UC Santa Cruz
UC Santa Cruz Electronic Theses and Dissertations

Title

Scaffolding for Success: When High School Science Teachers Scaffold Their Summative Classroom Assessments; Opportunities, Observations, and Outcomes

Permalink

https://escholarship.org/uc/item/5zg93073

Author

Couling, Joanne

Publication Date

2018

Copyright Information

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

SANTA CRUZ

**SCAFFOLDING FOR SUCCESS: WHEN HIGH SCHOOL SCIENCE
TEACHERS SCAFFOLD THEIR SUMMATIVE CLASSROOM
ASSESSMENTS; OPPORTUNITIES, OBSERVATIONS, AND OUTCOMES**

A dissertation submitted in partial satisfaction
of the requirements for the degree of

DOCTOR OF PHILOSPHY

in

EDUCATION

by

Joanne Couling

June 2018

The Dissertation of Joanne Couling
approved:

_____
Professor Trish Stoddart, Chair

_____
Professor Jerome Shaw, Member

_____
Professor Doris Ash, Member

_____
Professor Marco Bravo, Member

_____
Tyrus Miller, Ph.D.
Vice Provost and Dean of Graduate Studies

TABLE OF CONTENTS

**CHAPTER 3: TEACHER DESIGNED SUMMATIVE ASSESSMENT
SCAFFOLDS IN SCIENCE: STUDENT INTERACTIONS AND**

LIST OF FIGURES

CHAPTER 1

LIST OF TABLES

CHAPTER 1

CHAPTER 2

CHAPTER 3

Abstract

SCAFFOLDING FOR SUCCESS: WHEN HIGH SCHOOL SCIENCE TEACHERS
SCAFFOLD THEIR SUMMATIVE CLASSROOM ASSESSMENTS;
OPPORTUNITIES, OBSERVATIONS, AND OUTCOMES

Joanne Couling

The adoption of the Next Generation Science Standards (NGSS) has brought

with it a need for classroom assessments that measure students' ability to make sense

of, explain, and use science, with many of the performance expectations asking

students to apply scientific principles and evidence to produce an explanation. But

creating coherent written scientific explanations that demonstrate understanding of

scientific principles is not an easy task for students and selecting or creating

classroom assessments to properly evaluate this skill is not an easy task for teachers.

This study serves to incorporate both of these problems by exploring what happens

when scaffolds are included in the assessment process for the students of a group of

science teachers at an urban high school in Northern California who were receiving

coaching in scaffolding assessments. Presented as three distinct but related chapters I

examine how the scientific explanations of students changed when a detailed graphic

organizer/rubric scaffold, called SET4CER, was provided; I explore coaching as a

conveyor of professional development for high school science teachers to incorporate

scaffolding techniques into their classroom assessments; and I document how

students in a kinesiology class interact with teacher designed checklist scaffolds to

support a summative quiz taken under test conditions. The findings illuminate the

difficulties facing teachers in both assessment and scaffold design and indicate that

well-designed scaffolds add value to assessments by supporting students to give their best performance and ensuring that assessments evaluate students' conceptual understanding rather than their memorization or organizational skills. Analysis of the coaching cycles revealed that although the coaching process is not always straightforward, it has potential as a method of professional development delivery that facilitates adoption of the techniques being introduced.

ACKNOWLEDGEMENTS

INTRODUCTION

The origin of this dissertation research was driven from my own experiences as a high school science teacher, where I struggled with the nuances and difficulties that many teachers face when selecting or creating summative classroom assessments that are accessible and relevant to all students, and that adequately show what students know and can do. Classroom assessment is an important part of teacher practice and designing and choosing assessments for use in the classroom requires complex skills on behalf of the teacher (Bell, 2007). We also face the challenge of making sure our assessment systems evolve to stay up with advances in the fields of learning and measurement (Pelligrino, Chudowsky, & Glaser, 2001). With California's adoption of the Next Generation Science Standards (NGSS), which comprise performance expectations of what K-12 students need to do to demonstrate science proficiency, comes a shift in the focus of science standards from curriculum to achievement. Shifting focus in this way reminds teachers that NGSS assessment should be about students demonstrating understanding of science principles and not just recall of disconnected pieces of knowledge. For example, the performance expectations for the eight strands of *Matter and Its Interactions* at high school level state that *students who demonstrate understanding can*: construct and revise an explanation, plan and conduct an investigation, apply scientific principles and evidence to provide an explanation, use mathematical representations to support a claim, refine the design of a chemical system, use the periodic table as a model to make predictions, and develop a model to illustrate a phenomena (NGSS Lead States,

2013). In short, the performance expectations of the new standards indicate that assessments should focus on modeling, inquiry and explanation.

Scientific Explanations

At grades 9 through 12 the NGSS practice of constructing explanations requires students to produce claim-evidence-reasoning (C-E-R) type responses, i.e. make a claim, use supporting evidence to construct or revise an explanation, and apply scientific reasoning or theory to show how or why the evidence supports the claim (NGSS Lead States, 2013). An example of where teachers might assess this practice is when students produce laboratory reports in their science notebooks, which is most often a summative assignment. Laboratory activities within the NGSS are more likely to be focused on a guiding question to investigate or a problem to solve rather than the more traditional version of a step-by-step procedure to follow. Consequently, there is great importance in students being able to produce scientific explanations, a skill that does not necessarily come naturally to them (Ruiz-Primo, Li, Tsai, & Schneider, 2010).

Ruiz-Primo and colleagues (2010) explored the relationship between the quality of students' written scientific explanations in science notebooks and the quality of their learning in a study of 72 middle school science students. They were curious to find out if the students could naturally produce C-E-R explanations, without training, and also if the quality of C-E-R explanations was related in any way to their learning and achievement. Only 18% of the science notebooks were found to have all three aspects of C-E-R in their explanations. 40% produced only claims, that

is, statements without any supporting evidence. A common observation was that often students would describe data but would not use the data as evidence to support their claim. As a result, the authors conclude: "We are convinced that the use of prompts with an adequate level of guidance is what is needed to collect information that can be used for formative and summative assessment purposes" (Ruiz-Primo et al, 2010, p. 605).

Scaffolding Assessments

Throughout my studies as a Doctoral student in education I developed a belief in the benefits of scaffolding, and an interest in exploring the inclusion of scaffolding techniques into the assessment process, in particular the use of scaffolding tools for helping students when generating scientific explanations. Described as providing support to a learner that enables them to proceed further than they would have alone (Wells, 1999), the notion of scaffolding was first introduced to education by Wood, Bruner and Ross (1976) and is firmly embedded within the sociocultural theory of learning. The support, or scaffold, is provided by a *more knowledgeable other*, usually a peer or teacher, who assists the learner to expand beyond what they already know and move into their zone of proximal development (ZPD). The ZPD is a Vygotskian term most often described as the place just beyond a student's current level, a place that is further than they can reach alone but are able to reach if supported. Vygotsky described the ZPD as:

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

During assessment, rather than a peer or teacher, scaffolding comes from prompts or tools that are provided to students with the intention of supporting them to work in their ZPD throughout the assessment process (Poehner & Lantolf, 2005). Classrooms though, are full of students with different ZPDs so how the teacher manages this is critical (Black & Wiliam, 2006). Teachers need a good understanding of their students' abilities to provide high level assessment tasks that are positioned well within students' ZPDs and importantly are supported with well-designed, appropriate scaffolds that ensure the task is reachable (McNeill, Lizotte, Krajcik, & Marx, 2006). Poehner and Lantolf (2005) interpret Vygotsky as describing assessment in this mediational way when he said "we must not measure the child, we must interpret the child and this can only be achieved through interaction and cooperation with the child" (Vygotsky, 1978, as cited in Poehner & Lantolf, 2005, p.240).

For some, sociocultural theory places assessment as part of an evolving activity system, community of practice, or social situation where interaction occurs, with the unit of analysis shifted from the individual learner to a learner who is operating within the learning environment (Moss, 2008). The interaction may be physical or/and symbolic and when it involves some form of scaffolding the assessment allows teachers to move away from outdated traditional methods of assessments and instead encourages assessment activities that bring opportunities to prompt and support learning within the ZPD. Moss is a proponent of activity systems and defines this type of expansive learning as "a learner-operating-with-mediational-means" (Moss, 2008, p.228). While it is not my intention to frame my work within

Cultural Historic Activity Theory (CHAT), when the notion of providing learner attuned mediation is combined with Moss's description that an activity system "reminds us that individuals always have at least partially unique learning trajectories" (Moss, 2008, p.234) the possibilities arising from including scaffolding in the assessment process seems too important to ignore. Scaffolding offers the potential to provide a supportive assessment system that is based on more than just content knowledge recall for students who are learning in a sociocultural classroom. Students bring different experiences and culture to the classroom that shapes how they interact within that social environment, and also with an assessment. In order to understand what students are learning we need to move beyond assessing only what the student can do independently, and instead consider what they can do with help (Lidz, 1995). Hence, scaffolding is an essential component of a sociocultural learning environment, and while there are many forms of scaffolding Wells (1999) has highlighted the importance of the mediational role of artifacts, which is central to this study. In summary, assessments within sociocultural theory are meditational, and shaped by interaction. Providing scaffolding for students as they write scientific explanations has the potential help them to move beyond just stating a claim to producing a C-E-R style explanation that demonstrates their understanding of the scientific principles they have learned.

Getting Assessment Scaffolding into the Classroom

Research into teachers' assessment practices has consistently found that in-service teachers would benefit from additional training and support in effective

assessment practices (Campbell, 2013). With the changing focus of science learning brought about by the NGSS the message about classroom assessment couldn't be clearer: we need new assessments that measure students' ability to make sense of, explain, and use science through crosscutting disciplinary core ideas while engaging in scientific practices. This means that many science teachers need to change their classroom assessment practices, but for that to be successful, targeted professional development and support is critical (Herman & Butler Songer, 2014). Proficiency in assessment is known as assessment literacy, and when teachers are not proficient the consequences for their students can be dire (Stiggins, 2001). Creating assessments that are accessible to all students is an area where teachers already feel inadequately prepared, such that for many, their assessment expertise is obtained and developed on-the-job (Mertler, 2004; Zhang & Bury-Stock, 2003). For in-service teachers, any additional training most often comes in the form of professional development (PD) usually comprised of a day of mandatory attendance, for all faculty, at a program organized by the district or school, and is most likely decontextualized. These programs also tend to assume that knowledge comes from an expert, and that teachers can easily take away and apply what they have learned into their own classrooms (Donnelly & Sadler, 2009; Feiman-Nemser, 2001). Consequently, even if teachers find the PD content useful, without support most do not incorporate the new practice into their own curriculum. Teachers might be more successful in adopting assessment practices if they receive PD that is focused on their specific classroom environment

and they are supported throughout the integration process (Gibbons, Kimmel, & O'Shea, 1997).

While designing my proposed research these issues of teacher adoption of good practice kept coming to the forefront. If I were able to show the effectiveness of assessment scaffolds, how would this practice be adopted and implemented by teachers? I wanted to retain a sociocultural learning theory framework and after considering more traditional methods of professional development I was ultimately led to the relatively new and little researched method of Instructional Coaching. This is a model of coaching where the coach and teacher work collaboratively to develop, implement, and adopt a new practice (Fletcher, 2012; Knight, 2012). Collaboration is an important aspect of the sociocultural theory that develops understanding, and when it occurs with a more knowledgeable other the interaction supports the learner to work in their collective ZPDs, which helps the teacher to go further than they could alone (Vygotsky, 1978; Wells, 1999). More traditional methods of professional development require the teacher to adopt and implement an intervention without any collaboration, support, or assistance and as a result is often unsuccessful.

This study then, has two facets; the main focus is concerned with the importance of scaffolding classroom assessments in science, in particular scaffolding students' written scientific explanations. But it also considers the difficulties of getting professional development successfully into teachers' classrooms. It serves to answer the following questions:

1. Can the use of scaffolds support students to better show what they know and can do when used in the summative assessment process in science?

2. Is coaching an effective delivery method of classroom assessment professional development for science teachers?

To address these questions this dissertation comprises three distinct yet related chapters.

Chapter 1, "Scaffolding Students' Scientific Explanations", investigates the use of a graphic organizer/rubric scaffold to support students from three different high school science classes in their written scientific explanations. The scaffold, adapted from a BSCS Science Learning graphic organizer (www.BSCS.org, 2012), was developed into the Scientific Explanation Tool for Claim, Evidence, Reasoning, or SET4CER, during coaching sessions between myself and the participating teachers. The purpose of this chapter is to examine how the written scientific explanations of three different classes of high school students change when the SET4CER scaffold was provided. In particular, it focuses on two problems that teachers have encountered when asking their students to produce a scientific explanation as part of a summative assessment: producing a coherent, well written, and persuasive explanation in a claim, evidence, and reasoning format; and demonstrating understanding of the scientific principles learned. Analysis is presented as three unifying case studies as the science teachers and their classes provide three different cases of the use of SET4CER as an assessment scaffold. Student achievement is

discussed in terms of examining if the scores they achieve for a written scientific explanation increase when SET4CER is provided.

