughout the broadcast, the "scientist/guest" host did not have a computer in the broadcast booth and was expected to answer the questions or engage the STEM topics extemporaneously.

In most cases the banter between the host and "scientist/guest" host was designed by the host to pander to the perceived notion that there must always be a two-sided dichotomy when discussing STEM issues (similar to the "balance" approach used by journalists as discussed earlier in our introduction). Where there was no immediate question or caller/email or tweet the host would raise a piece of "topical" STEM "news" which was presented as a dichotomy with the host supporting one view and the "scientist/guest" host the other, often opposite, perspective. For instance if a headline in STEM news was about the prospect of the existence of life on other worlds, the "scientist/guest" host would be asked his position and the host would immediately take the other side in an adversarial fashion to provide the perceived balance prevalent in journalistic practice. The issue of the actual STEM content and how the host might skew the perception of the topic in the public mind to provide these polarized decisions was rarely if ever discussed between the host and the "scientist/guest" host. This process extended into all areas and fields of STEM, from Climate change to anti-vaxxers to Darwin vs Creationists.

Using this adversarial approach it was a common practice for the regular host to attempt to undermine and find some fault, no matter how minor, with the explanations provided by the "scientist/guest" host. The positioning of trying to undermine science/the scientist in all areas discussed was prevalent throughout the hour, in an almost game show manner. The task of the host was implicitly to be the foil against STEM findings and any researcher. The regular host's language and discussion was often not around how to clarify or make the STEM information understandable or discernible to the listener, but was more of a "contest" between STEM research and the public, where the goal was to "stump" the "scientist/guest" host and thereby "prove" the fallacy of scientific method. Common statements throughout the run of the show to that point have been "How do you know?" and "You think you are so smart" (the latter clearly an attempt to frame the "scientist/guest" host as an elitist). In fact, early on during the show's inception there was a substantial portion of the show dedicated to "stumping the science chump" where putting the "scientist/guest" host into the position of saying "I don't know" was a goal of the show host.

Insights into the Public Interests in Science through Call-in Questions

Often, questions asked by the callers (or submitted electronically; note that we did not have access to the questions submitted by email by listeners to fully comment on this) appeared to be "one-offs" in that they did not derive from earlier questions or discussions between the show host and the "scientist/guest". Callers frequently appeared to struggle with framing and centering their question when they were asking it and they often left their discussion/question/sentence/voice trail off leading the "scientist/guest" host to broadly discuss the topic they had raised with the caller, or the host, asking further or elaborative questions during the conversation. It was notable that in the three shows chosen for analysis that all callers appeared to be male (despite a listenership that was roughly split 50-50 between males and females (Bureau of Business Management) callers were rarely female with this host) having male names, using male referents, and having stereotypical male vocal patterns and tones.

Examples of the types of questions submitted by listeners are found in Table 2. Of the three broadcasts which were part of this study there were six, four and seven listener participants in the hour-long broadcasts (with short news and advertising breaks four times in each broadcast hour). In some cases the listener asked a single question, listened to the response, and then thanked the "scientist/guest" host and hung up. In other cases there were extended dialogues with follow-up questions between the caller and the "scientist/guest" host.

Table 2:
Examples of questions asked by the listeners.

[Phone in] Raw meat, if it is cured as in Proscuitto or salami it does not spoil. But boil it and it spoils. Why?
[Emailed in] Since the moon is held in place by the pull of the earth and the pull of the sun, would any change to the weight, shape or size change its orbit/distance from the earth? If so, what would be the effects if it moved closer, and also if it moved further from the earth? I ask this because every now and then there have been discussions about mining the moon for resources and that could, over time, eventually change the moon. I hope this isn't too confusing. I look forward to hearing what you can tell me about this.
[Phone in] I'm trying to create a science experiment here. Basically, I'll give you a little story about what I'm trying to accomplish here. I was watching this show called "Brad Meltzer's Decoded" and they had some expert there talking about magnetic vortexes. So they had this so-called expert on magnetic, I'm not exactly sure how you'd become a certified expert on magnetic vortexes, but he claimed to have one. A miniature one. He had a 3 magnets he had (word unclear) he had put into the shape of a star and basically the effect he was saying if you stand at the edge of the vortex, two people, standing opposite each other, one will appear to grow taller and one will appear to shrink. Now they re-created the effect on TV. The people that took part in [cell phone cutting out] I put two magnets...[voice faded out]
[Emailed] Why are there so many sinkholes in Florida? Any idea?
[Emailed] Is quicksand real or is it an invention of bad 70's television writers?

Question topics were from a broad range of science domains (although there was a somewhat greater tendency towards physical science/astronomy/engineering topics, possibly reflecting the academic background of the "scientist/guest" host) touching on topics (across many shows) exceeding those in this study, all of science, math, engineering and technology and even psychology and sociology (Zurawski, Bowen; unpublished data). In addition to the variety across subjects, the topics raised by listeners or the show host were often ones which could be considered "sensational" or "dramatic". This effect was exacerbated by the host (as described earlier) participating as a "foil" or challenger for statements by the "scientist/guest" host – who was answering the question asked by the caller – in that the regular show host "polarized" or split topics into "for/against" perspectives in what was an obvious attempt to raise the drama(tic) level. Thus, listener questions, which were seemingly submitted for informational purposes, were turned into "debates" by the show host, in many cases obfuscating the factual responses that the "scientist/guest" host was attempting to provide in response to the question.

Based on their comments and the interactions with the hosts, the questions asked by the listeners appeared to derive from many sources (see Table 3).

Table 3:
Apparent sources of questions from the listening public

Own "everyday" experience
From previous radio call-in show topics (host & science expert discussions)
Newspaper and other news items
TV/Movie entertainment
Past school experiences

Finally, occasionally the conversation between the show host and the "scientist/guest" host influenced the questions asked by the listeners. For instance:

1. Host and "scientist/guest" host discuss "dreams" for 10:20.
2. Phone in questions from callers (4 in total) then deal with "dreams" for the next 7:30.

Following the host/"scientist" discussion of dreams listeners the listeners phoned in with the following questions about dreams and dreaming (names below are pseudonyms):

- Corey (male): While I was just wondering, you were talking about dreams and stuff, anyone who has quit smoking and drinking, mostly with smoking, you take the patch and usually you sleep with it and your dreams are so real. They totally change your everyday dreams. I was just kind of wondering how, if they did studies on it, or why that happens....
- Jimmy (male): Do animals dream, or is it something that's a side effect of the evolution of man? Is it something like as our brain got more sophisticated it's like an off-gassing type of sort of thing or do other animals all dream? And why?
- Dan (male): Why, I get a good night's sleep or whatever, but I never seem to remember dreams. That happened since I was a teenager.
- Bernard (male): I was wondering, I always dream about people and places I don't know anything about and I was wondering if you had any comment on that?

This "seeding" of the broadcast with ideas to stimulate calls from the listeners is unlike the process which can occur in the surveys usually used to understanding what science topics the public is interested in within which there are no such "prompts". As can be seen in these questions the influence of the topic of "dreams", as first introduced by the show host, provides a rich substrate for listeners to engage with a science topic linked to their personal experiences (which most of the questions reflected). It is unlikely that this sort of detail about interests in a specific science topic, such as dreams in this case, could be gleaned from a survey designed to understand public interest in various STEM topics.

From the preceding narrative descriptions of practices engaged in by the participants involved in broadcasting a science call-in show the following observations deserve highlighting:

- stereotypical, and normative, journalistic practices of creating tension/drama play-out even in a call-in show dealing with science, technology, engineering and mathematics topics
- the host often acted as an antagonist (versus as a protagonist) with the "scientist/guest" host, seemingly aiming more for entertainment and obfuscation on STEM issues rather than being informative
- journalistic practices of providing "balance" were present in the radio broadcast through the framing of STEM issues (by the host) as being polarized into opposite for/against perspectives
- science was presented in a "competitive" frame where the host was striving for "I know this and you don't" types of discussions, and this was exacerbated by the host and the producer having computers, access to the internet, and advance knowledge of the questions; which were resources not provided to the "scientist/guest" host. Pre-show preparation by the host/producer on topics unknown to the "scientist/guest" host further emphasized this "competitive" aspect.
- control of topics, and the length of engagement with the public, lay mostly in the hands of the station employees – the show host and producer – the "scientist/guest" host often had little influence on topics/engagement with the callers
- public interest in STEM issues was wide-ranging but was often somewhat unfocused in that the callers were interested in the topics, but were often unsure how to phrase their questions around those topics (as the quoted calls & emails suggest).

Conclusions and Implications

The structural design of this call-in show is typical of that found in talk radio in that it has a "caller", an "expert", a "studio host" and a listening audience (Hutchby, 1995). What is revealed in this analysis is that the producer has a role in the topics and information conveyed in the show to the host and "scientist/guest" host, and further that the producer played a role in the "antagonistic role" of the show host. The process through which "questions" flowed before reaching the airwaves meant that the calls that reach the "scientist/guest" host to be answered were filtered first by the producer (w.r.t. phone calls) and secondly by the regular show host (who was able to filter all listener input) before the questions were actually asked of the "scientist/guest" host. This is quite similar to a process we have described previously (see Bowen (2011, 2014)) where it was an unknown producer who decided which science stories a science journalist was going to create a news segment about on a particular day. In other words, which STEM topics make it on air during the call-in show were determined in large part by people in news media without a background in science who have an eye on protecting the interests of the news media outlet itself (i.e., ratings, sales, etc), in this case of the call-in radio broadcast both the producer and the regular show host. Even in this venue of call-in radio, unexpectedly we might add, there is a buffer between the listening public and the "scientist| that mediate the STEM issues that are considered publicly relevant and of public interest. This "control" by the host of the topics and how the "scientist/guest" host and the listeners get to participate in the construction of the on-air "science" is similar to that which has been reported for talk radio broadcasts about politics (Thornborrow, 2001), although the role of the producer in screening calls and providing web-information to the host for discussion was unreported in that study. This similarity is particularly true when it comes to the asymmetry resulting from interventions by the host that limit the contributions which are possible from the other participants in the on-air dialogue, despite the contrast between the socio-political content in Thornborrow's work and the more fact-based content in the science show (which the host appeared to attempt to make controversial and polarized).

Topics often raised by the host (and sometimes by the "scientist/guest" host) often reflected the "spectacle" nature of STEM/science (Davis, 1997) so that listeners were entertained and engaged. This was emphasized by the host's use of oppositional strategies, such as "you say X, but what about Y" (see Hutchby, 1992) to polarize the conversation and develop controversy. This strategy would seem to be adopted to negate any "grey areas" and to present the STEM findings as black or white without any nuance (see Maille et al, 2010). In many ways the approach resembles that of a cross-examination in court, whose purpose is to raise doubt about the initial explanation by the witness, and in this instance to raise doubt in the explanation provided by the "scientist/guest" host. Thus, despite the "scientist's" efforts to answer questions in a straightforward manner presenting the best information available from STEM research, the listener's experience provided a considerable opportunity for them to doubt the answers provided by the "scientist/guest" host (based on the arguments of the show host) and instead adopt an understanding about the issue quite at odds with the current scientific understandings. This emphasis on non-scientific explanation is similar to that which is present in journalistic practices in print media (Dearing, 1995). These practices have considerable implications for what the public hears and learns about STEM disciplines (and how they frame their understanding of those disciplines), how science information is conveyed (combatively, if the show is taken as an example), and how knowledge is conceived and grows (in a polarized, oppositional manner as modeled in the show). Given this, it is worth asking if call-in shows such as this one actually do STEM subjects any service at all, or do they cause more issues in public understanding than they resolve?

One positive role that a "scientist" in a phone-in show can serve is as an effective interpreter of language contrasts between the use of terms in STEM fields and that which is in "general usage" by the public. Somerville and Hassol (2011) identified numerous words – such as bias, theory, and manipulation – that had different meanings to the general public than they would to researchers. Discussion of these terms by a "scientist" affords the opportunity for these differences to be discussed, and the dialogues between the caller or the host and the "scientist/guest" host provided a greater opportunity for these different understandings of terms to come to the fore. This might well be particularly educative for the listening public as there is evidence that call-in radio shows can lead to the listeners actively participating in the construction of meaning (Tankel, 1998).

Nevertheless, there is almost an "inevitability" for the existence of communication problems between the STEM research and the public (Neidhardt, 1993; p. 348), including journalists, because of the increase in complexity of language in research that is occurring (Hayes, 1992). This is considerably problematic as "when people cannot understand, they have to believe. And whether they believe or not is a matter of trust." (Niedhardt, 1993; p. 348) which means that when trust is lost, so is the acceptance of the STEM research. It is easy to see how this would happen in a context where incorrect information is presented (such as when the host framed science knowledge oppositionally against some type of "denier" claim raising skepticism about the established science itself) because of a lack of understanding by the journalist of the STEM research itself, and a lack of time exists (through what is arguably a cultural practice of journalism; Bowen (2014)) to allow the journalist to move beyond the "instant expert" role that they both are put into and accept (ibid) so that they cannot actually developing expertise in the area. In the radio call-in show the host clearly demonstrates this "instant expert" approach of journalism through the role he takes in the show whereby he portrays himself as knowledgeable about the topic under discussion through his (and his producer's) use of the internet on-the-fly to frame his discussion with the "scientist/guest" host.

The heavy prevalence of male callers (there were female callers in other shows that were observed, but they had a very low prevalence) is not atypical for call-in shows. Tannen (1990) provided a similar description of a call-in show which had a similar 50-50 split in listenership, but which even on topics where one would think there would be a high female participation (i.e., abortion) there was low (~10%) representation of women callers. Various explanations are offered for this, notable amongst them are the suggestion that "fewer of them call in because to do so would be putting themselves on display, claiming public attention for what they have to say, catapulting themselves onto center stage." (Tannen, 1990; p. 288). Thus, it is actually difficult to draw conclusions about what the "public interest" in STEM might be when almost half of the potential "public" are (supposedly) choosing not to participate in public call-in shows (Note: a recent change in the call-in show to a host with a less confrontational style has, reportedly, led to an increase in female callers. This has interesting implications.). This limits the utility of studying questions asked in shows such as this to discern any specifics about what STEM issues the public might be interested in.

The current, and ongoing for some time now, decline in specialist STEM/science journalists (Mooney, 2010; Wilson, 2000; Boykoff & Mansfield, 2008) offers little hope that the quality of STEM journalism will improve, suggesting that the problems detailed in past research studies of science in the news media will likely do nothing but increase. The teaching of journalism practices in journalism schools can contribute to the very problem of news media presentations of STEM/science discussed above (see Bowen, 2014) and changes in how journalists view themselves and their role in society also leads to doubt that science journalism will improve. This change in view includes a moving away from the "ideals" discussed earlier (from Deuze, 2001) to such an extent that "notions of objectivity, ethics, detachment, and even the public service ideal" are going "down the drain" (Deuze, 2001; p. 10) such that journalists do not see themselves as having responsibilities to educate the public about STEM issues, but rather that their job is one of providing an "entertaining coverage of science" (Hansen, 1994; p. 130; also Bowen, 2014)

A radio call-in show offers many opportunities for the public to develop their understanding of STEM findings and research. However, our study suggests that to accomplish that it would have to adopt a different structure than that which is used for other subjects such as politics. The for/against polarizing approach adopted in this show poorly reflects the structure of science presented by Hodson (1998, 2003) and has the potential to generate as much misunderstanding about STEM research as it does understanding of it. It is clear to us that these shows require a host and a producer who need to understand STEM research and its practices as much as they are as they need to be familiar with the needs of journalism and broadcasting (specifically the need to generate interest and engagement by their listenership). In other words, in this show the seemingly limited knowledge about STEM research of the producer and host biases and distorts the very STEM information the "scientist/guest" host is able to present to the public (a mediating process also documented elsewhere in television reporting in Bowen, 2014). We consider this to be problematic as the intent of the "scientist/guest" host is to promote STEM (and particularly scientific) literacy through his participation in the broadcast and in our view the structuring of the program (through engaging in "typical" call-in show and journalistic practices) restricts that.

Overall, we feel it is important to note that there is an interaction effect between news media and what is learned in “school science” about STEM disciplines and their practices. In general school science develops an understanding of science that the public “has” when it leaves school. Traditionally this has been a very structured or definitive view of science that little resembles science as it is engaged in by scientists (see Bowen & Rodger (2008) for how this perspective plays out when discussing global warming issues in news media forums). Thus, as adults, the public encounters the news media (re)presentations of STEM issues in news media, even in science call-in shows, which also offer a description of research and knowledge growth in STEM disciplines that also doesn’t well reflect STEM research as it is practiced. Ultimately, the news media presentation of STEM issues, through the mediating impact of the show host (at least in this case) and producer, reinforce the distorted perspective of STEM practices and claims that were originally developed in the traditional science teaching experienced by much of the public in their schools. This would seem to arise because the interests of the commercial radio station lie in “entertainment” as opposed to “education” or accurate information and in that they are treating STEM topics no differently than any other topic these days, whether it be sports or politics.