In Chapter 2, "Exploring the Role of Coaching in Professional Development: A Study of Three Science Teachers and how they Respond to Coaching in Techniques to Scaffold Students' Written Scientific Explanations", I provide an in-depth exploration of the effectiveness of coaching as a conveyor of professional development via descriptive case studies of two of the teachers from chapter one, and one other. The purpose is to glean rich, detailed information about how teachers respond to the coaching process that adds to the literature supporting its potential as a method of administering professional development. Analysis is presented as descriptive chronological narratives of each teachers' initial approach to scaffolding assessments, the nature of their participation in the coaching process, and their evolving adoption of the target professional development.

Chapter 3, "Teacher Designed Summative Assessment Scaffolds in Science: Student Interactions and Achievements", presents an exploration of how the students in a high school kinesiology class respond to and interact with teacher designed assessment scaffolds. Student interaction with the scaffolds and their performance is considered in response to three teacher designed checklists to support summative quiz questions. The purpose, by considering the responses of all students rather than only report accounts where the scaffold is successful, is to produce a richer analysis of the effectiveness of those scaffolds, which not only provides better feedback to the

teacher but also contributes to the lack of valuable information regarding scaffold design to support assessments.

Collectively, these chapters provide valuable information for future studies that examine the design and implementation of assessment scaffolds for summative classroom assessments. It also provides a rich and descriptive account of the intricacies of the coaching process and the nuances surrounding its potential as method of effective professional development adoption.

References

Bell, B. (2007). Classroom assessment of science learning. In S. K. Abell, & N. G. Lederman, (Eds.), *Handbook of Research on Science Education* (pp. 9651006). New York, NY: Routledge.

Black, P., & Wiliam, D. (2006). Developing a theory of formative assessment. In J. Gardner (Ed.), *Assessment and learning* (Chapter 5, pp. 81-100). Thousand Oaks, CA: Sage.

BSCS.org. (2012). Scientific Explanation Tool. Retrieved from: https://bscs.org/tools/scientificexplanationtool.

Campbell. C. (2013). Research on teacher competency in classroom assessment. In J. H. McMillan (Ed.), *Sage Handbook of Research on Classroom Assessment* (pp.71-85). Thousand Oaks, CA: Sage Publications, Inc.

Donnelly, L. A. & Sadler, T. D. (2009). High school science teachers' views of standards and accountability. *Science Education, 93*(6), 1050-1075.

Feiman-Nemser, S. (2001). From preparation to practice: Designing a continuum to strengthen and sustain teaching. *Teachers College Record, 103*(6), 1013-1055.

Fletcher, S. J. (2012). *Coaching: An Overview.* In Fletcher, S. & Mullen, C. (Eds.). The SAGE Handbook of Mentoring and Coaching in Education (pp.24-40). Thousand Oaks, CA: Sage Publications, Inc.

Gibbons, S., Kimmel, H. & O'Shea, M. (1997). Changing teacher behavior through staff development: Implementing the teaching and content standards in science. *School Science and Mathematics: 97*(6) October 1997, 302-309.

Herman, J., & Butler Songer, N. (2014) Assessment for the Next Generation Science Standards. [webinar] In the NSTA web seminar series. Retrieved from: http://learningcenter.nsta.org/products/symposia_seminars/NGSS/webseminar .aspx.

Knight, J. (2012). *Coaching to improve teaching: Using the instructional coaching model.* In C. van Nieuwerburgh (Ed.), Coaching in Education (pp.3-24). London, UK: Karnac.

Lidz, C.S. (1995). Dynamic assessment and the legacy of L.S. Vygotsky. *School Psychology International, 16*(2), 143-153. doi 10.1177/0143034395162005.

McNeill, K.L., Lizotte, D.J., Krajcik, J. & Marx, R.W. (2006). Supporting students' construction of scientific explanations by fading scaffolds in instructional materials. *The Journal of the Learning Sciences,* 15(2), 153-191. doi:10.1207/s15327809jls1502_1.

Mertler, C. A. (2004). Secondary teachers' assessment literacy: Does classroom experience make a difference? *American Secondary Education, 33*(1), 49-64.

Moss, P.A. (2008). Sociocultural implications for assessment I: Classroom assessment. In P.A. Moss, D.C. Pullin, J.P. Gee, E.H. Haertel, & L.J. Young (Eds). *Assessment, equity, and opportunity to learn* (pp.222-258). New York: Cambridge University Press.

NGSS Lead States (2013). *Next Generation Science Standards: For States, By States.* Washington, DC: The National Academies Press.

Pelligrino, J.W., Chudowsky, N., & Glaser, R. (2001). *Knowing what students know: The science and design of educational assessment*. National Academy Press, Washington, D.C.

Poehner, M. E., & Lantolf, J. P. (2005). Dynamic assessment in the language classroom. *Language Teaching Research, 9*(3), 233-265. doi: 10.1191/1362168805Ir166oa.

Ruiz-Primo, M. A., Li, M., Tsai, S. P., & Schneider, J. (2010). Testing one premise of scientific inquiry in science classrooms: Examining students' scientific explanations and student learning. *Journal of Research in Science Teaching, 47*(5), 583–608. doi:10.1002/tea.20356.

Stiggins, R. J. (2001). The unfulfilled promise of classroom assessment. Symposium *Paper presented at Looking Back and Forward: Measurement issues for the 21st Century,* NCME annual conference, New Orleans, LA, April 2000.

Vygotsky, L. S. (1978). *Mind in Society: The development of higher psychological processes.* London: Cambridge University Press.

Wells, G. (1999). *Dialogic Inquiry: Towards a sociocultural practice and theory of education.* New York: Cambridge University Press.

Wood, D., Bruner, J.S., & Ross, G. (1976) The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry, 17*, 89-100.

Zhang, Z., & Burry-Stock, J. A. (2003). Classroom assessment practices and teachers' self-perceived assessment skills. *Applied Measurement in Education, 16*(4), 323-342. doi: 10.1207/S15324818AME1604_4.

CHAPTER 1

SCAFFOLDING STUDENTS' SCIENTIFIC EXPLANATIONS

Introduction

Constructing explanations (science) and designing solutions (engineering) is one of the 8 science and engineering practices of the Next Generation Science Standards (NGSS) as presented in A Framework for K-12 Science Education (NRC, 2012). These practices contribute to the revised learning goals for science, and together with core ideas and crosscutting concepts, are intended to develop students' understanding of the nature of science while mirroring the practices of professional scientists: "The goal for students is to construct logically coherent explanations of phenomena that incorporate their current understanding of science, or a model that represents it, and are consistent with the available evidence" (NRC, 2012, p.52). At grades 9 through 12 the practice of constructing explanations requires students to produce responses where they make a claim, use supporting evidence to construct or revise an explanation, and apply scientific reasoning or theory to show how or why the evidence supports the claim (NGSS Lead States, 2013). This type of explanation is often known as claim-evidence-reasoning or C-E-R and is important because it provides the opportunity for assessments where students can demonstrate scientific sense making because constructing scientific explanations can contribute to developing students' understanding of science concepts and also of how science is done (Quinn, Lee, & Valdes, 2011). When students write scientific explanations, they are using the language of science to write claims supported with evidence and linked

with scientific principles. Participating in scientific literacy in this way helps students to make sense of science in the natural world such that writing a scientific explanation about a phenomenon provides students with more than just *doing science*, instead they are using scientific principles to deepen their understanding (NRC, 2012).

One example of where teachers might assess this practice is when students produce laboratory reports in their science notebooks, which is most often a summative assignment. Laboratory activities within the NGSS are more likely to be focused on a guiding question to investigate or a problem to solve rather than the more traditional version of a step-by-step procedure to follow. Consequently, there is great importance in students being able to produce scientific explanations, a skill that does not necessarily come naturally to them (Ruiz-Primo, Li, Tsai, & Schneider, 2010). Ruiz-Primo and her colleagues (2010) explored if students could naturally produce C-E-R style explanations without training, and also if the quality of their explanations was related in any way to their learning and achievement. Of the 72 middle school science students who participated in the study, only 18% of the students' science notebooks were found to have all three aspects of C-E-R in their explanations, and 40% produced only claim statements without any supporting evidence or scientific principles. A common observation was that students would often describe data but would not use the data as evidence to support their claim. As a result, the authors concluded: "[w]e are convinced that the use of prompts with an adequate level of guidance is what is needed to collect information that can be used for formative and summative assessment purposes" (Ruiz-Primo et al, 2010, p. 605).

Research into the use of prompts, or scaffolds, has shown that when scaffolds are provided and students are given repeated exposure and practice with the scaffold, all students have been able to improve the quality of their C-E-R scientific explanations (Gotwals & Butler Songer, 2013; Kang, Thompson, & Windschitl, 2014). A scaffold can be described as assistance in the form of material, linguistic, cultural, social, or conceptual support (Kang et al, 2014).

Building on prior research, this study serves to examine how the scientific explanations of students from three high school teachers changed when a scaffold was provided. Scaffolding is a key aspect of sociocultural learning theory and in terms of assessment is described as a meditational means to focus on interpreting what a student can do with help rather than what they are currently able to do (Poehner & Lantolf, 2005). The scaffold used by all teachers in this study is the *scientific explanation tool for claim evidence reasoning*, or SET4CER (appendix A), which is a combined graphic organizer and model that was developed based upon the McNeill, Lizotte, Krajcik, and Marx (2006) framework for scientific explanation where reasoning is the link between evidence and claim, shown in the graphic below.
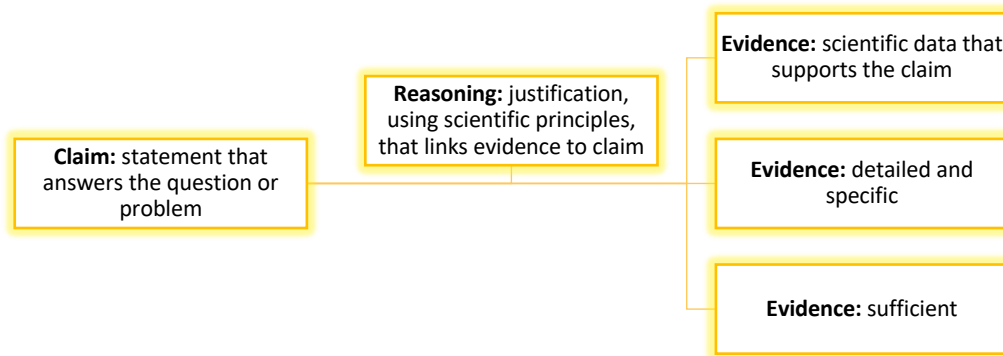
Figure 1. Schematic showing C-E-R relationship in a scientific explanation

Context of Study: Site and Participants

This study was conducted with three high school science teachers of varying experience and backgrounds; Roxanne, Vanessa, and Jessica (pseudonyms) and their students from a culturally diverse urban high school in Northern California, called S.B. High school (pseudonym). The student body comprises 43% Asian/Filipino, 16% European American, 15% Hispanic and 6% African American, with 20% of students reported as multi-racial. There is an established international program at the school, primarily for students from China who live with local host families.

Roxanne has a Bachelor's degree in biology, a Master's degree in education and has been teaching biology for 12 years. She cannot recall any prior assessment training and the last professional development she received was a session on grading approximately 4 years prior to this study. She is currently teaching biology and

17

honors biology. For this study, she elected to work with a class of 10 low-tracked sophomore students whom she is struggling to make progress with and is unhappy about being assigned to teach this year. With regard to writing scientific explanations during laboratory reports she said: "Some of them won't even write a conclusion, it's not even there" (Roxanne, initial interview).

Vanessa has a Bachelor's and Master's degree in forest science. She worked as a research assistant for over ten years, getting her teaching credential later in life. Originally a middle school teacher but for the last 14 years she has been teaching high school biology and integrated science. This is her first year at S.B. High where she currently teaches biology, honors biology, and is preparing to teach AP environmental science. For this study, she elected to work with a regular biology class comprising 5 freshmen and 20 sophomore students. She describes the freshmen in this class as high achieving.

Jessica has a Bachelor's degree in kinesiology with an emphasis in athletic training, and a Master's in education with CLAD. She has been teaching at S.B. High for 20 years, beginning in the athletic department and moving to science for the past 16 years. She is currently the department chair and teaches kinesiology, anatomy and physiology, AP biology, and biotechnology. For this study, she elected to work with her kinesiology students, a mixture of 30 juniors and seniors.

Throughout the duration of this study each of the teachers had been receiving individual coaching from the researcher in the process of scaffolding classroom assessments. This paper focuses on the part of the study that examined changes in

students' written scientific explanations when the SET4CER scaffold tool was used. Observations regarding coaching are recorded in chapter two.

Guided Literature Review

This paper draws from previous research in science education that focuses on (a) scaffolds as a tool to support student learning and assessment, and (b) developing students' written scientific explanations. I discuss prior research that examines how students have been shown to benefit from the use of scaffolds in learning and assessment, concentrating on literature from the two types of scaffolds that apply to this study: graphic organizers and modeling. This is followed with a summary of the findings from the growing literature on students' written scientific explanations, in particular C-E-R style explanations. It should be noted that while there are proponents of C-E-R who incorporate a fourth aspect, rebuttal, into scientific explanations, literature in this area is not discussed as the rebuttal aspect did not feature in the explanations studied for this project.