Finally, it needs to be noted that this is a study of a call-in show in a “particular” broadcast market (perhaps not reflective of others) and examines the practices of a particular host and producer. Although the general structure of the show compares with studies of other call-in shows, the actions of this particular host may well have influenced both the actions of the producer and the way in which the public callers engaged with STEM issues with the “scientist/guest” host; in other words no broad conclusions can be drawn from just this study. What was obvious, however, was that there was considerable interest in being able to ask STEM questions and have them addressed by someone who was presented as an authority. However, this interest often dealt with “personal” topics (i.e., dreams) and within that perhaps lays a lesson that higher engagement in science derives from allowing individuals to find “personal” connections to science topics (such was also concluded in a completely different setting by Cakmakci et al., 2012).

References

Almeida, M. A. (2013). Epidemics in the news: Health and hygiene in the press in periods of crisis. *Public Understanding of Science*, *22*(7), 886–902.

Bergold, J., & Thomas, S. (2012). Participatory research methods: A methodological approach in motion. *Forum: Qualitative Social Research*, *13*. http://nbn-resolving.de/urn:nbn:de:0114-fqs1201302

Bowen, G. M. (2011). Insights on the media’s practices and representations of (global warming) science: Confusing the public, educating school children? *Journal for Activist Science and Technology Education*, *3*(1), 52-79.

Bowen, G. M. (2014). Trajectories of socio-scientific issues in news media: Looking into the future. In S. Alsop & L. Bencze (Eds), *Activist Science and Technology Education*. Berlin: Springer.

Bowen, G. M., & Rodger, V. (2008). Debating global warming in media discussion forums: Strategies enacted by persistent deniers. *Canadian Journal of Environmental Education*, *13*(1), 89-106.

Boyes, E., & Stanisstreet, M. (1992). Students’ perceptions of global warming. *International Journal of Environmental Studies*, *42*, 287-300.

Boykoff, M. (2008). Media and scientific communication: A case of climate change. Geological Society, London, *Special Publications*, *305*, 11-18.

Boykoff, M., & Boykoff, J. (2004). Balance as bias: Global warming and the US prestige press. *Global Environmental Change*, *14*, 125-136.

Boykoff, M. T., & Mansfield, M. (2008). ‘*Ye Olde Hot Aire*’*: Reporting on human contributions to climate change in the UK tabloid press*. Retrieved May 5, 2008, from http://www.iop.org/EJ/article/1748-9326/3/2/024002/erl8_2_024002.html

Bucchi, M., & Mazzolini, R. G. (2003). Big science, little news: Science coverage in the Italian daily press, 1946–1997. *Public Understanding of Science*, *12*, 7-24.

Cakmakci, G., Sevindik, H., Pektas, M., Uysal, A., Kole, F., & Kavak, G. (2012). Investigating Turkish primary school students' interest in science by using their self-generated questions. *Research in Science Education*, *42*, 469-489.

Carvalho, A. (2007). Ideological cultures and media discourses on scientific knowledge: Re-reading news on climate change. *Public Understanding of Science*, *16*, 223-243.

Charney, D. (2003). Lone geniuses in popular science. *Written Communication*, *20*, 215-241.

Curtis, D. (2007). *It's not news, it's fark: How mass media tries to pass off crap as news*. New York: Gotham Books.

Davis, S. G. (1997). *Spectacular nature: Corporate culture and the Sea World experience*. Berkeley: University of California Press.

Dearing, J. W. (1995). Newspaper coverage of maverick science: Creating controversy through balancing. *Public Understanding of Science*, *4*, 341-361.

Deuze, M. (2001). Educating 'new' journalists: Challenges to the curriculum. *Journalism and Mass Communication Educator*, *56*(1), 4-17.

Dimopolous, K., & Koulaidis, V. (2003). Science and technology education for citizenship: The potential role of the press. *Science Education*, *87*(2), 241-256.

Dispensa, J. M, & Brulle, R. J. (2003). Media's social construction of environmental issues: Focus on global warming — A comparative study. *International Journal of Sociology and Social Policy*, *23*(10), 74-105.

Donald, A. (2012). *How many scientists does it take to make a discovery?* The Telegraph, 17 Sept, 2012. Retrieved February 9, 2015 from http://www.telegraph.co.uk/news/science/9548376/How-many-scientists-does-it-take-to-make-a-discovery.html

Einsiedel, E. F. (1992). Framing science and technology in the Canadian press. *Public Understanding of Science, 1*, 89-101.

Fairclough, N. (1992). *Discourse and social change*. Cambridge: Polity Press.

Halkia, K., & Mantzouridis, D. (2005). Students' views and attitudes towards the communication code used in press articles about science. *International Journal of Science Education*, *27*(12), 1395-1411.

Hansen, A. (1994). Journalistic practices and science reporting in the British press. *Public Understanding of Science*, *3*, 111-134.

Hayes, D. P. (1992). The growing inaccessibilityof science. *Nature*, *356*, 739-740.

Hodson, D. (1998). *Teaching and learning science: Towards a personalized approach*. Buckingham, UK: Open University Press.

Hodson, D. (2003). Time for action: Science education for an alternative future. *International Journal of Science Education*, *25*(6), 645-670.

Hutchby, I. (1992). The pursuit of controversy: Routine skepticism in talk on 'Talk Radio'. *Sociology*, *26*(4), 673-694.

Hutchby, I. (1995). Aspects of recipient design in expert-advice giving on call-in radio. *Discourse Processes*, *19*(2), 219-238.

Hyslop, J. H. (1899). Newspaper science. *Science, New Series*, *10*, 695-696.

Latour, B., & Woolgar, S. (2013). *Laboratory life: The construction of scientific facts*. Princeton University Press.

Lewenstein, B. V. (2001). Who produces science information for the public? In J. H. Falk (Ed.), *Free-choice* *science education: How we learn science outside of school* (pp. 21-43). New York: Teachers College Press.

Maille, M. E., Saint-Charles, J., & Lucotte, M. (2010). The gap between scientists and journalists: The case of mercury science in Québec's press. *Public Understanding of Science*, *19*(1), 70-79.

McBean, G. A., & Hengeveld, H. G. (2000). Communicating the science of climate change: A mutual challenge for scientists and educators. *Canadian Journal of Environmental Education*, *5*, 9-25.

McComas, K., & Shanahan, J. (1999). Telling stories about global climate change: Measuring the impact of narratives on issue cycles. *Communication Research*, *26*(1), 30-57.

Miller, J. D., Augenbraun, E., Schulhof, J., & Kimmel, L. G. (2006). Adult science learning from local television newscasts. *Science Communication*, *28*(2), 216-242.

Mooney, C. (2010). *Do scientists understand the public?* Cambridge, MA: American Academy of Arts and Sciences. Retrieved on 27 August 2014 from https://www.amacad.org/content/publications/pubContent.aspx?d=1101.

Myers, G. (1996). Out of the laboratory and down to the bay: Writing in science and technology studies. *Written Communication*, *13*(1), 5-43.

Neidhardt, F. (1993). The public as a communication system. *Public Understanding of Science*, *2*, 339-350.

Nelkin, D. (1995). *Selling science: How the press covers science and technology*. New York: Freeman.

Pellechia, M. G. (1997). Trends in science coverage: a content analysis of three US newspapers. *Public Understanding of Science*, *6*, 49-68.

Pettersen, S. (2005). Norwegian health journalists' ability to report on health research: A concern to science education? *Nordic Studies in Science Education*, *1*, 5-16.

Scanlan, C. (2003). *Writing from the top down: Pros and cons of the inverted pyramid*. Accessed 27 August 2014 from http://www.poynter.org/news/media-innovation/12754/writing-from-the-top-down-pros-and-cons-of-the-inverted-pyramid/.

Schibeci, R. A. (1990). Public knowledge and perceptions of science and technology. *Bulletin of Science and Technology Society*, *10*, 86-92.

Shanahan, J., & Good, J. (2000). Heat and hot air: Influence of local temperature on journalists' coverage of global warming. *Public Understanding of Science*, *9*, 285-295.