Scaffolds

Siegel, Wissehr, and Halverson (2008) define a scaffold as: *a temporary support that is later removed.* Other definitions include "…temporary supporting structures provided by people or tools to promote learning of complex problem solving" (McNeill & Krajcik, 2006, p.6). In this study I approach scaffolding as a means to provide access to assessments in terms of language and structure so that students are able to better understand what the assessment is asking of them, and also as a means of guidance to help students with generating a response that really shows

what they know and understand about the science principles that are being assessed. The notion of scaffolding was first introduced by Wood, Bruner and Ross (1976), and the importance of scaffolding students to show what they know and can do has become nested within the sociocultural theory of learning where scaffolding involves providing support from a more knowledgeable other such that a learner is able to proceed further than they could alone. In sociocultural classrooms this has been developed into questioning from the teacher that initiates a discourse to guide and support students to work in their zone of proximal development (Wells, 1999). The zone of proximal development (ZPD) is a Vygotskian term and is best described as the distance between what a student can do with and without help. For Vygotsky, scaffolding provides the help that supports students to just beyond (proximal) their otherwise demonstrated competence. Once the concept is mastered, the support is removed (Vygotsky, 1978). Wells (1999) describes scaffolding students' learning in this way as "operationalizing Vygotsky's concept of working in the zone of proximal development" (p.127). The provision of scaffolding during assessments allows teachers to take their evaluations deeper than more traditional recall assessments by providing prompts that support students to work in their ZPD throughout the assessment process (Moss, 2008; Poehner & Lantolf, 2005).

Hence, scaffolding is an essential component of a sociocultural learning environment, and while there are many forms of scaffolding Wells (1999) highlights the importance of three features of scaffolding that are central to sociocultural theory: discourse to co-construct knowledge, how the learning activity is situated, and the

mediational role of artifacts. This study focuses on scaffolding assessments that are generally more for summative purposes, for example laboratory reports, end of unit tests, and what Kang, Thompson and Windshitl (2014) refer to as *medium cycle assessments*, assessments that are given after 2-3 weeks of instruction within a particular unit. Incorporating Well's principles into summative assessments like these poses challenges for teachers as the norm in such situations would be for students to work independently without discourse of any kind. However, replacing discourse with contextualized, written scaffolds (artifacts) such as graphic organizers allows teachers to incorporate the essential scaffolding principles into their summative assessments without compromising the essence of the assessment purpose.

Siegel, Menon, Sinha, Promyod, Wissehr, and Halverson (2014) explored the effects of using scaffolds in written classroom assessments for middle school students, both native English speakers and English language learners (ELs). Their classroom assessment tasks were specifically designed to follow recommendations by NGSS to engage students in scientific sense making and the use of multifaceted language. The study compared assessments modified for ELs with unmodified assessments and found that both sets of students (ELs and non-ELs) benefitted from the use of scaffolds to help elicit responses and to comprehend and organize their thinking, such as graphic organizers to help with responses. Other studies (for example Abedi, Hofstetter, & Lord, 2004; McNeill & Krajcik, 2006) also report improved access to assessments when scaffolds are used. Methods of scaffolding can

be incorporated into the question or to help students generate their response. While both approaches are equally important, this particular study focuses on the latter.

Kang et al (2014) describe scaffolds as important tools for helping students to engage in more advanced thinking when generating scientific explanations. In a study of the impact of written assessment scaffolds on the quality of students' scientific explanations Kang and colleagues collected assessment tasks and samples of student responses from 33 first year teachers over a two-year period. Over 700 responses from students of varying academic backgrounds were examined using 76 teacher-designed assessment tasks. The aim of their study was to better understand the role of scaffolding in students' written explanations and they found that when used effectively, assessment scaffolds provided better opportunities for students to show what they know. Just as importantly, they found that the quality of the scaffold was more effective than the quantity of scaffolds and conclude that simple, thoughtful scaffolding practices are sufficient to make classroom assessments more accessible to *all* students.

The process of constructing coherent evidence-based responses, or C-E-R responses, can be challenging for students because good scientific explanations demand a degree of complexity that is new to many of them. However, while prior research provides limited direction on types of scaffolds to support students in such complex tasks as writing scientific explanations, the theories of scaffolding for ELs are built upon substantial literature about practices that can be applicable to all students (Kang et al, 2014). One of the practices they recommend is *instructional*

modeling, which is where the teacher provides an example(s) of the structure, format or answer that students can fully or partially emulate. For example, the teacher may make available a sample of completed work that clearly indicates what they are expecting in terms of detail, depth, and overall finished product. This type of scaffold is particularly effective at scaffolding disciplinary language and can also provide motivational support to students (Walqui, 2006). Another recommended scaffold is *bridging,* where teachers provide support for students to make connections between prior knowledge and new learning, including language. One well-known way this is achieved during instruction is with the use of a KWL chart. At the start of a unit the teacher may ask students to complete this type of graphic organizer to help them organize their thoughts. Students write down what they already know (K), and what they would like to find out, or want (W) to know about a particular concept or topic. At the end of instruction students then complete the (L) section with what they have learned (Kang et al, 2014). Bridging techniques can also be incorporated into graphic organizers that scaffold students in the assessment process by helping them organize their thoughts before writing. SET4CER demonstrates a good example of bridging in the first section where it asks a series of where and what questions.

In a review of the literature on graphic organizer research, Wills and Ellis (2005) report that graphic organizers have been found to be helpful in developing students' thinking and learning skills in a variety of content areas. The use of a graphic organizer can make content easier to understand and learn by helping students separate important information from large bodies of text, or by guiding focus to

important, connected pieces of material. Helping students to organize their thinking in this way leads to improvements in analytical, communication, and writing skills (Wills & Ellis, 2005).

Explanations

Constructing and defending explanations has been growing steadily as a practice in science education, and since the onset of NGSS is now considered a central or core practice. Sandoval (2003) and others (see for example Driver, Newton & Osborne, 2000) have shown that constructing explanations in science involves a tight relationship between students' conceptual and epistemic understanding of science; relationships that Sandoval says need to be discussed and extended in the classroom by the teacher.

Sandoval (2003) scaffolded students in their scientific explanation writing with a software program called ExplanationConstructor, a domain-specific technology program that provides guidance to students about what a good scientific explanation entails within a particular problem with regard to articulating a claim and how to use available evidence to support that claim. This was an exploration of how high school students' conceptual understanding of a science domain interacts with their knowledge development, or epistemological ideas, about the nature of science when writing scientific explanations of natural selection. He found that students tended to recognize the need for causal coherence between data and claim but often failed to use data explicitly which Sandoval attributed to students both experiencing difficulty in interpreting the available data and also to a general lack of seeing explicit

evidence as necessary for a good explanation. As a result, he hypothesizes that if students are provided with disciplinary scaffolds grounded within explicitly epistemic structures then they might be helped to use disciplinary concepts in their scientific explanations.

Sandoval's analysis did not focus on quantifying how well the data was used in explanations, i.e. quality and/or quantity of data, but rather to assess the level of students' sensitivity to the need to use data as evidence in support of their claims. He found that for the most part students were successful in articulating a claim and explanation in terms of natural selection theory but were less successful in the skill of interpreting some data relating to the problem. Students had access to charts, graphs, and field notes. Some data were clear and compelling while others were more obscure. Sandoval found that many students made use of the clearer data and cited it to support their claims, however many students also found the more obscure data difficult to understand and although they were able to warrant their claim because they had clearly accessed the data, very few made use of the more obscure data. Sandoval says this is because they did not understand the data and were therefore unable to cite how it supported their claims. He also suggests that another reason that students did not use explicit evidence might be because either they didn't understand it or they did not consider it a necessary component of their explanation. Instead they think that data is supposed to be used as a means to generate a claim and not as specific evidence to support a claim. So, while students in the most part demonstrated an understanding that it is important to support a claim with evidence, many of them

lacked important epistemic strategies such as viewing a lack of available data as a weakness (in support of a claim), for example, if a group could not provide sufficient data in support of their claim this might have been an indicator that an alternative claim may provide a better explanation.

Sandoval's study also highlights a deeper consideration of the relationship between explanations and scaffolds:

> If students can write natural selection explanations, that suggests they understand the theory and how to apply it to answer particular problems. If students can write coherent and well-supported explanations it suggests they understand the epistemic game of explanation (Sandoval, 2003, p7).

So, teachers need an awareness of the importance of scaffold focus – what do you want your scaffold to do? Is the assignment about writing the explanation or about assessing understanding of content? Or both? The level of support that the scaffold provides can be varied according to what the teacher is looking for and what support students might need. Context (content) specific scaffolds can help students with hints about what content to use as supporting evidence, i.e. to connect the science to their explanations, for example a teacher might incorporate a sentence frame that directs students to specific information in the lab data table: *The temperature increased from___˚C to ___˚C.* The use of context supports like this can help students deepen their conceptual understanding and also their knowledge of what constitutes good evidence in a scientific explanation (McNeill & Krajcik, 2006).

Generic scaffolds help with the framework of an explanation and are applicable to all contexts, not just science. Some studies have found that when generic scaffolds are used repeatedly students have been able to transfer the scaffold

across content areas (McNeill & Krajcik, 2006). A written reminder to *use appropriate evidence from your data to support your explanation* is an example of a generic scaffold. McNeill and Krajcik (2006) compared context and generic scaffolds in an attempt to identify the strengths and weaknesses of each type. They found that students who received context specific scaffolds learned more about writing explanations, with improvements in their evidence and reasoning scores, and also more about the science content. The context specific scaffold they used told students exactly what to include and what not to include as evidence. In this particular study, such information assisted with students' learning about the properties of substances and consequently provided deeper understanding of the content while supporting the task of writing an explanation. Students who received this scaffold were able to perform better on a multiple-choice test about content taken later in the unit, while students who received only generic scaffolds did not perform so well on the test.

Sandoval and Millwood (2005) examined the scientific explanations from high school students studying natural selection looking at the quality of the argument rather than the structure of the argument. However, they define quality as something that "includes judgments about the structure of arguments and their conceptual adequacy" p. 24. Which means that as teachers help students to develop their scientific explanations there should be some analytical measures in place that not only assess that students are making the right kinds of arguments but also that those arguments make sense. In this case, their focus was on how students coordinated their claims with supporting evidence; what kind of data did students use as evidence and

how did they explain how the data supports their claims within their explanations? Their study led them to conclude that the data that students use and how they explain that data (rhetorical reference) is a reflection of their implicit beliefs about the nature of science. Constructing an explanation requires a conceptual understanding by the student of relevant theories and how they apply to a specific problem, and also an understanding of what makes a good explanation (Sandoval & Millwood, 2005).

Berland and Reiser (2009) see an alternative way of looking at scientific explanations that is different to C-E-R. They also recognize three components, which they identify as (1) making sense of phenomena, (2) articulating those understandings, and (3) persuading others, and they note that these components are not necessarily equal. Making sense of phenomena requires students to recognize and understand how specific data supports the claim(s) they are making. If students do not understand their data it is possible they may not provide sufficient or accurate data as evidence, as found by Sandoval (2003), see above. Articulation requires students to write their explanations using science language and incorporate the science principles that provide justification, or reasoning, for their claim. While defending one's explanation and persuading others might seem to fit more easily into the "argumentation" field, it is important to remember that when students' written accounts articulate well-presented science principles that link their evidence to their claims, and when they provide sufficient and appropriate evidence in their explanations, they are using persuasion. As Berland and Reiser so accurately put it "(w)hen attention is paid to this goal of persuasion, students move beyond

articulating their understandings, by working to convince their community of the scientific accuracy of their explanations" (Berland & Reiser, 2009, p.30).

Berland and Reiser (2009) examined the successes and challenges experienced by 53 middle school science students as they constructed and defended their written scientific explanations using the C-E-R framework developed by McNeill et al (2006). They found that students either wrote closely following the C-E-R framework or they wove the components together making it difficult to distinguish one from the other. However, all students produced an accurate claim that was supported with evidence so Berland and Reiser attempted to use their own sensemaking-articulating-persuading framework to explain this distinction in how students presented their explanations. They found that when students wove the components together, even if their accounts included good evidence and appropriate scientific principles, that writing in this way makes it difficult for the reader to determine fact from inference, or as they put it, the explanations lacked persuasion, even when students attempt to persuade by using statements such as *our data shows it is true.*

Like Berland and Reiser's study, much of the prior research into written scientific explanations has tended to focus on middle school students, with studies of high school students mainly concentrating on warranting claims and/or specifically how students make sense of and use evidence. Consequently, the literature on how students incorporate scientific principles into their explanations is sparse, particularly at the high school level, even though many have reported that this is the area where

students struggle the most (Bell & Linn, 2000; McNeill & Krajcik, 2008; McNeill et al, 2006).