Shapin, S. (1992). Why the public ought to understand science-in-the-making. *Public Understanding of Science*, *1*(1), 27- 30.

Somerville, R. C. J., &Hassol, S. J. (2011). Communicating the science of climate change. *Physics Today*, 48-53.

Tankel, J. D. (1998). Reconceptualizing call-in talk radio as listening. *Journal of Radio Studies*, *5*(1), 36-48.

Tannen, R. (1990). *You just don't understand: Women and men in conversation*. New York: Ballintine Books.

Thornborrow, J. (2001). Questions, control and the organization of talk in calls to a radio phone-in. *Discourse Studies*, *3*(1), 119-143.

Wilson, K. M. (1995). Drought, debate, and uncertainty: measuring reporters' knowledge and ignorance about climate change. *Public Understanding of Science*, *9*, 1-13.

Zurawski, R. (2011). *Media mediocrity — waging war against science: How the television makes us stooped!* Halifax: Fernwood Publishing.

Journal of Research in STEM Education
ISSN:2149-8504 (online)
Vol 1, No 2, December 2015, PP 142-155

RESEARCH REPORT

Testing an Adapted Model of Social Cognitive Career Theory: Findings and Implications for a Self-Selected, Diverse Middle-School Sample

Christian E Mueller[1], Alfred L Hall , Danielle Z Miro

University of Memphis, USA

Abstract: *We tested an adapted version of social-cognitive career theory (SCCT; Lent et al., 1994, 2000) with a self-selected, diverse sample of middle-school students attending a Saturday STEM Academy asking, "Is SCCT valid for examining career choice goal-intentions among a sample of students already expressing interest in math and science-related subjects and careers?" According to SCCT, choosing a STEM-related career involves the complex interplay of personal and contextual factors, many of which become increasingly salient during the middle-school years. There is reason to believe that SCCT may function differently for students who are self-selected, such as those found in the present sample. Main findings in the full regression model showed that math/science motivation (T1), family support for engineering (T1), outcome expectancies (T2), and interest (T2) were significant predictors of (T2) goal intentions; whereas self-efficacy was non-significant as has been shown in much previous research. Relatedly, we found several measurement issues with the SCCT variables among this sample, thus partially answering the larger research question. Implications of the present findings and suggestions for future research are discussed in the context of the career-choice literature, theoretical and practical implications of SCCT, and relatedly, possible measurement issues arising from using SCCT with self-selected, middle-school samples.*

Keywords: Social cognitive career theory, Self-selected, Middle school, STEM

Social-cognitive career theory (SCCT) (Lent, Brown, & Hackett, 1994, 2000) has gained prominence as a framework for examining career-related choices across numerous populations, including science, technology, engineering, and mathematics (STEM) careers (e.g., Wang, 2013a). SCCT proposes that career development and career-related choices are best explored as emerging from the complex interaction between individual factors (i.e., self-efficacy, interests, outcome expectancies, and goals), background/personal factors (e.g., ethnicity, predispositions, and gender), and prior learning and achievement—all viewed through the lens of social-cognitive theory (Bandura, 1986). Further, SCCT places emphasis on the role of distal and proximal influences on career decision-making, which are assumed to have direct and indirect influences (Lent, Lopez, Lopez, & Sheu, 2008). For example, among middle- and high-school aged students, previous investigations have found that both teacher and parental support have positive associations with career decision-making, self-efficacy, and prediction of future goals (Gushue & Whiston, 2006; Zebrak, Le, Boekeloo, & Wang, 2013b). Yet, despite its prominence, Lent & Brown (2006) acknowledged that SCCT has inherent conceptual and measurement issues that must be addressed if accurate hypothesis testing is to occur with applications of their theory. In the present study, we hypothesized that the self-selective nature of the present sample (i.e., students attending a Saturday STEM academy) might impact interrelationships among SCCT variables and that social cognitive variables

[1] Corresponding author. Department of Counseling, Educational Psychology and Research, 100 Ball Hall, University of Memphis, Memphis, TN 38152. E-mail: cemuellr@memphis.edu

Mueller, C. E., Hall, A. L., & Miro, D. Z. (2015). Testing an adapted model of social cognitive career theory: Findings and implications for a self-selected, diverse middle-school sample. *Journal of Research in STEM Education*, *1*(2), 142-155.

may have varying degrees of influence within this sample.

Concurrently, much research has applied SCCT within the STEM fields (e.g., Wang, 2013a), with diverse student populations (e.g., Turner & Lapin, 2002), with middle- and high-school populations (e.g., Jiang & Zhang, 2012; Wang, 2013b), and among all three (e.g., Navarro, Flores, & Worthington, 2007). These studies are important and have relevance for the present study for two reasons. First, SCCT has been a surprisingly robust predictor of career-related choices in the STEM fields. However, to date, few studies exist that have examined SCCT in the context of middle-school aged students already showing high levels of interest or motivation in science and math, which are precursors to later STEM study. This may be problematic because, as Lent and Brown (2006) acknowledged, "When most people's ratings 'top the chart,' this creates ceiling effects, range restriction, and negative skew in the distribution of […] scores" (p. 25). Here, they were referring specifically to self-efficacy, however, there is reason to believe that these issues may exist for other SCCT factors, as well. Second, increasingly, researchers are turning attention to understanding the non-cognitive factors that help sustain student interest, motivation, and persistence along the STEM educational pipeline (e.g., Andersen & Ward, 2013; Wang & Degol, 2013). Many of these non-cognitive factors represent core constructs in the SCCT model. For example, Lent et al. (2015) found that, among other things, self-efficacy was a significant predictor of academic persistence among undergraduate engineering students, even across gender and racial/ethnic lines. As the nation increases in ethnic and gender diversity, understanding how to retain the best and brightest students all along the STEM educational pipeline becomes increasingly important in sustaining U.S. global competiveness in STEM fields.

Theoretical Framework and Review of Relevant Literature

Social Cognitive Career Theory

As previously discussed, SCCT assumes that individual career choices and goal intentions are best explained through the interaction of numerous factors, including individual, background/personal, and other proximal/distal influences (see Figure 1). The interplay among environmental factors, social factors, and personal factors, in concert directly and indirectly shape outcome expectations and increase motivation in determining career and education paths (Loera, Nakamoto, Oh, & Rueda, 2013). According to Lent, Brown, & Hackett (2002), SCCT is rooted in Albert Bandura's social cognitive theory (Bandura, 1986), which posits that individual actions, choices, and goals are influenced by personal attributes (e.g., self-efficacy), external environmental conditions, and overt behavior. In addition, "Within this triadic system, people become both 'products and producers of their environment' (Wood & Bandura, 1989, p. 362), with the potential for self-regulation" (p. 261).

Career choice, then, is understood within a framework emphasizing individual choice as emanating from the interplay between person factors, such as self-efficacy, outcome expectations, and interests; and other factors, including learned experiences and previous barriers and supports in the environment, such as higher or lower levels of parental or teacher support (Figure 1; Lent et al., 1994). Specifically, the social-cognitive constructs in SCCT include self-efficacy, an individual's perceived competence in a specific area that may predict effort expenditure or persistence (Bandura, 1977); outcome expectations, perceived consequences or successes about performance on a task (Lent et al., 1994); and goal orientation, an individual's determination to engage in a task or activity (Bandura, 1986). Further, person inputs include predispositions, gender, race/ethnicity, and health status that are interrelated to background or contextual factors, such as family or parental support (Byars-Winston & Fouad, 2008). Distal contextual factors can be viewed as perceived barriers or supports to an interest, goal, or action (Lent et al., 1994, 2000). Moreover, learned experiences and the environment shape an individual's self-efficacy, outcome expectations, and goals. SCCT also proposes a bi-directional model between person inputs and other person inputs, which directly influence learning experiences (Navarro et al., 2007).