The nature of writing explanations is clearly complex, and while there are a variety of approaches, most are based upon Toulmin's (1958) model of argumentation, which incorporates claim, data, and warrant. McNeill and her colleagues (2006) developed their claim-evidence-reasoning model from Toulmin's model, modernizing the terminology and making the concepts more accessible to the current education community. This study has adopted the McNeill and colleagues' scientific explanation model and employs a sociocultural perspective in its focus on scaffolding as a supportive assessment method to guide students beyond content knowledge recall.

Research Questions

When writing about science, students are used to reporting their data but often find it difficult to make links between the data they have collected and the scientific principles that they have learned about (Ruiz-Primo et al, 2010). In particular they often struggle to describe how their data supports a conclusion or claim, and in many cases, to even know that they should do this. Consequently, this study focuses on two particular problems that teachers have encountered when asking their students to produce a scientific explanation as part of a summative assessment: producing a coherent, well written, and persuasive explanation that incorporates claim, evidence, and reasoning; and demonstrating their understanding of the scientific principles learned. Grounded in the belief that scaffolding is the best way we can help students

in this process the overarching question guiding this part of my research is *can the use of scaffolds support students in their written scientific explanations so that when a scaffold tool is used they are able to produce explanations that incorporate all three aspects of C-E-R and also demonstrate understanding of scientific principles?*

RQ#1: How does the use of a scaffolding tool affect the structure of students' written scientific explanations?

RQ#2: To what extent does the provision of an assessment scaffold support students in demonstrating their understanding of scientific principles?

Table 1
Research Questions, Data Collected, and Details of Analysis

Research Question	Collected Data	Analysis
1. How does the use of a scaffolding tool affect the structure of students' written scientific explanations?	Student work with and without SET4CER Semi structured teacher interviews	Comparison of components of explanations when scaffolded and un-scaffolded. Explanation detail
2. To what extent does the provision of an assessment scaffold support students in showing their understanding of scientific principles?	Student work with and without SET4CER. Semi structured teacher interviews.	Comparison scores of 2 or 3 for reasoning component when scaffolded and un-scaffolded

Analysis

Data Collection

Semi-structured teacher interviews. At the beginning of the study all three teachers participated in a semi-structured audio recorded interview (Bell, 2010) to

31

find their backgrounds, how their students currently approached written scientific explanations, and to introduce the concept of scaffolding assessment practices. The teachers were asked to select a class to focus on, and to consider an area(s) in an upcoming unit where a summative assessment scientific explanation would best fit.

The teachers were interviewed as often as possible after students had produced a scientific explanation using SET4CER or other scaffolded assessment. They were asked to reflect on how students had performed and to discuss and plan for areas where students might need more or less support in future assignments. Sometimes additional reflections were carried out by email exchange. All audio interviews were transcribed.

Classroom observations. Classrooms were observed several times throughout the study. Vanessa's class was observed as she built on the concept of writing in a C-E-R format. Students had previously been introduced to writing in this way and they had discussed as a class how to best approach this kind of writing before incorporating into several following summative assessments that were collected. Roxanne did not feel comfortable teaching her students about C-E-R writing so I introduced this concept to her students and observed several following lessons. Jessica's class was observed twice within the same unit as students were learning about the key scientific principles of lever systems in the body that she hoped would be incorporated into scientific explanations later in the unit. Field notes were written during the observations.

Artifacts. Teachers shared unit assignments, related assessments, and student work with me, including copies of completed SET4CER handouts. A common un-scaffolded scientific explanation was collected from all students (appendix C) at the beginning of the study. This assignment was created based on the *Instant Laboratory* (Ward's Natural Science Establishment, 2003) that all students undertake in the first week of every school year.

Table 2
Data and Artifacts Collected

Participant	Interview	Observations	Un-scaffolded Explanation	Scaffolded Explanation
Roxanne	Initial 2 reflections Emails	Three	Instant Lab Bird Beak	CER quiz CER test SET4CER
Vanessa	Initial 1 reflection Emails	One	Instant Lab	Bird Beak Natural selection EF comparison SET4CER
Jessica	Initial 2 reflections Emails	Two	Instant Lab	Biomechanics quiz Biomechanics activity Sticker activity SET4CER

Method of Analysis

The three life science teachers and their classes provide three similar detailed and specific cases of analysis that lend towards a case study approach (Lichtman, 2010). However, while they went through similar processes regarding teacher coaching, initial un-scaffolded assignment, and several assessments using SET4CER

and other scaffolds throughout this study, because each teacher and her students are separate and the demographics of each classroom are different my analysis is focused less on case comparison and more as a report of *how and what* in each case so as to unify the potential of the graphic organizer assessment scaffold in three different scenarios. A cross-case comparative analysis to examine variations between the teachers (Patton, 2002) is included in the summary but this is not intended to be the primary focus of analysis.

Scaffolding students' written scientific explanations with SET4CER can provide students with the necessary prompts to access deeper learning so that they are more able to explain and reason about why something has happened. Providing a scaffold like SET4CER also makes it clear to students what the teacher is expecting, the teacher is making their framework explicit which encourages deeper thinking and allows students to make their thinking explicit (McNeill & Krajcik, 2006; 2008). McNeill and Krajcik concentrated on scaffolds that provide both generic (how to write a scientific explanation) and content support as they investigated how to reduce the complexity of writing explanations for students. Building on this, my analysis is focused on changes noted in how students scored for writing a scientific explanation with and without using the scaffold, in particular, improvements from 0 or 1s to 2s and 3s. I explore changes in each section: claim, evidence, and reasoning, and also in the overall presentation of the explanation, and also the content included, i.e. scientific principles.

Un-scaffolded Assignment. At the beginning of data collection each class completed the same un-scaffolded assignment (appendix C). In this assignment students were provided with a sample of data from *The Instant Laboratory*, which is an activity that every student completes at the start of each school year. They were also provided with a short paragraph explaining the scientific principles related to the laboratory and were asked to *write a scientific explanation that states the affect of temperature on time to emergence (of the sponge animal in the Instant Laboratory investigation).* This assignment is referred to as *un-scaffolded* for the purpose of analysis and is used to compare students' explanations with and without the SET4CER scaffold. Essentially this assignment does include a small amount of contextualized scaffold support for students as they are familiar with the data and the assignment asks them to provide a claim supported by evidence and reasoning. The scientific explanation rubric (appendix B) was developed and used to score the explanations. Analysis focuses on if the use of SET4CER moves students from scores of 0 or 1 to scores of 2 or 3 for each category.

Scaffolded Assignments. Each teacher incorporated the use of SET4CER into a content specific assignment after spending class time teaching their students about C-E-R style explanations and practicing with some class examples. Vanessa's biology students had been studying natural selection and evolution and they had recently worked on a project where they each conducted some research on a biome of their choosing and then created a shoebox model of their biomes. Her scaffolded assignment occurred towards the end of the unit and required students to *explain why*

a specific producer and a specific consumer in your project biome is well suited to its

environment.

Roxanne's biology students had been studying adaptations, survival, and food webs. Her scaffolded assignment also occurred towards the end of the unit and students were asked to *explain the impact of the loss of this predator (coyotes) on the squirrel, deer, and blackberry populations only.*

Jessica's kinesiology students were about three quarters of the way through a long biomechanics unit and had just completed learning about the 3 lever systems of the body. Their scaffolded assignment involved an activity where students identified key anatomy involved in each of the lever systems by sticking colored stickers onto their partner and then taking measurements to identify the lever system. After taking photographs of the stickers and recording the measurements students were asked to *write a scientific explanation that identifies the class of lever for the 1st MTP joint and the elbow joint.*

Adapted versions (to each assignment) of the initial scientific explanation rubric were used to score the explanations. Analysis was focused on if the use of SET4CER moves students from scores of 0 or 1 to scores of 2 or 3 for each category.

Vanessa

All 24 of Vanessa's students completed the un-scaffolded assignment but only 18 of these 24 also submitted work in the scaffolded assignment. When reporting on scores achieved and providing specific examples without the scaffold the work of all 24 students is included. When examining changes that occurred between un-

scaffolded and scaffolded explanations only the work of the 18 students who completed both assignments is discussed.

Claim. The scientific explanation rubric describes a claim as *a statement that answers the original question*. Students can score in the range from: *does not make a claim* (score 0) to *makes an accurate and complete claim* (score 3). In the assignment without the SET4CER scaffold 5 of the 24 students (almost 21%) did not score full marks for their claim. One student did not make a claim, two made inaccurate claims and two gave claims that were vague or incomplete. All students who made claims began their explanation with a claim statement and one student separated their explanation into claim, evidence and reasoning sections. When students used the scaffold to help with their explanations the number of students missing full marks in this category was reduced such that only 1 of 18 students (5.5%) who completed both assignments earned less than 3 points. This student wrote a claim that while accurate, was incomplete (score 2 points).

Evidence. The scientific explanation rubric describes evidence as *scientific data or observations that support the claim*. Evidence should be appropriate and there should be sufficient evidence included. Scores range from: *inappropriate or vague evidence* (score 0) to *specific and sufficient* (score 3). There was no significant difference between students who scored full marks for evidence with and without the scaffold, just under 46% without and 44.5% with. However, six students who completed both assignments achieved a gain in their evidence scores when using the scaffold.

Table 3

Gains in Evidence Scores with SET4CER for six of Vanessa's students

Student	Evidence score un-scaffolded	Evidence score with scaffold
202	0	3
206	1	3
212	2	3
213	1	2
215	0	1
217	2	3

There was also an increase in students who achieved a score of 2 or 3 for evidence when they were provided with the SET4CER scaffold. In the un-scaffolded assignment 17 of 24 students (71%) provided sufficient and specific evidence and earned a score of 2 or 3. For the 18 students who completed both assignments 13 (72%) fell into this category. This rose to 16 (89%) when using the scaffold, shown below:

Figure 2. Chart comparing the number of Vanessa's students who scored 2 or 3 for evidence with and without scaffold.

Reasoning. The scientific explanation rubric describes reasoning as *a justification that links the evidence to the claim. It shows why the data counts as evidence to support the claim by using the appropriate and sufficient scientific principles*. Possible scores range from: *does not provide reasoning* (score 0) to *provides appropriate and sufficient scientific principles* (score 3). Often the reasoning component is the area that students find the most difficult, and they commonly repeat their claim here as a form of persuasion instead of including scientific principles (McNeill et al, 2006). In the un-scaffolded assignment 7 of 24 (29%) students made no attempt at reasoning. When students were provided with the SET4CER scaffold only 1 of 18 (5.5%) made no attempt at reasoning. This means that with the scaffold 94% of students attempted to include a reasoning component into their explanations, compared to 71% in the un-scaffolded assignment. Of the 18 students who completed both assignments 12 (67%) attempted reasoning in the un-scaffolded assignment, rising to 17 (94%) who attempted reasoning when the scaffold was provided.

To score 2 or 3 for reasoning means that a student had successfully incorporated some or all of the required scientific principles into their reasoning, producing a persuasive argument. In the un-scaffolded assignment only 9 of Vanessa's students achieved a score of 2 or 3 for reasoning, 7 of these also completed the scaffolded assignment. When the SET4CER scaffold was provided there was an increase to 15 students who scored 2 or 3 for reasoning, see figure 3 below. This shows a significant improvement from 39% in the un-scaffolded assignment to 83% when scaffolded. Although students were able to incorporate some scientific

principles when the scaffold was used, Vanessa had hoped to see better use of the unit

vocabulary:

> Some students thought that as long as they began their final sentence with "therefore" it was all good. I have started including a list of essential vocabulary with my assignments just as a reminder to use some "sciency" language and I think it has helped a bit (Email exchange with Vanessa, 2016)
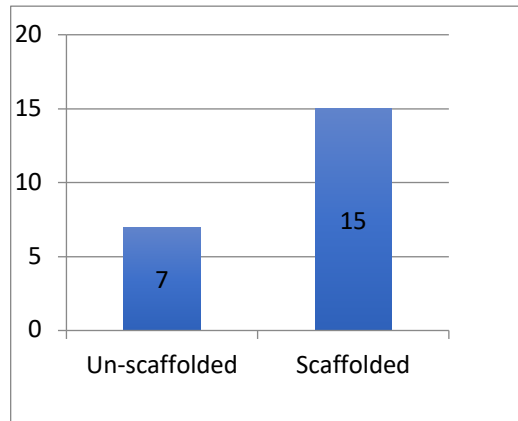
Figure 3. Chart comparing the number of Vanessa's students who scored 2 or 3 for reasoning with and without scaffold.