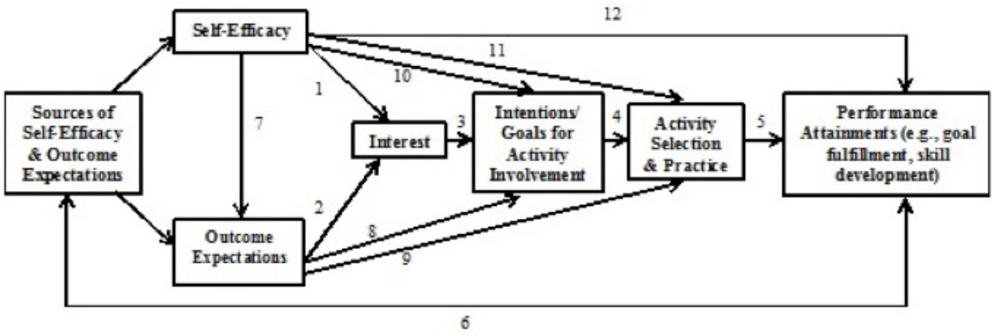

Figure 1. *Social Cognitive Career Theory Reprinted from "Toward a Unifying Social Cognitive Theory of Career/Academic Interest, Choice, and Performance," by R. W. Lent, S. D., Brown, & G. Hackett, 1994, Journal of Vocational Behavior, 45, 79-122. Reprinted with permission.*

SCCT and STEM

With respect to STEM, SCCT has proven to be very robust in predicting career goals and intentions among STEM-related fields. Lent, Singley, Sheu, Schmidt, and Schmidt (2007) found that in predicting engineering goals, progress factors such as outcome expectations, self-efficacy, and environmental supports were influential to career decision choice. These findings suggest there is a need to examine how students monitor and frame their STEM goals, which environmental supports or resources they use, and their STEM self-efficacy. Furthermore, interplay between environmental factors (e.g., STEM labs), social factors (e.g., supportive family, teachers, or engineering professionals), and personal factors (e.g., student satisfaction) shape outcome expectations and increase motivation in determining career and educational paths (Loera et al., 2013).

In determining college majors, SCCT has allowed researchers to examine pathways of persistence as well as deterministic factors that result in individuals majoring in STEM (Lent et al., 2008). Nauta, Epperson, and Kahn (1998) found that individuals with higher perceived abilities and self-efficacy had increased aspirations to enter a STEM career and suggested that self-efficacy mediated the relationship between abilities and STEM-career intention. Of relevance for the present study, many of these studies have focused on students who were already majoring in STEM, rather than those intending to major in STEM, thus leaving a void in understanding the possible supports and barriers for entrance into or sustaining interest along the STEM educational pipeline (Wang, 2013a). SCCT is well-suited to fill this gap via examination of STEM-related social-cognitive variables in populations of highly efficacious and motivated students. In particular, as more STEM-focused curricula and programs are developed, implemented, and researched for their effectiveness in sustaining student interest, motivation, and persistence along the STEM educational pipeline, it is imperative that we gain understanding of these social-cognitive variables when student STEM interest is inchoate, such as coinciding with the middle-school years or before (Maltese & Tai, 2010).

SCCT and Middle/High School Student Populations

Identifying barriers and supports to STEM majors is vital, yet when viewing the entirety of the STEM educational pipeline and persistence therein, there is much research showing that interest in STEM domains (i.e., math and science) forms during the middle-school years, or before. For example, Maltese and Tai (2010) found that retrospective accounts by scientists and engineers reported developing an interest in science well before middle school, which in part, carried them all the way through the STEM education pipeline. Interestingly, Maltese and Tai found that for women their interest in science was related to activities at school, whereas for men, interest was more related to self-initiated activities. In addition, Lindley (2005) suggested that contextual barriers faced by adolescents is a key component to career decision-making and interest formation in STEM, which is well suited to be studied under an SCCT framework.

SCCT posits that perceptions of the self, tasks, and performance on such tasks is influential to adolescent career development and defining later goals to develop ability or demonstrate ability in a subject area (Schultheiss, 2008). To which, in a qualitative study of young adolescents, Shoffner, Newsome, Minton, and Morris (2015) found that careers in STEM were associated with negative outcome expectations, particularly among females. Further, Cantrell, Pekcan, Itani, and Velasquez-Bryant (2006) found that the influence of contextual variables, such as race/ethnicity and socioeconomic status, decreased with middle school students when design science was implemented, suggesting that SCCT is an appropriate framework to examine STEM-related issues with younger populations.

Purpose of the Present Study and Research Questions

Collectively, all of the above research suggests a continued need for understanding how SCCT functions in predicting STEM career decision-making across varied and diverse populations. As such, in the present study we were interested in exploring the larger research question "Is SCCT valid for examining career choice goal-intentions among a sample of students already expressing interest in math and science-related subjects and careers?" To answer this larger question we answered three distinct, but interrelated questions:

1. How strongly do personal and environmental antecedents relate to the social-cognitive factors of self-efficacy, outcome expectancy, interest, and goals with our sample?
    - Hypothesis: We hypothesized significant positive relationships between personal and contextual antecedents and SCCT social-cognitive variables.
2. How strongly do the social-cognitive variables of self-efficacy, outcome expectancy, and interest collectively predict career choice goal-intentions with our sample?
    - Hypothesis: We hypothesized significant positive relationships between our main predictors of self-efficacy, outcome expectancy, and interest, and our outcome variable of career choice goal-intentions.
3. Are there potential measurement issues presented in using SCCT variables (e.g., ceiling effects) with our self-selected sample?
    - Hypothesis: Although we were not certain about the exact nature of potential measurement issues, we hypothesized that the self-selected nature of the present sample would present some issues that needed to be addressed before answering research questions one and two (Lent & Brown, 2006).

Methodology

Participants and Procedure

Participants (N = 186) included sixth- through eighth-graders attending a Saturday STEM Academy program in a Mid-southern city during the spring 2015 semester. This program was designed to provide an opportunity for middle-school students to explore STEM-based projects and activities to help develop student interest in STEM and to potentially encourage them to enroll in and complete a high school STEM program of study. Students came from several schools within a larger urban district and chose to attend the academy after a district-wide solicitation effort. There were three Saturday Academy sessions and only students who completed both pre- and post-surveys (n = 104) were included for subsequent analysis. Table 1 presents all descriptive statistics for the sample, and shows that the sample was predominantly female (52.9%), African-American (45.2%), and comprised of sixth-graders (46.2%). IRB approval was obtained prior to pre and post-test data collection.

Table 1
Descriptive Statistics of Sample

Number		104
Gender	Male	49 (47.1%)
	Female	55 (52.9%)
Ethnicity	African/Black American	47 (45.2%)
	American Indian/Alaskan Native1	1 (1.0%)
	Asian & Pacific American	18 (17.3%)
	Latina/Latino/Hispanic American	12 (11.5%)
	White American	14 (13.5%)
	Other	12 (11.5%)
Grade	Sixth	48 (46.2%)
	Seventh	33 (31.7%)
	Eighth	23 (22.1%)

Instrument (Pre-Test)

The pre-test instrument was administered at the beginning of the program and included demographic questions, questions about attitudes toward STEM, and perceived barriers or supports toward STEM goal pursuit. These measures were initially developed and validated by the Advancing Women and Men in Engineering (AWE) program (2008). In the present study, some items were removed from consideration for use on the final scales and subsequent analysis when they violated assumptions of normality (i.e., too much skewness or kurtosis) or when removing items improved the overall reliability estimates for the present sample. Items that were retained were kept in their original form. Details about each of the scales that were administered on the pre-test measure are provided next.

Demographics. Gender, ethnicity, and grade-level served as the main demographic correlates.

Motivation for math/science ($\alpha = .70$, 8 items). Eight of the original 14 items were retained for the adapted scale and used in the present study, which is designed to measure students' motivation for math and science. Respondents were asked to rate on a four-point scale— strongly disagree (1) to strongly agree (4)—their agreement to items such as, "I look forward to science class in school" and "I like learning how things work."

Math/science future value. Math/science future value was assessed by a single-item asking students to rate on a three-point scale—Not Important, Somewhat Important, Very Important—"How important is it to you to do future work that allows you to use math, computer, engineering or science skills?"

Teacher-support/family-support/community-other support for engineering. Three single-item statements were used to measure students' perceived teacher, family, and community/other support for engineering. Respondents were asked to initially respond to "Has anyone talked to you about becoming an engineer?" If students responded "yes," they were prompted to select among teacher, family, and community/other; responses were then dichotomized to yes (1) or no (0). Each item was treated as an independent factor in subsequent regression analyses.

Instrument (Post-Test)

The post-test measure was administered at the conclusion of the program. Scales were initially developed by Fouad, Smith, and Enochs (1997) and Fouad and Smith (1996) and based on SCCT (Lent et al., 1994, 2000). All of the post-test measures had to be adapted in the same manner as the pre-test measures, including dropping items that violated assumptions of normality and removing individual items when overall reliability estimates were improved. All items that were retained on each scale were kept in their original form. Details about each of the scales that were administered on the post-test measure are provided next.

Math/science interest (α = .76, 13 items). Thirteen of the original 20 items were retained for the adapted scale. The original scale developed by Fouad and her colleagues was designed to measure students' interest in math and science-related activities. Respondents were asked to rate each item as to "How interested you are in these things…" according to dislike, not sure, and like. Examples of retained items include "…Visiting a science museum," "…Using a calculator," and "…Solving computer problems."

Math/science efficacy (α = .80, 7 items). Seven of the original 12 items were retained for the adapted scale. The original scale developed by Fouad and her colleagues was designed to measure students' confidence in their ability to do a series of math and science-related activities. Respondents were asked to rate each item as to how confident they were in terms of their ability on items ranging from "Determine the amount of sales tax on clothes I want to buy" to "Design and describe a science experiment that I want to do." Respondents rated their confidence/ability levels from very low ability (1) to very high ability (5).

Math/science outcome expectations (α = .71, 4 items). Four of the original seven items were retained for the adapted scale. The original scale developed by Fouad and her colleagues was designed to measure students' beliefs "Regarding the consequences of their potential mathematics and science-related course activities and achievements" (Navarro et al., 2007, p. 325). Respondents were asked to rate how much they agree or disagree with items such as "If I learn math well, I will be able to do lots of different types of careers" and "If I do well in science, then I will be better prepared to go to college." Respondents rated these items on a five-point scale, ranging from strongly disagree (1) to strongly agree (5).

Math/science goal intentions (α = .74, 6 items). All original scale items were retained and were designed to measure middle-school students' future goals to pursue math and science-related courses and careers (Navarro et al., 2007). Respondents were asked to rate their agreement and/or disagreement on a five-point scale—strongly disagree (1) to strongly agree (5)— with items such as "I plan to take math classes in high school" and "I intend to enter a career that will use science." Respondents rated these items on a five-point scale, ranging from strongly disagree (1) to strongly agree (5).

Data Analyses

In the present study, we employed a longitudinal research design to explore the larger research question of "Is SCCT valid for examining career choice goal-intentions among a sample of students already expressing interest in math and science-related subjects and careers?" In order to answer the larger research question, we conducted two sets of regression analyses to explore smaller, but related questions. With respect to the first question, "How do personal and contextual antecedents influence social cognitive variables (interest, self-efficacy, outcome expectancy, & goals) in SCCT with a self-selected, diverse middle-school sample?" we conducted four separate hierarchical regressions. For each regression, the demographic correlates of gender, grade and ethnicity were entered at step 1 of each analysis. At step 2, personal factors (motivation for math/science & math/science future value) and contextual factors (teacher, family, &/or community/other support for engineering) were added as additional predictors for each of the four outcomes.

With respect to the second question, "How well does an adapted version of SCCT work in examining career choice goal intentions for this sample?" we conducted a final regression analysis in which the same demographics were entered at step 1, the same personal and contextual factors were entered at step 2, and as an additional step to test the entire model, the social-cognitive variables of interest, outcome expectancy, and self-efficacy were entered at step 3. Goal intentions served as the outcome variable in the model (see Figure 1).

Table 2
Means (SDs) for Pre- and Post-Test Measures

Variable (Scores 1 to 5)	N (observations total)	Group M (SD)
Pre-Test Motivation for Science/Technology Scale	102	3.36 (.38)
Math/Science Future Value Scalea	104	2.66 (.53)
Post-Test for Math/Science Interest Scale	104	2.52 (.34)
Post-Test for Mat/Science Efficacy Scale	104	3.95 (.73)
Post-Test for Math/Science Outcome Expectancy Scale	102	4.46 (.54)
Post-Test for Math/Science Goals Scale	102	4.25 (.62)

Results

For sake of clarity, results are presented and organized according to how they relate to the primary research questions presented earlier. Table 2 presents means and standard deviations for the entire sample across all pre and post-test measures. Results for the first two research questions are presented next. Discussion around research question three regarding measurement issues is addressed in the Discussion portion of the study.

RQ1: How strongly do personal and environmental antecedents relate to the social-cognitive factors of self-efficacy, outcome expectancy, interest and goals with our sample?

Regressions 1 and 2: Non-significant findings. Non-significant findings were obtained in the hierarchical regressions conducted with self-efficacy and outcome expectancy as main outcomes. Neither the demographic correlates at step 1, consisting of gender, grade, and ethnicity, nor the personal and contextual antecedents, consisting of motivation for math/science, math/science future value and teacher-support, family-support, and community-other support for engineering at step 2, had a significant effect when regressed onto self-efficacy and outcome expectancy. Table 3 presents full findings from these two models.

Regressions 3 and 4: Significant findings. Conversely, when examining interest and goals as the main outcomes, both of the hierarchical regression models were significant. In the third regression, predicting time 2 interest in math/science as the main outcome, the demographic variables of gender, grade, and ethnicity explained a significant proportion of the variance with an adjusted R2 = .06, $p < .05$; however, only gender ($\beta = -.272$, $p < .01$) was significant. This implies that females have significantly less interest in math/science than males within this group. At step 2, in adding the time 1 personal and contextual antecedents, the model again explained a significant proportion of the variance in time 2 interest with an adjusted R2 = .31, $p < .001$. In the presence of the personal and contextual antecedents, all of the demographic variables including gender, became non-significant. In this case, only the time 1 personal factors of motivation for math/science ($\beta = .429$, $p < .001$) and math/science future value ($\beta = .179$, $p < .05$) were significant positive predictors of time 2 math/science interest, therefore showing that higher levels of motivation and valuing of math/science at time 1 led to higher levels of math/science interest at time 2 for this self-selected sample. The contextual antecedents of teacher-support, family-support, and community-other support for engineering were not significant in predicting interest.

Last, in the fourth regression predicting time 2 goals as the main outcome, the model at step 1 with all of the demographic variables was non-significant. However, the model at step 2, with the time 1 personal and contextual variables added, explained a significant proportion of the variance in time 2 goals with an adjusted R2 = .17, $p < .001$. Furthermore, in addition to the personal variable of motivation for math/science ($\beta = .362$, $p < .001$), the contextual variable of family-support for engineering ($\beta = .233$, $p < .05$) was a significant positive predictor of time 2 goals. In this case, having higher levels of motivation for math/science and higher perceived levels of family support for pursuing engineering as a career, impacted future goals about math/science at time 2. Table 3 presents all results from these significant models.

Table 3
RQ1 Results: Hierarchical Regression Analyses Predicting Post-Math/Science Interest, Efficacy, Outcome Expectancies, & Goals

	Post-math/science efficacy		Post-math/science outcome expectancies		Post-math/science Interest		Post-math/science goals	
Variable	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Controls								
Gender	-.096	-.015	-.197*	-.139	-.272***	-.133	-.084	.022
Ethnicity	-.034	-.011	-.126	-.087	-.998	-.035	.042	.116
Grade	.005	-.011	-.003	-.026	-.263	-.039	.051	.021
Direct Effects								
Motivation for Science/ Technology Scale		.244**		.187*		.429****		.362****
Math/Science Future Value Scale		.108		.138		.179**		.109
Teacher Support Engineering Scale		-.038		-.010		-.003		-.041
Family Support Engineering Scale		.082		.093		.156*		.233**
Community/Other Support Engineering Scale		.025		-.010		-.013		-.095
F total	.370		2.117		3.171**		.421	
R^2	-.019	.025	.033	.063	.061	.314	-.018	.166
Δ F		1.891		1.626		8.252****		5.244****
Δ R^2		.091		.077		.280		.221

* p < .10
** p < .05
***p < .01
****p < .001

RQ2: How strongly do social-cognitive variables of self-efficacy, outcome expectancy and interest collectively predict career choice goal-intentions with our sample?

As can be seen, the full model was not significant at step 1 (demographics only), but was significant at step 2 with an adjusted R2 = .17, p < .001 with the personal and contextual factors added, and an adjusted R2 = .51, p < .001 at step 3 with the social-cognitive variables added. This indicates that the social-cognitive variables were robust predictors of career goal intentions, as they explained an additional 33% of the variance, even after controlling for the other variables. As seen in Table 4, gender and ethnicity were non-significant at all three steps in the model, thus indicating that neither gender nor ethnicity showed any effect for this self-selected sample. Interestingly, grade was not significant at the first two steps of the regression, but did retain significance at step 3, demonstrating there were significant differences across grades for career goal intentions. Further, these differences only manifested in the presence of the social-cognitive variables, thus indicating that perhaps some of the variability explained by grade at step 3 was more a reflection of a statistical anomaly rather than being indicative of a meaningful finding (see Shieh, 2006, for discussion on suppression in multiple regression).