Observed changes. In overall scores for their scientific explanations 13 of the 18 students (72%) who submitted work in both assignments showed improvement in the evidence and/or reasoning sections of their explanation when the assignment was scaffolded. Many of the improvements involved a two-point increase in reasoning. In this assignment Vanessa was looking for students to discuss adaptations and natural selection. Using SET4CER students were producing reasoning that incorporated some of the scientific principles they had been learning about. Student 220 describes adaptations and discusses why they are important features: *to survive in the arid conditions:*

A Cardon Cactus is well suited to the Atacama Desert because it first has adapted to its freakishly dry conditions and their cores contain porous, which are tiny holes to let the plant ventilate out the heat it absorbs as well as water to keep it hydrated. Second, they have prickly and sharp thorns to protect them against any predators that may arise. Third, they have woody vertical rib like structures that allow the columnar cactus to store water to survive in the arid conditions. Fourth, since they stores [sic] over a ton of water the Cardon Cactus have a framework of hardwood vertical rods, which extremely long to help stiffen ribs and support all the excessive weight. Lastly, they also haves [sic] stomata, which is used to control gas exchange that open only after dark where the cacti absorb carbon dioxide since it happens during the cooler night time, which cause them to be very water efficient. Because of these reasons and many more, the Cardon Cactus are well suited to their environment in the Atacama Desert (scaffolded assignment, student 220).

Figure 4 below compares total scores for scientific explanations by students in this class with and without an assessment scaffold. When the assignment was un-scaffolded students scored, on average, between 5 and 7 out of a possible 9 points. When a scaffold was provided there is a clear shift to higher overall scores, with 13 of 18 students scoring 8 or 9 points. Also important is the change in the number of students who achieved low scores. When a scaffold was provided only 1 student scored under 5 points compared to 4 students without a scaffold.
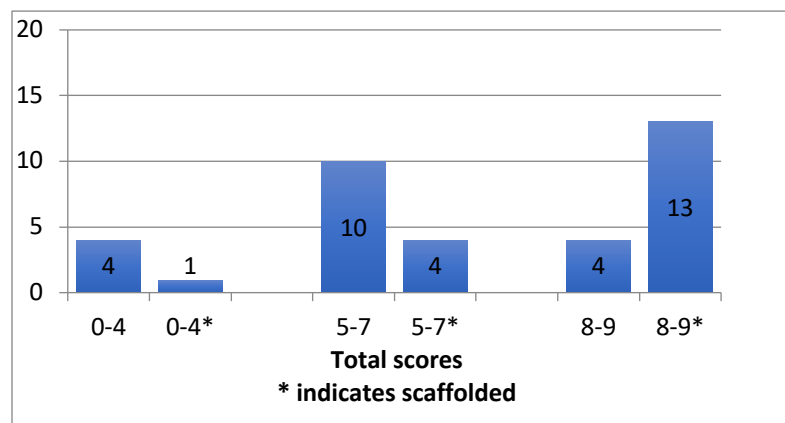
Figure 4. Chart comparing the number of Vanessa's students in each score range (total scores) with and without scaffold.

Roxanne

All 10 of Roxanne's students completed both assignments.

Claim. In the un-scaffolded assignment 8 of the 10 students produced a claim statement, and half of these achieved full marks for their claim, although while some students were able to make clear and accurate claims, these claim sentences were also their total explanation. When the scaffold was provided all 10 students produced a claim statement but only 3 achieved full marks. When using the scaffold 9 of the 10 students specifically started their explanations with an *I claim…* statement.

Evidence. In the un-scaffolded assignment none of the students in this class were able to score full marks for evidence, in fact only 2 of the 10 made any attempt to discuss the data and achieve a score at all, both earned 1 point in this category. This means that 80% of the students in this class did not provide evidence or provided vague or inappropriate evidence in the un-scaffolded assignment. When a scaffold was provided all but three of the students, so 70%, achieved an evidence score, as shown in figure 5 below. This also means that 70% were able to make an improvement in their evidence scores when SET4CER was used. In addition, 4 of the 7 students who achieved an evidence score provided evidence that was specific and sufficient, i.e. scored 2 or 3 in this category.
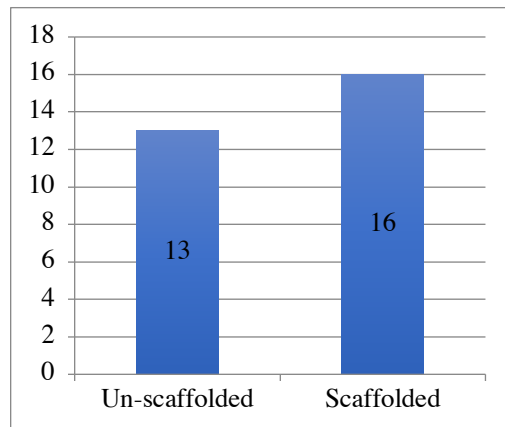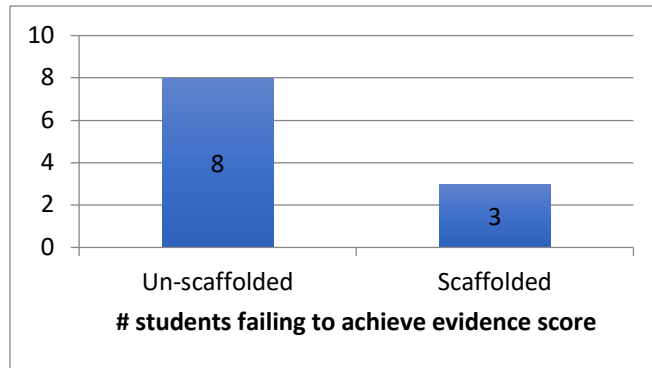
Figure 5. Chart comparing the number of Roxanne's students who scored zero for evidence with and without scaffold.

Reasoning. The scientific explanation rubric awards a score of 1 if students correctly mention how their evidence is linked to their claim but do not incorporate scientific principles, but students would achieve a score of 0 if they do not make this link. Consequently, students could achieve a score of 0 while still having made an attempt at reasoning. In the un-scaffolded assignment only 1 student achieved a reasoning score (score of 1) although 2 students had made an attempt at reasoning. When the scaffold was provided 5 students made an attempt at reasoning but 2 of these achieved a score of 0 because they did not link evidence and claim. Overall, in the scaffolded assignment 7 of the 10 students (70%) showed a small improvement in the reasoning component moving from a score of 0 to a score of 1 or 2.

Because this is a class of lower achieving students I have focused on reasoning attempts rather than full marks, with the teacher being able to use this information formatively to guide further learning and assignments. Without the scaffold only 2 students attempted to incorporate reasoning into their explanations, and only one of them scored any points (score of 1). When the scaffold was used 8 of

the 10 students attempted reasoning, of these students 7 (85%) scored 1 point, and 1 scored 2 points:
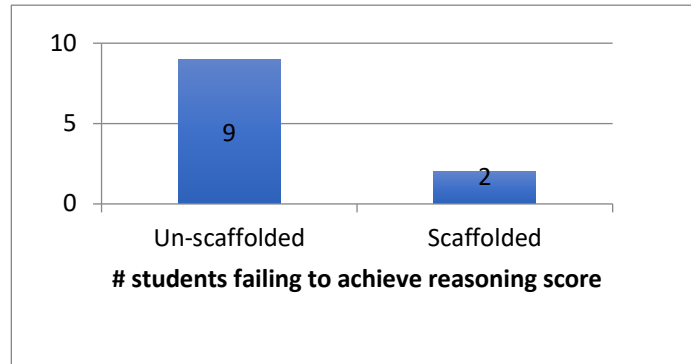
Figure 6. Chart comparing the number of Roxanne's students who scored zero for reasoning with and without scaffold.

Observed changes. Comparing students' explanations with and without the scaffold it is clear that when the scaffold was provided 80% of this class were able to show small improvements, i.e. moving from scores of 0 to 1, and one of the students was indicated as *showing the beginnings of a good CER.*

Overall, students in this class demonstrated great improvement in the structure and detail of their explanations when the assignment was scaffolded. Improvements included incorporating some evidence such as talking about the food web and offering interpretations. For example, student 104 moved from a total score of 0 when un-scaffolded to a score of 7 with the scaffold:

> Un-scaffolded: The data each kid got are very likely to be similar to everyone elses [*sic*]. Just take the first one, the duck has a close similarity to the rest of the other duck type sponges (student 104).

> Scaffolded: My claim is that the populations will decrease. In the temperate forest it seems that there are more predators than preys. With more predators eating or killing the preys it will decrease the population of both preys and predators. No preys [*sic*] means no food for the predators, which will kill both

of them. But if there were less predators and more preys it will increase the population and increase the food source for man and animals (student 104).

Likewise, student 109 was able to score in the evidence and reasoning categories when the scaffold was provided:

Un-scaffolded: In this lab the affect of temperature on time of emergence is that the lower the water temp is the slower the time it will take (student 109).

Scaffolded: The deer, squirrel and blackberry population will thrive once the coyote population dies off. Since the coyotes are gone the 3 populations will have a better chance of reproduction because less animals will kill the 3 pops. Even though more predators for they prey their chance of surviving will be better. The prey will have more offsprings [*sic*] which makes their population stronger and giving their population more years of life (student 109).
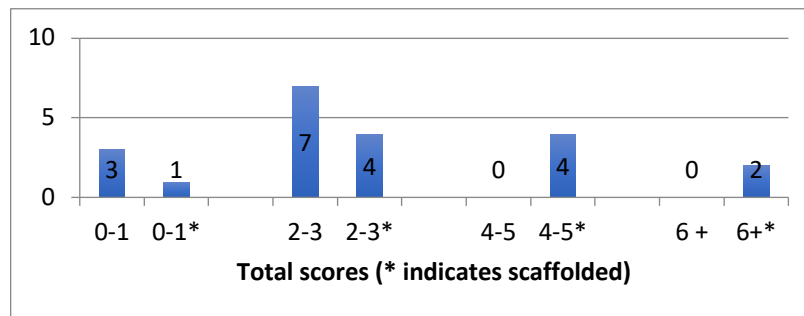
Figure 7. Chart comparing the number of Roxanne's students in each score range (total scores) with and without scaffold.

Figure 7 compares overall scores for the scaffolded and un-scaffolded assignments. In the un-scaffolded assignment most students scored 2 or 3 total points out of a possible 9, and no student scored higher than 3. When the assignment was scaffolded 60% of students earned a total score of above 3, with 20% scoring 6 or above. While these gains are small, what is important is that students appear to be producing explanations with a better structure when provided with a scaffold such that they are attempting more than a claim statement. Providing a scaffold helped almost all of the students with the structure of their explanation in that they were able

to score in all categories. Clearly the teacher has a long way to go with this class in terms of writing scientific explanations but the first step is getting them away from zeros and using SET4CER has shown to be helpful in this.

Jessica

In Jessica's class, 26 students completed the un-scaffolded assignment and 28 students completed the scaffolded assignment. When comparing scores for both assignments only the work of the 26 students who submitted work for both will be discussed or included in calculated percentages.

Claim. In the un-scaffolded assignment 96% of the students produced a claim statement with close to 85% scoring full marks. When scaffolded, 100% of the students were able to produce a claim statement with 96% scoring full marks.

Evidence. For this class, the use of a scaffold did not help students make improvements to their scores for evidence; in fact, scores were lower in the scaffolded assignment. 22 students (84.5%) scored 2 or 3 for evidence in the un-scaffolded assignment, but only 10 of the 26 students (38.5%) who completed both assignments scored 2 or 3 for evidence when a scaffold was used, as shown below.
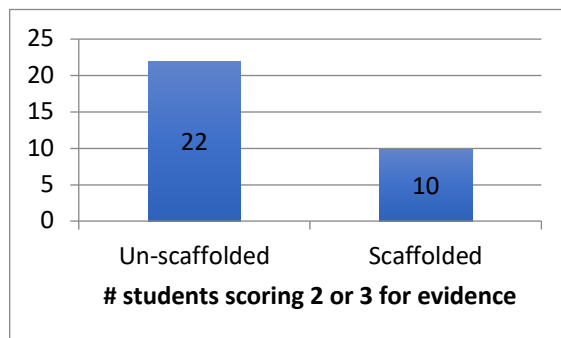
Figure 8. Chart comparing the number of Jessica's students scoring 2 or 3 for evidence with and without scaffold.

Of the students that did not score full marks for evidence in the un-scaffolded assignment it is clear that they were looking at the data and intending to use it as support for their claim, but their evidence is vague and missing specifics, for example:

> "The numbers and averages for each show that the higher temperature had lower numbers and the lower temperature had higher numbers" (student 302, un-scaffolded assignment).

However, in the scaffolded assignment the 16 students who did not score well for evidence completely overlooked the data that they had collected and consequently produced explanations with little or no supporting evidence, even though they have correctly identified the lever systems. For example:

> The class for the 1st MTP Joint would be a second-class lever. This can be classified as a 2nd class lever because when you stand on your toes standing on your 1st MTP joint can form an axis. The elbow joint can be classified as a 3rd lever, the joint can be considered as the axis and force can be applied to when you lean it on something and it is also applied when you move you elbow up and down (student 324, scaffolded assignment)

In this assignment students created a color-coded system representing the common elements of the lever system of the 1st MTP joint when the joint is performing plantar flexion, and the elbow joint, and then placed colored stickers (representing their system) onto one of their four group members, took photographs and recorded measurements in a data table. The assignment then asked them to *write a scientific explanation that identifies the class of lever for the 1st MTP joint and the elbow joint.* While most students were able to give a good description of the definitions of classes of lever and the rules that define each class, many failed to

make the link between their collected data table of measurements as a source of

evidence to prove their claims for lever class, for example:

> The 1st MTP joint is a second class lever system. The elbow joint is a 3rd class lever system. The 1st MTP joint is a second class lever system because the weight is located between the pivot and the force. The elbow joint is a 3rd class lever system because the force is applied between the resistance and the pivot (student 302, scaffolded assignment).