With respect to the personal and contextual variables in the full regression model, there were some interesting findings as well. At step 2, time 1 math/science motivation (β = .362, p < .001) and time 1 family support for engineering (β = .233, p < .05) were significant positive predictors of time 2 career goal intentions. At step 3, in the presence of the social-cognitive variables, both of these became non-significant; however, family support for engineering did still approach significance (β = .140, p = .08). Among the social-cognitive predictors, time 2 interest (β = .223, p < .05) and time 2 outcome expectancy (β = .497, p < .001) retained significance, but contrary to previous findings in the literature, self-efficacy was not significant. Table 4 presents the full findings of this final regression.

Table 4
RQ2 Results: Testing Adapted Version of SCCT with Goals as Main Outcome

Covariates		Model 1		
	Gender	-.84		
	Grade	.042		
	Ethnicity	.051		
	F total	.421		
	R^2	.013		
Covariates			Model 2	
	Gender		.022	
	Grade		.116	
	Ethnicity		.021	
	Motivation for Science/Technology Scale		.362****	
	Math/Science Future Value Scale		.109	
	Teacher Support Engineering Scale		-.041	
	Family Support Engineering Scale		.233**	
	Community Support Engineering Scale		-.095	
	Δ F		5.244****	
	R^2		.166	
	Δ R^2		.221	
Covariates				Model 3
	Gender			.124
	Grade			.164**
	Ethnicity			.049
	Motivation for Science/Technology Scale			.141
	Math/Science Future Value Scale			-.006
	Teacher Support Engineering Scale			-.026
	Family Support Engineering Scale			.140*
	Community Support Engineering Scale			-.092
Direct Effects				
	Post Math/Science Self-Efficacy Scale			.223**
	Post Math/Science Outcome Expectations Scale			.104
	Post Math/Science Interest Scale			.497****
	Δ F			22.612****
	R^2			.513
	Δ R^2			.334

* p < .10
** p < .05
***p < .01
****p < .001

Discusson

Situated in the literature surrounding career goal theory, social-cognitive theory, and STEM-education persistence, the present study sought to explore how robustly SCCT predicted career goal intentions among a sample of middle-school students already showing high levels of interest in math and science. Given the difficulty in examining all aspects of the model simultaneously, we sought to answer the larger question through exploration of narrower, but related questions. Findings from the present study add to the growing SCCT literature in distinct and important ways—our general conclusion is that indeed, SCCT is a robust way to examine career goal intentions among self-selected students. As with any good theory, however, the story is not always as straightforward as it seems. In the next few sections, we explore the implications of select findings as they relate to previous research; they are organized and presented in relation to the three research questions posed at the beginning of the study.

RQ1: How strongly do personal and environmental antecedents relate to the social-cognitive factors of self-efficacy, outcome expectancy, interest and goals with our sample?

Not surprisingly, gender was found to be a significant predictor of student interest and outcome expectancy, which is supported by previous findings (Lent, Sheu, Gloster, & Wilkins, 2010). However, of note, the impact of gender lessened and was mitigated by personal factors such as math/science motivation and math/science future value, two antecedents appearing to help females in the present sample. Referring back to the findings of Maltese and Tai (2010), it may be important to consider how personal factors, gender, and school-based interventions intersect, as Maltese and Tai found that school-based activities were listed as a significant influence on developing math/science interest for females. It bears worth mentioning again that the females in the present study were from a self-selected group. To gain a broader understanding of how motivation and valuing influence decision-making along the STEM pipeline, future researchers should test these motivational relationships across broader samples and at different points along the pipeline (Anderen & Ward, 2013; Perez, Cromley, & Kaplan, 2014; Wang & Degol, 2013). Nevertheless, it is encouraging that once again, positive outcomes seem to result from an increased interest in math and science, and that interventions can be designed to increase interest all along the STEM education pipeline (Hidi & Renninger, 2006).

Interestingly, but again not surprisingly, family support for engineering was the only significant contextual predictor of future goals for our sample. In essence, for students already showing high levels of interest in math and science, perceived family support for engineering was significant in predicting future goals, which points to the importance of family in helping students persist along the STEM pipeline. It is also encouraging that neither gender nor ethnicity was significant with respect to goals—these findings are all the more encouraging given the longitudinal design of the study. Previous literature has also shown this to be true; for example, Ferry, Fouad, and Smith (2000) found that parental encouragement directly impacted their child's math/science learning experiences. Although we used the term "support" differently than in previous studies, and there are certainly more comprehensive ways to capture the true nature of what is meant by support, we are nonetheless encouraged that family support, in any form, appears important for career decision-making along the STEM education pipeline for this group.

RQ2: How strongly do social-cognitive variables of self-efficacy, outcome expectancy and interest predict career choice goal-intentions with our sample?

With regards to SCCT as a theoretical framework for understanding career choice goal intentions among self-selected middle schoolers, we found that the broader model was indeed robust, with the several significant social-cognitive predictors. Interestingly, self-efficacy was a non-significant factor in our model, which further highlights the complicated nature of career choice. Typically, self-efficacy is found to be a significant predictor in the SCCT model (Lent et al., 2015), however, it seems natural that students with strong STEM self-efficacy are more likely to pursue STEM-related programs (Byars-Winston & Fouad, 2008), such as was the case with the present sample attending a Saturday STEM academy. In the present sample, not only was there a ceiling effect for self-efficacy, but skewed scores also led to violations of normality on many items. This fact is further explored in the next section. Perhaps of more relevance for the present study are ways in which high levels of interest and self-efficacy can be sustained all along the STEM education pipeline. We refer the reader to related literature on the developmental nature of interest and self-efficacy (e.g., Britner & Pajares, 2006; Hidi & Renninger, 2006; & Usher & Pajares, 2008).

Present findings also pose the question as to if there are alternative SCCT models that may better help understand this population and better fit the data. Our results reflected those of Lent et al. (2005) that found that although their measurement model provided good fit to their data, structurally, it did not ideally fit their data, suggesting the theoretical model had some inadequacies in assessing social-cognitive relationships. Clearly SEM models would need to be conducted in order to test simultaneous SCCT relationships with the present sample; however, findings from present regression models do at least, in part, indicate the nature of these relationships. While our findings extend support for SCCT as a whole with this sample, in that math/science interests and outcome expectancies accounted for 50% of the variance in goal intentions, developing alternative frameworks may assist in improving understanding of STEM-related career choices among highly-efficacious students. Perhaps, future models should place more emphasis on contextual factors and their direct or indirect pathways to goal intentions (Schaefers, Epperson, & Nauta, 1997).

Methodological Implications

With respect to the methodological implications of using SCCT with a self-selected sample, we found main issues around violations of normality (e.g., excessive skewness or kurtosis), and in some cases, correspondingly low reliability estimates. Our findings supported Navarro et al.'s (2007) call for developing instruments that are domain specific and utilize many measures to capture a broader range of SCCT's constructs. In addressing measurement and conceptualization issues of SCCT constructs, Lent and Brown (2006) suggested that researchers design instruments that mirror the operationalization and context of the construct. Additionally, for educational researchers, further study on predictors and contextual factors that affect student interest, participation, and self-efficacy in STEM (especially for middle school populations of students from disadvantaged backgrounds) is needed. In the present study, we were left with dropping some items due to violations of normality. We ultimately decided against the use of transformational procedures as these procedures are often difficult to interpret, and our target audience was not just educational researchers. In future, we would suggest adapting the items more specifically to population and context as suggested by Navarro et al. (2007), and Lent and Brown (2006).

Limitations

One of the limitations of this study was the non-normality of data within this sample. Because of the nature of the sample, students who attended the STEM Saturday Academy likely had high math/science interest and self-efficacy prior to data collection, which certainly impacted the predictive utility of these constructs in subsequent analyses. We also did not examine differences of math or science beliefs individually, but rather conjointly. This may have impeded some of the findings, such as females being more likely to have negative attitudes towards mathematics than boys (Yee & Eccles, 1998).