> The 1st MTP Joint is described as a second-class lever because the weight is located between the pivot. The metatarsophalangeal joints form the pivot, which is identified by the red sticker on the foot. The resistance is the weight of the body and the force is applied to the calcaneus bone or heel by the gastrocnemius and soleus muscles through the Achilles tendon. The elbow joint is described as a third-class lever system because the force is applied between the resistance and the axis. The joint is the axis and the resistance is the forearm, wrist, and hand. In addition, the force is the biceps muscle when the elbow is flexed (student 319, scaffolded assignment).

Compare these to the explanation from student 304 who includes measurements from

their data table into the explanation as evidence to support the claims for each joint

classification:

> 1st MTP Joint is a second-class lever system. The Resistance force is located between the axis or pivot which is the 1st MTP joint and effort force which is located at your gastrocnemius. In this class, the effort arm (14 cm) is longer than the resistance arm (10 cm), which is the same representation as the wheel barrel. Also just like a wheel barrel the force is the lower leg, and the axis is the located at the metatarsals, and the weight is in between. Elbow Joint third class lever system, which is where the force is between the weight and the force. This implies that resistance arm is longer than the effort arm, which it is in this case. Since the effort arm is 6.2 cm and the resistance arm is 14 cm. Just like someone picking an item up with a shovel the force applied is between the resistance force and the axis (student 304, scaffold assignment)

Jessica had designed the assignment with the intention for students to use their own

measurements as evidence. She provided the following information to me:

> Evidence collected will be from the measurements that students made in the tables. They should have data that supports the definition below:

By definition a second-class lever has a greater Effort Arm as compared to the Resistance Arm (EA > RA) and both arms are on the same side of the pivot. A third-class lever has an Effort Arm that is less than the Resistance Arm (EA < RA) and both arms are on the same side of the pivot (collected artifact).

Sandoval (2003) noted that when students do not understand data they are unable to cite it as evidence. It seems appropriate to hypothesize that many of these students failed to understand the importance of the purpose of the data they had collected, and it's potential as evidence to support their claims. What is not clear is if students did not understand the purpose of the data they collected, or if they did not understand how to transfer those measurements to support their claims, or if they felt that their written definitions acted as evidence.

Reasoning. In the un-scaffolded assignment 16 students (61.5%) achieved a reasoning score, compared to 100% of students when a scaffold was provided. Of the 16 who scored in the un-scaffolded assignment, 9 scored 2 or above for reasoning, i.e. just over half (56%) of students who achieved a reasoning score provided reasoning that incorporated scientific principles. In the scaffolded assignment, while every student achieved a reasoning score, the percentage scoring 2 or 3 increased slightly to 65% scoring in that range. Overall 15 of the 26 students (58%) who completed both assignments showed improvement in their reasoning scores, and 9 of those improved by 2 or more points when using the scaffold.
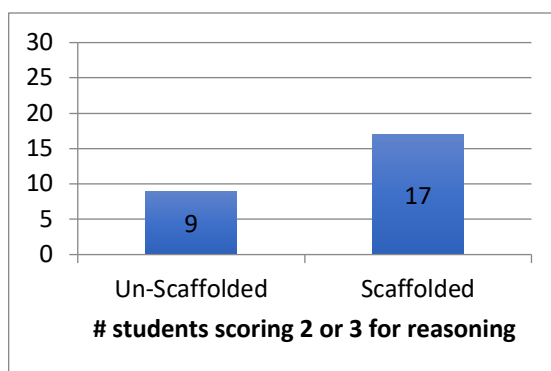
Figure 9. Chart comparing the number of Jessica's students who scored 2 or 3 for reasoning with and without scaffold.

Interestingly 7 students wrote their explanation out of order, so E-C-R, or R-C-E instead of C-E-R. While this did not affect their scores since they were scored on the degree of inclusion of these components, it is interesting that many of the older students chose to write this way in this and other assessments that were collected.

Observed changes.
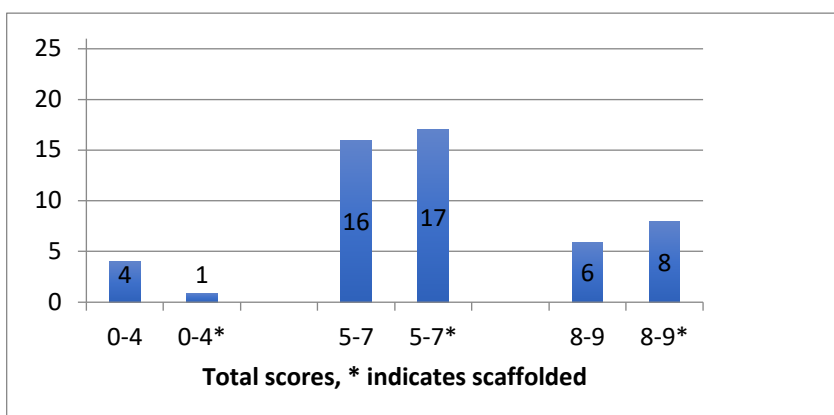
Figure 10. Chart comparing number of Jessica's students in each score range (total scores) with and without scaffold.

Figure 10 above, compares total scores achieved by Jessica's students for their written scientific explanations with and without a scaffold. The chart clearly shows that when scaffolded fewer students achieved a total score of 4 or under, while more

students achieved a total score of 8 or 9 points. The number of students scoring in the middle range, 5-7, remained fairly constant, although given the overall gains, these are likely now a different group of students. Employing a scaffold to support the writing assessment has clearly helped shift the lower scoring students towards higher total marks.

Summary

Vanessa's, Roxanne's, and Jessica's classes provided three different cases of the use of SET4CER as an assessment scaffold, with each class needing something slightly different from the scaffold. Roxanne's lower achieving students had struggled to write more than a claim statement as their whole explanation, and not all of them had managed even that. They needed more help than the others in learning about the components of an explanation and where to find the information they could use. Vanessa's class of sophomores had been attempting to write explanations but needed help putting in more detail so they could move into the more proficient scoring range. In particular they benefited from support with incorporating scientific principles such that their reasoning attempts became more than just attempts. Jessica's upperclassmen were taking an advanced science class and had become very set in old ways of reporting on science. They also benefitted from scaffolding to help with reasoning attempts and the incorporation of scientific principles.

Common Findings

Table 4, below, shows how the students of all three teachers scored in the common un-scaffolded assignment and also in their class specific scaffolded

assignments. With the exception of Jessica's scaffolded evidence scores (discussed in more detail below) the table clearly shows an increase in the number of students achieving higher scores in all C-E-R aspects when an assessment scaffold was provided to students. In total, 54 students completed both assignments, and all but one of those were able to produce a claim statement when the scaffold was provided, with many students electing to begin their explanations with an *I claim* statement. As mentioned earlier, analysis of the scientific explanations of Roxanne's students was not focused on scores of 2 or 3 for evidence and reasoning but rather attempts at evidence and reasoning, however even when this information is included in the table it is clear that small gains were also made in this area for her class. All three classes showed an increase in students being able to incorporate scientific principles into their explanations and achieving scores of 2 and above in this category. When a scaffold was provided 48 of 54 students made an attempt at reasoning. Without the scaffold this number was 31 students.

Table 4
Scores by class and category un-scaffolded and scaffolded

	Claim statement	Score 2 or 3 evidence	Reasoning attempt	Score 2 or 3 reasoning
Roxanne				
Un-scaffolded	8/10 (80%)	0/10 (0%)	2/10 (20%)	0/10 (0%)
Scaffolded	10/10 (100%)	4/10 (40%)	5/10 (50%)	1/10 (10%)
Vanessa				
Un-scaffolded	19/24 (79%)	17/24 (71%)	17/24 (71%)	9/24 (38%)
Scaffolded	17/18 (94%)	16/18 (89%)	17/18 (94%)	15/18 (83%)
Jessica				
Un-scaffolded	25/26 (96%)	22/26 (85%)	12/26 (46%)	9/26 (35%)
Scaffolded	28/28 (100%)	10/28 (36%)	28/28 (100%)	18/28 (64%)

Figure 11, below, shows the mean total scores from all three classes. Although Roxanne's students started at a lower mean score the graph shows that overall scores improved significantly, by 2 points on average, when the assessment scaffold was employed. Jessica's students also made excellent gains in mean total scores for their class, with the average total score increasing from 6.1 to 8 when the scaffold was used. Jessica's totals are suffering from the evidence mishap, where almost half the students scored only 0 or 1 for evidence. However, the delta in total mean score is only 0.2 points with this mishap, which highlights the increases made in the reasoning component.
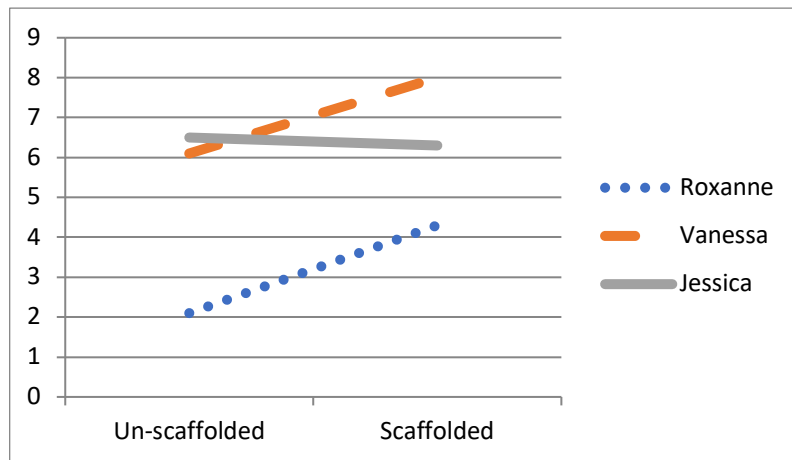
Figure 11. Chart comparing changes in mean total scores from all three teachers with and without scaffold.

When used effectively, assessment scaffolds provide opportunities for students to show what they know about the science they have learned (Kang et al, 2014). The teachers who participated in this study were very happy with how the use of SET4CER helped their students with the structure and content of their explanations

and have continued to use the scaffolding tool with other classes. They were also

happy with the progress their students had made towards writing persuasively;

students were setting out their explanations out clearly, beginning with a claim

statement followed by a discussion of data supported by scientific principles.

Jessica: I feel my students doing better all ready. You could see their
             confidence as they worked through the lesson. Huge difference from
             yesterday's class.

Vanessa: I know that some of them who are good writers can just write a
             paragraph but for those who struggle it definitely helps them to have
             this structure, and that way for the kids who are at the lower end, it has
             definitely helped them.

Roxanne: And then on the reasoning part I think there were some kids that
             showed improvement I think just because it was laid out that way.

Independent Findings

Vanessa. In overall increases Vanessa's students showed the most progress

with their explanations, making notable advances in all categories. Vanessa was

encouraged by the results of using SET4CER with her students and noted that:

> Most students found the organizer helpful, especially the examples
> included…. I think they've improved some since last fall. I hope they don't
> forget over the summer... (exit interview, Vanessa).

She also found that while proving to be a useful tool, her students needed additional

prompting in some areas of the organizer even after they'd spent time in class

learning how to use it together and creating some class examples:

> The "what evidence" section was used variously -- some students were very
> specific, others super vague. The "where will you look for reasoning" was
> very puzzling for some students. I had to reinforce the idea that their notes,
> labs, textbook (!) might be helpful.

Jessica. In the previous school year Jessica had removed the biomechanics unit from her kinesiology syllabus because she said the students had been finding it too hard for the past few years. She acknowledges that it is a difficult unit but felt that her current students were at a higher level and wanted to try to reintroduce the unit while also incorporating more formative assessment and student check-ins and she was keen to incorporate scientific explanations into the summative assessment process. Jessica introduced a variety of scaffolds into her assessment processes in addition to SET4CER, and in assessments that did not incorporate SET4CER Jessica's students often wrote their explanations more as a conclusion and less as a persuasive explanation. This could be because her students, as juniors and seniors, had become proficient in the REEPEPA format of reporting on science and many found the change to a C-E-R only format quite challenging. REEPEPA is a common high school science reporting method for laboratory experiments. The report has three sections: REE – a results, evidence, and explanation discussion; PE – possible errors from experimental design; PA – a conclusion section discussing practical applications.

Regarding the drop in the evidence scores in the scaffolded assignment, Jessica felt that students had completely missed the purpose of the measurements they took and recorded, and instead had freaked out about the amount of text instructions accompanying the assessment and had consequently approached the sticker activity, completion of SET4CER, and writing the explanation as separate assignments: "I'm wondering if they didn't even read what I wrote because it was long" (Jessica,

interview 3). I mentioned above that it was not clear if some students misinterpreted the assignment or indeed if the assignment highlighted a lack of understanding in writing scientific explanations or understanding of the content. Perhaps the students who did not make use of their data were not sufficiently supported into their ZPD by the scaffold. In this case a more dynamic scaffolding process might be beneficial involving teacher-student interactions to consider student responsiveness to the scaffold (Poehner & Lantolf, 2005)

Roxanne. Roxanne was starting from a place where many of her students had only written one sentence in the un-scaffolded assignment, no one had included evidence or attempted reasoning. This was not unusual. When describing how students reported their laboratory investigations Roxanne told me "some of them won't even write a conclusion, it's not even there" (Roxanne, initial interview). Consequently, the goals for her class were different and the aim was to move them towards making attempts in all areas of C-E-R. When using SET4CER all students made improvements in some areas. Most attempted to utilize evidence and were making steps towards thinking about specifics and details, and while they weren't quite there in terms of incorporating scientific principles, in most cases it seemed that students were beginning to realize they needed more than a claim statement such that in the overall structure of their explanations the whole class showed improvements.

Conclusions

The discussion above highlights the success of employing scaffolding during the assessment process in three quite different science classrooms. As mentioned

earlier, research has shown that all students have been able to improve the quality of their C-E-R scientific explanations when scaffolds were provided (Gotwals, & Butler Songer, 2013; Kang et al, 2014), and this study supports those findings. While SET4CER is a generic scaffold we were able to be successful across science disciplines by working individually with each teacher to implement scaffolding techniques into their classrooms. This proved to be a key factor in making the tool work for the needs of each classroom and set of students in both quality of explanation and content understanding. The importance of tailored professional development that individually helps teachers to implement new pedagogy into their classrooms cannot be overlooked and is discussed in chapter 2.

In this study the older students were able to better demonstrate their understanding of the scientific principles they had learned by improving the reasoning component of their explanations when using the SET4CER assessment scaffold; Vanessa's 9th and 10th grade class made similar progress with demonstrating understanding of scientific principles but they also made good gains in their writing structure with almost 100% of the class making attempts in all areas of explanation when the scaffold was used. As the quality of their explanations improved using the scaffold Vanessa was able to use the assessments formatively and began to concentrate on how to get her students to incorporate more science vocabulary into their explanations. For the class of lower achieving students, the scaffold helped them to begin writing scientific explanations that comprised more than just a claim statement. Consequently, scaffolding assessments works, and as Kang and

colleagues' study of scaffolding written assessment tasks concluded "providing effective scaffolding is essential, not optional, when trying to support students in meeting twenty first century standards" (Kang et al, 2014, p.702).

This study has focused on using an assessment scaffold to help students produce coherent, well written, and persuasive explanations with a claim-evidence-reasoning structure, that demonstrate their understanding of scientific principles. When students write scientific explanations in the C-E-R format, they are asked to demonstrate their understanding of science by using scientific principles to explain science phenomena. While writing about science in this way can strengthen their understanding of scientific principles (McNeill & Krajcik, 2008) it is not always an easy task for them, especially if they have been used to writing about science in a more traditional way. Students often find difficulties in supporting their claims with evidence and scientific reasoning, which, according McNeill and Krajcik (2008) is not surprising because traditionally in science the curriculum is taught in a way that requires students to memorize and reproduce facts, so being asked to back up one's statements with scientific principles is not something students are always used to. In addition, traditional curriculum materials often do not support teachers in teaching scientific inquiry processes like this to their students. Constructing an explanation requires a conceptual understanding of the relevant scientific principles, how they apply to a specific problem, and also an understanding of what makes a good explanation (Sandoval & Millwood, 2005), it's a complex process. Scaffolding should provide just enough guidance that students are able to make independent

progress. This study has shown that scaffolding tools that make the scientific explanation process explicit for students can help to ensure that they understand and continue to use these processes. As science education develops and more teachers adapt to NGSS teaching and assessment guidelines there continues to be a need for research and development of such tools to help both teachers and students be successful (McNeill et al, 2008).

References

Abedi, J., Hofstetter, C. H., & Lord, C. (2004). Assessment accommodations for English language learners: Implications for policy-based empirical research. *Review of Educational Research, 74*(1), 1-28.

Bell, J. (2010). *Doing your research project: A guide for first-time researchers in education, health and social science.* Maidenhead, UK: McGraw-Hill Education.

Bell, P., & Linn, M. C. (2000). Scientific arguments as learning artifacts: Designing for learning from the web with KIE. *International Journal of Science Education*, *22*(8), 797–817. doi:10.1080/095006900412284.

Berland, L. K., & Reiser, B. J. (2009). Making sense of argumentation and explanation. *Science Education*, *93*(1), 26-55.

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

Gotwals, A. W., & Butler Songer, N. (2013). Validity evidence for learning progression-based assessment items that fuse core disciplinary ideas and science practices. *Journal of Research in Science Teaching*, *50*(5), 597–626.

Kang, H., Thompson, J., & Windschitl, M. (2014). Creating opportunities for students to show what they know: The role of scaffolding in assessment tasks. *Science Education*, *98*(4), 674-704. doi:10.1002/sce.21123.

Lichtman, M. (2010). *Qualitative research in education: A user's guide*. Los Angeles, CA: Sage Publications Inc.

McNeill, K. L., & Krajcik, J. (2006). Supporting students' construction of scientific explanation through generic versus context- specific written scaffolds. American Educational Research Association Annual Meeting, San Francisco, CA. Retrieved from: http://www.umich.edu/~hiceweb/iqwst/Papers /McNeill_Krajcik_AERA2006.pdf.

McNeill, K.L., Lizotte, D.J., Krajcik, J., & Marx, R.W. (2006). Supporting students' construction of scientific explanations by fading scaffolds in instructional materials. *The Journal of the Learning Sciences, 15*(2), 153-191. doi:10.1207/ s15327809jls1502_1.

McNeill, K.L., & Krajcik, J. (2008). Scientific explanations: Characterizing and evaluating the effects of teachers' instructional practices on student learning. *Journal of Research in Science Teaching, 45*(1), 53-78. doi:10.1002/tea.20201.

McNeill, K.L., & Krajcik, J.S. (2012). *Supporting grade 5-8 students in constructing explanations in science: The claim, evidence, and reasoning framework for talk and writing.* Boston, MA. Pearson.

Moss, P.A. (2008). Sociocultural implications for assessment I: Classroom assessment. In P.A. Moss, D.C. Pullin, J.P. Gee, E.H. Haertel, & L.J. Young (Eds). *Assessment, equity, and opportunity to learn* (pp.222-258). New York Cambridge University Press.

National Research Council: Committee on conceptual framework for the new K-12 science education standards. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas.* Retrieved from http://www.nextgenscience.org/next-generation-science-standards.

NGSS Lead States. (2013). *Next Generation Science Standards: For States, By States.* Washington, DC: The National Academies Press.

Patton, M. Q. (2002). *Qualitative research & evaluation methods.* Thousand Oaks, CA: Sage Publications, Inc.

Poehner, M. E., & Lantolf, J. P. (2005). Dynamic assessment in the language classroom. *Language Teaching Research, 9*(3), 233-265. doi:10.1191/1362168805lr166oa.

Quinn, H., Lee, O., & Valdes, G. (2011). Language demands and opportunities in relation to Next Generation Science Standards for English language learners : What teachers need to know. Downloaded from: http://mes.sccoe.org/resources/ALI%202012/11_KenjiUL%20Stanford%20Final%205-9-12%20w%20cover.pdf#page=44.

Ruiz-Primo, M. A., Li, M., Tsai, S. P., & Schneider, J. (2010). Testing one premise of scientific inquiry in science classrooms: Examining students' scientific explanations and student learning. *Journal of Research in Science Teaching, 47*(5), 583–608. doi:10.1002/tea.20356.

Sandoval, W. J. (2003). Conceptual and epistemic aspects of students' scientific explanations, *The Journal of the Learning Science, 12*(1), 5-51. Downloaded from http://www.jstor.org/stable/1466633. Accessed: 17-08-2016.

Sandoval, W. A., & Millwood, K. A. (2005) The quality of students' use of evidence in written scientific explanations, *Cognition and Instruction*, *23*(1), 23-55, doi:10.1207/s1532690xci2301_2.

Siegel, M. A., Wissehr, C., & Halverson, K. (2008). Sounds like success: A framework for equitable assessment: How to revise written assessments for English language learners, *The Science Teacher*, *75*(3), 43–46. Retrieved from http://people.uncw.edu/kubaskod/NC_Teach/Class_5_Assessment/Sounds_Like_Success_Equitable_Assessment.pdf.

Siegel, M. A., Menon, D., Sinha, S., Promyod, N., Wissehr, C., & Halverson, K. L. (2014). Equitable written assessments for English language learners: How scaffolding helps. *Journal of Science Teacher Education (25)*, 681-708.

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

Vygotsky, L. S. (1978). *Mind in Society: The development of higher psychological processes*. London: Cambridge University Press.

Walqui, A. (2006). Scaffolding instruction for English language learners: A conceptual framework. *International Journal of Bilingual Education and Bilingualism*, *9*(2), 159–180. doi:10.1080/13670050608668639.

Ward's Natural Science Establishment (2003). The Instant Laboratory: Introduction to experimental design and data presentation lab activity.

Wells, G. (1999). *Dialogic Inquiry: Towards a sociocultural practice and theory of education.* New York: Cambridge University Press.

Wills, S., & Ellis, E.S. (2005). The theoretical and empirical basis for graphic organizer instruction. Downloaded from: http://www.calhoun.k12.al.us. *Retrieved December*, *13*, 2016.

Wood, D., Bruner, J.S., & Ross, G. (1976) The role of tutoring in problem solving, *Journal of Child Psychology and Psychiatry*, *17*, 89-100.

Appendix A

Developing a Scientific Explanation using CER

What is the question that you want to answer?
What evidence do you need to answer this question?
Where will you look to find the scientific principles to provide reasoning?

<center>**Support for your explanation**</center>

Evidence	Scientific Principles

<center>**Scientific explanation = Claim + Evidence + Scientific Reasoning (CER)**</center> **My claim is __(answer to the question)__ because ___(Evidence followed by scientific reasoning)__**

Question:	Examples: - Did a chemical reaction occur? - How would the hawk population change if there were no seeds available? - Can animals survive in the desert? - Do levers make work easier?
Claim:	Examples:

A statement that answers the original question.	- A chemical reaction occurred. - The hawk population would decrease if there were no seeds. - Only animals that are adapted to desert conditions are able to survive in the desert. - Levers can sometimes make work easier.
Evidence: *Specific data that supports the claim, such as numbers from a data table or observations. Data needs to be appropriate and you need enough data to support your claim (e.g. high and low numbers from the data table)*	Examples: - The density of butanol is 0.81g/cm^3, whereas the density of layer A is 0.87 g/cm^3 - The hawk eats squirrels, rabbits and sparrows. Squirrels only eat seeds. Sparrows eat seeds and grasshoppers. - Camels have humps to store fat, they don't sweat much, they have really big feet and they have fur. - When the effort distance was 60 cm the effort force was 2.5N, but when the effort distance was 20 cm the effort force was 8N.
Reasoning: *Justification that links the claim and the evidence. Include appropriate and sufficient scientific principles.* *This is where you include the science you have learned. Look in your class notes, resources provided by the teacher, introductory or background text for a lab, or other available texts.*	Examples: - Density is a property of a substance along with melting point and color. In a chemical reaction, new substances are formed with different properties to the original substances. The differences in densities between butanol and layer A indicate that a chemical reaction occurred and that layer A is a new substance. - Organisms in a food web are affected by other organisms in the same food web even if they are not directly linked to them. Without seeds the squirrels would have no food so they would die out and the sparrows would have less food so their numbers would decrease. This means there will be fewer hawks because they will have less food to eat. - Camels have adaptations that help them survive in desert conditions. By storing fat in their humps, they are able to go longer without food and because they don't sweat much they lose less water from their bodies and can go longer without water. Their big feet make it easier to walk on the sand and their fur keeps them warm at night as it can get extremely cold in the desert at night. - Doing work is the ability to move an object. If it takes less force the work feels easier. A lever can help the work feel easier depending on the position of the fulcrum, effort, and load.