Implications

This research study has implications for education researchers and STEM educators wanting to understand sources/antecedents for these social-cognitive factors—where do they come from and how do we help develop them? Before this question can be answered, more research measurements and analyses are needed to assess their effect on student development. Future investigations should also test the effectiveness of such questions via hypothesis testing and examining the effectiveness of such STEM programs.

Moreover, among self-selected students, sustaining their STEM motivation is crucial for continuity in the STEM education pipeline and to ensure a viable STEM workforce in the future. For educational researchers, further study on predictors and contextual factors that affect student interest, outcome expectancy, and self-efficacy in STEM (especially for middle-school populations of students from disadvantaged backgrounds) is needed. These findings also suggest that as STEM educators and program developers work to provide meaningful opportunities to develop long-term goal intentions in STEM, student interest, grade level, and outcome expectancies must also be addressed in the recruitment, support, and retention efforts of various STEM programs. Developing authentic learning experiences in STEM that are related to real-life application may address students' perceptions of STEM and their career-related choice decisions.

Author Note

This research was supported in part by a grant from the National Science Foundation (DRL 1311876).

References

Advancing Women and Men in Engineering (AWE) & National Girls Collaborative Project. (2008). *Pre-Activity Survey for Middle School-Aged Participants-Engineering [Measurement instrument]*. Retrieved from: AWEonline.org.

Anderson, L., & Ward, T. J. (2013). Expectancy-value models for the STEM persistence plans of ninth-grade, high-ability students: A comparison between Black, Hispanic, and White students. *Science Education*, *20*, 1-27.

Bandura, A. (1986). *Social foundations of thoughts and action: A social cognitive theory.* Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, *84*, 191-215.

Byars-Winston, A. M., & Fouad, N. A. (2008). Math and science social cognitive variables in college students: Contributions of contextual factors in predicting goals. *Journal of Career Assessment*, *16*, 425-440.

Britner, S. L., & Pajares, F. (2006). Sources of science self-efficacy beliefs in middle school students. *Journal of Research in Science Teaching*, *43*, 485-499.

Cantrell, P., Pekcan, G., Itani, A., & Velasquez-Bryant, N. (2006). The effects of engineering modules on student learning in middle school classrooms. *Journal of Engineering Education*, *95*, 301-309.

Ferry, T. R., Fouad, N. A., & Smith, P. L. (2000). The role of family context in a social cognitive model for career-related choice behavior: A math and science perspective. *Journal of Vocational Behavior*, *57*, 348-364.

Fouad, N. A., & Smith, P. L. (1996). A test of a social cognitive model for middle school students: Math and science. *Journal of Counseling Psychology*, *43*, 338-346.

Fouad, N. A., Smith, P. L., & Enochs, L. (1997). Reliability and validity evidence for the Middle School Self-Efficacy Scale. *Measurement and Evaluation In Counseling And Development*, *30*, 17-31.

Gushue, G. V., & Whiston, M. L. (2006). The relationship among support, ethnic identity, career decision self-efficacy, and outcome expectations in African American high school students. *Journal of Career Development*, 33, 112-124.

Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist*, *41*, 111-127.

Jiang, Z. P., & Zhang, Z. R. (2012). Using social cognitive career theory to predict the academic interests and goals of Chinese middle vocational-technical school students. *Public Personnel Management*, *41*, 59-68.

Lent, R. W., & Brown, S. D. (2006). On conceptualizing and assessing social career constructs in career research: A measurement guide. *Journal of Career Assessment*, *14*, 12-35.

Lent, R. W., Brown, S. D., & Hackett, G. (2002). *Social cognitive career theory in D. Braun and Associates (Eds.) Career choice and development* (pp. 255-311). San Francisco, CA: Jossey-Bass.

Lent, R. W., Brown, S. D., & Hackett, G. (2000). Contextual supports and barriers to career choice: A social cognitive analysis. *Journal of Counseling Psychology*, *47*, 36-49.

Lent, R. W., Brown, S. D., & Hackett, G. (1994). Toward a unifying social cognitive theory of career and academic interest, choice, and performance. *Journal of Vocational Behavior*, *45*, 79-122.

Lent, R. W., Lopez, A. M., Lopez, F. G., & Sheu, H. B. (2008). Social cognitive career theory and the prediction of interests and choice goals in the computing disciplines. *Journal of Vocational Behavior*, *73*, 52-62.

Lent, R. W., Miller, M. J., Smith, P. E., Watford, B. A., Hui, K., & Lim, R. H. (2015). Social cognitive model of adjustment to engineering majors: Longitudinal test across gender and race/ethnicity. *Journal of Vocational Behavior*, *86*, 77-85.

Lent, R. W., Sheu, H. B., Gloster, C. S., & Wilkins, G. (2010). Longitudinal test of the social cognitive model of choice in engineering students at historically Black universities. *Journal of Vocational Behavior*, *76*, 387-394.

Lent, R. W., Singley, D., Sheu, H. B., Schmidt, J. A., & Schmidt, L. C. (2007). Relation of social-cognitive factors to academic satisfaction in engineering students. *Journal of Career Assessment*, *15*, 87-97.

Lindley, L. D. (2005). Perceived barriers to career development in the context of social cognitive career theory. *Journal of Career Assessment*, *13*, 271-287.

Loera, G., Nakamoto, J., Oh, Y. J., & Rueda, R. (2013). Factors that promote motivation and academic engagement in a career technical education context. *Career and Technical Education Research*, *38*, 173-190.

Maltese, A.V., & Tai, R.H. (2010). Pipeline persistence: Examining the association of educational experiences with earned degrees in STEM among U.S. Students. *Science Education*, *95*, 877-907.

Nauta, M. M., Epperson, D. L., & Kahn, J. H. (1998). A multiple-groups analysis of predictors of higher level career aspirations among women in mathematics, science, and engineering majors. *Journal of Counseling Psychology*, *45*, 483-496.

Navarro, R. L., Flores, L. Y., & Worthington, R. L. (2007). Mexican American middle school students' goal intentions in mathematics and science: A test of social cognitive career theory. *Journal of Counseling Psychology*, *54*, 320-335.

Perez, T., Cromley, J. G., & Kaplan, A. (2014). The role of identity development, values, and costs in college STEM retention. *Journal Of Educational Psychology*, *106*(1), 315-329.

Schaefers, K. G., Epperson, D. L., & Nauta, M. M. (1997). Women's career development: Can theoretically derived variables predict persistence in engineering majors? *Journal of Counseling Psychology*, *44*, 173-183.

Schultheiss, D. E. (2008). Current status and future agenda for theory, research, and practice of childhood career development. *The Career Development Quarterly*, *57*, 7-24.

Shieh, G. (2006). Suppression situations in multiple linear regression. *Educational and Psychological Measurement*, *66*, 435-447.

Shoffner, M. F., Newsome, D., Minton, C. A., & Morris, C. A. (2015). A qualitative exploration of the STEM career-related outcome expectations of young adolescents. *Journal of Career Development*, *42*, 102-116.

Turner, S., & Lapin, R. T., (2002). Career self-efficacy and perceptions of parent support in adolescent career development. *The Career Development Quarterly*, *51*, 44-55.

Usher, E. L., & Pajares, F. (2006). Sources of academic and self-regulatory efficacy beliefs of entering middle school students. *Contemporary Educational Psychology*, *31*, 125-141.

Wang, M., & Degol, J. (2013). Motivational pathways to STEM career choices: Using expectancy–value perspective to understand individual and gender differences in STEM fields. *Developmental Review*, *33*(4), 304-340.

Wang, X. (2013a). Modeling entrance into STEM fields of study among students beginning at community colleges and four-year institutions. *Research in Higher Education*, *54*(6), 664-692.

Wang, X. (2013b). Why students choose STEM majors: Motivation, high school learning, and postsecondary context of support. *American Educational Research Journal*, *50*(5), 1081-1121.

Wood, R., & Bandura, A. (1989). Social cognitive theory of organizational management. *Academy of Management Review*, *14*, 361-384.

Yee, D. K., & Eccles, J. S. (1988). Parent perceptions and attributions for children's math achievement. *Sex Roles*, *19*, 317-333.

Zebrak, K. A., Le, D., Boekeloo, B. O., & Wang, M. Q. (2013). Predictors of intent to pursue a college health science education among high achieving minority 10th graders. *Current Issues in Education*, *16*, 1-11.

J-STEM

Journal of Research in STEM Education

j-stem.net