Appendix B
Pre-assessment Activity – scientific explanation rubric

Component	Levels			
	0	**1**	**2**	**3**
Claim A statement or conclusion that answers the original question.	Does not make a claim.	Makes an inaccurate claim. ----- e.g. rate of emergence depends on capsule color.	Makes an accurate but vague or incomplete claim. ----- e.g. rate of emergence depends on water temperature.	Makes an accurate and complete claim. ----- e.g. emergence rate increases as water temperature increases.
Evidence Scientific data that supports the claim. The data needs to be appropriate and sufficient to support the claim.	Does not provide evidence, or only provides inappropriate or vague evidence. ----- e.g. the data shows me it is true.	Provides insufficient evidence to support the claim, does not include specific data. ----- e.g. when the water temperature is higher the animal emerged faster	Provides 1 of the following 2 pieces of evidence: - specific data (e.g. numbers) from the investigation when the rate was faster with higher temperatures. - Specific data (e.g. numbers) from the investigation when the rate was slower with lower temperatures.	Provides both of the following pieces of evidence: - specific data (e.g. numbers) from the investigation when the rate was faster with higher temperatures. - Specific data (e.g. numbers) from the investigation when the rate was slower with lower temperatures.
Reasoning A justification that links the evidence to the claim. It shows why the data counts as evidence to support the claim by using the appropriate and sufficient scientific principles.	Does not provide reasoning, or only provides reasoning that does not link evidence to the claim. ----- e.g. all of the capsules dissolved	Repeats the evidence and links it to the claim but does not include scientific principles. ----- e.g. the increased temperature melts the capsule faster	Provides 1 of the following 2 reasoning components: - an increase in T produces increase in KE. - increased KE means molecules collide more often and with more energy so take less time to emergence	Provides both of the following reasoning components: - an increase in T produces increase in KE. - increased KE means molecules collide more often and with more energy so take less time to emergence

Adapted from McNeill & Krajcik, 2012

Un-scaffolded Assignment

A group of students have been learning about the effects of temperature on dissolving. They read the paragraph below and then repeated the *Instant Laboratory* investigation.

Heating up a solvent gives the molecules more kinetic energy. The more rapid motion means the solvent molecules collide with the solute molecules more often and the collisions occur with more force. Both factors increase the rate at which the solute dissolves.

In the investigation they collected the following data:

Sponge Type	Capsule Color	Water temperature (˚C)	Capsule width (mm)	Capsule length (mm)	Time to emergence (s)
Duck	Purple	40	5	20	99
Duck	Purple	60	6	21	66
Duck	Purple	60	5	21	56
Goat	Yellow	40	6	20	136
Goat	Yellow	50	6	20	108
Goat	Yellow	40	6	21	288
Sheep	Pink	40	5	20	196
Sheep	Pink	60	6	20	61
Sheep	Pinky red	50	5	21	72
Horse	Yellow	50	6	21	162
Horse	Yellow	50	6	20	98
Horse	Yellow	60	6	21	58
Pig	Purple	40	5	20	188
Pig	Blue	50	5	20	82
Pig	Purple	50	5	20	67

In the space below write a **scientific explanation** that states the affect of temperature on time to emergence. Remember to support your claim with evidence and reasoning.

CHAPTER 2

EXPLORING THE ROLE OF COACHING IN PROFESSIONAL
DEVELOPMENT: A STUDY OF THREE SCIENCE TEACHERS AND HOW
THEY RESPOND TO COACHING IN TECHNIQUES TO SCAFFOLD
STUDENTS' WRITTEN SCIENTIFIC EXPLANATIONS.

Abstract

This paper explores coaching as a conveyor of professional development for high school science teachers to incorporate scaffolding techniques as they assess their students. Through a series of coaching cycles, I (coach and author) aimed to assist three teachers to scaffold their students when writing scientific explanations for summative assessment. Qualitative analysis of these coaching cycles is used to explore the effectiveness of coaching in this role via descriptive case studies of Vanessa, Roxanne, and Henry, three high school science teachers. These case studies reveal a spectrum of take-up and adoption of scaffolding techniques which are discussed in terms of teacher responsiveness to coaching.

Introduction

Effective professional development (PD) has been shown to be an important factor in increasing student performance (Cornett & Knight, 2009). However, teachers often view PD sessions as boring and of little or no value (Wilson & Berne, 1999), and consequently take little to nothing away from the sessions. Feiman-Nemser (2001) and others describe professional development as mandatory attendance at a program organized by the district or school, or as part of a graduate school program. Both approaches assume that the professional development knowledge comes from experts and that teachers can take away what they've learned and apply it in their own classrooms. Even though many teachers are able to self-identify areas in which they need more training, school or district mandated PD programs often do not take this into consideration. Instead the programs are frequently skewed towards out dated models of teaching and learning and are decontextualized, and so do not help teachers bring new knowledge into their practice (Feiman-Nemser, 2001). From a review of the literature Wilson and Berne (1999) distinguish effective PD as something that requires teacher collaboration so that there is collective participation in implementation. They go on to say that to be successful PD should be focused in areas of teachers' curriculum and instruction with specific needs, should be embedded in teacher work, with adequate time provided, and be inclusive and accessible. For assessment related professional development with in-service teachers, Zhang and Burry-Stock (2003) recommended tailoring the PD to suit

the needs of the individual teacher, their content area and the grade level of their students.

Coaching as a Transport for Successful Professional Development

Coaching is a fairly recent addition to education, following on the heels of mentoring, and many universities are still to incorporate coaching as part of teacher preparation programs (Fletcher, 2012). Educational coaching has been little researched, with much of what we currently know about the successes of coaching coming from the world of business, where it is fairly common. This means that much of the available coaching material has been written for education practitioners by business experts and as a consequence, is often based upon business literature (Fletcher, 2012). To ensure an education perspective Fletcher recommends that in addition to this available business-oriented material, which has value, those interested in educational coaching also need access to education research-based information, which unfortunately is sparse.

Instructional coaching. One of the more recent education-based models is instructional coaching, which has been developed by Jim Knight at the University of Kansas. His model contains the components that prior research has highlighted as necessary for teacher uptake of PD. For example, for a teacher who might be concerned about issues of accessibility in her assessments administered to her English language learner students (ELs), a traditional PD session might describe how using assessment scaffolds, such as sentence frames or incorporating images has been shown to help ELs access assessments, but it does not provide any support or help for

the teacher to implement this new learning into her own classroom, with her own students beyond providing the teacher with the idea. The instructional coaching model is a partnership model built upon seven principles designed to do exactly that - improve the likelihood of teachers adopting new practices into their classrooms following PD (Cornett & Knight, 2009; Knight, 2007).

The Coaching Model: A Sociocultural Perspective

Instructional Coaching is described as a partnership model because the coach and teacher work collaboratively to develop, implement, and adopt a new practice (Fletcher, 2012; Knight, 2009). Collaboration is the key to successful coaching. While the coach may offer suggestions based upon prior experience it is not the intention that the teacher just receives tools from the coach without collaborating on design and implementation plans. Collaboration is a social interaction process that develops understanding, and when it occurs with a more knowledgeable other, in this case a peer, the interaction supports the learner to work in their collective zone of proximal development (ZPD), helping the teacher to go further than they could alone (Vygotsky, 1978; Wells, 1999). Thus, it is the very essence of sociocultural learning, as teacher and coach establish one-on-one dialogues that identify a problem and then the more knowledgeable other, in this case the coach, helps to mediate to a solution that the teacher then implements. The importance of such joint dialogues is also a central tenant to the sociocultural theory of learning (Tharp & Gallimore, 1988). More traditional methods of professional development require the teacher to adopt

and implement an intervention without any collaboration, support, or assistance from peers.

Instructional Coaching Adaptation for This Study

The Knight model has seven principles, or steps, that begin with an interview to establish a coaching partnership rapport. After this initial meeting, the coach and teacher collaboratively plan an intervention and co-design an observation form. The coach then models the new practice in the teacher's classroom while the teacher observes using the co-designed observation form. The teacher and coach meet in a teacher directed conference to discuss the teacher observations and then the teacher teaches using the new practice while the coach observes, using the same observation form. Following this, the teacher and coach meet again to collaboratively explore the data they have collected. The coach provides ongoing support as needed by the teacher until they are ready to continue implementing the new practice without support (Knight, 2007; Knight & Cornett, 2008). For this study, Knight's instructional coaching model was adapted slightly as the intended new practice is centered in assessment rather than instruction. Collaboration remains central to the model but instead of collaborating on the design of an observation form, teacher and coach collaborate on the design of scaffolding tools for students to use during assessment. The adapted model is shown schematically below. The six stages comprise one coaching cycle. When the coach and teacher meet in step 5, a reflection stage, they have the option to return to stage 2 for a second (or third) cycle until they

both feel that the aim(s) have been met and the area the teacher raised for concern has been addressed. At this point the partnership moves into stage 6.

(1) Interview	(2) Research & Design	(3) Collaborate
One on one coach/teacher interview. Teacher identifies an area of concern. Coaching partnership rapport estabilshed.	Coach researches evidence-based solutions for planned intervention.	Teacher/coach collaboration in co-design of assessment scaffolds and scoring rubrics.

(4) Implement	(5) Reflect	(6) Fading
Teacher implements intervention. Scores using rubric.	Teacher/coach meet to analyze results. Return to stage 2-4 or 6 as necessary	Coach continues to support teacher until teacher is ready to move on alone.

Figure 1. Coaching cycle phases.

Case Study Methodology

I used a case study methodology to examine the potential of coaching as a tool for successful professional development implementation. Case study data provides rich and detailed information facilitating an holistic and context specific analysis (Patton, 2002). The richness of the information gained is more important that the ability to generalize (Lichtman, 2010). I collected and transcribed audio recordings of all meetings; field notes, observation notes, and email conversations. All data was compiled into case matrices to facilitate analysis of the multiple data sources (Miles & Huberman, 1994). Case matrices provide visual displays of data, like a map, to simplify the process of comparing large amounts of sequential text, allowing multiple variables to be considered at once. Visualizing key variables not only highlights

possible connections but also provides an indication of how the variables might influence each other (Miles & Huberman, 1994).

Drawing on these case matrices I created chronological narratives framed within coaching cycles of each teacher. These narratives describe the teachers' initial approach to scaffolding assessments, the nature of their participation in the coaching process, and their evolving adoption of the target professional development. This is followed by my evaluation of the success of anticipated professional growth.

Table 1
Data collected, purpose, and focus of analysis

Data source	Purpose	Analysis focus from case matrix
Initial interview	Understand teacher background & experience with CER Identify area for coaching Establish coaching partnership	Highlight connections Measure anticipated professional growth
Transcribed audio from coaching meetings	Monitor collaboration Student achievement Measure intervention Highlight relationship	Reflective dialogues Responsiveness Student achievement Effective & ineffective strategies
Email communication	Exchange resources Planning Continuing collaboration	Responsiveness Reflective dialogues
Observation	Informal Gain understanding to support coaching/collaboration	Strategies Responsiveness to coaching
Post coaching reflection	Understanding of degree of continued use of PD	Reflective dialogues Effective & ineffective strategies Responsiveness Measure anticipated professional growth

The case studies reveal the large amount of dedication, effort, and tensions that need to be considered when introducing new pedagogies with teachers. The findings are discussed in terms of coaching to support professional development implementation and also acknowledges the tension of teacher beliefs. The observations provide valuable descriptive accounts that add to existing literature in both of these areas.

Case Study Selection

This paper comprises part of a larger research effort (doctoral dissertation research) with eight teachers of the science department at S.B. High school (pseudonym), a culturally diverse urban high school in Northern California. The overall study serves to examine students' written scientific explanations during assessments when their teacher has provided scaffolding. The teachers each received coaching in scaffolding techniques and claim-evidence-reasoning (C-E-R) style explanations (see chapter one). The coaching process involved a series of meetings and observations with each of the participating teachers. As these sessions increased I began to notice that their responsiveness to coaching lay on a spectrum. The differences were not overt, for example teachers were neither approving or disapproving of being coached, or of the progressions that we took. Rather, as the number of meetings and observations grew, and as I began to analyze the changes in student responses when assessments were scaffolded and I was able to evaluate the effectiveness of each coaching cycle, I became more aware of the differences in teacher responsiveness and cooperation. For this reason, I began to investigate where

each teacher fell on a spectrum of responsiveness to coaching, and also of successful

professional growth.

Because of these subtle differences, I found the need to create a *teacher take-up rubric* (TTUR) (appendix A) to reflect more on their responsiveness. Each teacher

was scored on their response to and engagement with coaching, scaffolding, and use

of scientific explanations. The scores translated to a spectrum.

None	Minimal	Moderate	Full

Figure 2. Spectrum of teacher response to coaching.

Based on their place on the spectrum I selected three teachers to look at more

closely. Vanessa scored highly on the TTUR, being very responsive to coaching, her

take-up of scaffolding techniques, and implementation and continued use of C-E-R

scientific explanations. I have selected Vanessa as someone high on the spectrum in

all areas. Henry came across as very enthusiastic in adoption of scaffolding and

responsiveness to coaching, but I came to understand that all was not quite as it

appeared on the surface. I selected him as someone who sits mid-range on the

spectrum. Finally, Roxanne, a teacher who began by strongly requesting help with her

low track students but did not respond as expected, represents low spectrum

positioning.

Table 2

Case study teachers – degree of responsiveness and outcome

Teacher	Coaching	Scaffolding	Scientific Explanation	Professional Growth
Vanessa	High	Moderate	Full	High
Henry	High	Minimal/Moderate	Moderate	Moderate
Roxanne	Minimal	None/minimal	Minimal	Low

Coaching Vanessa

Vanessa, a veteran teacher who joined the S.B. High school science department at the beginning of the school year in which this study took place, has both a Bachelors' and Masters' degree in Forest Science and worked for 10 years as a research assistant before she became a teacher. I met with Vanessa in a coaching capacity for the first time early in the second semester, where she told me she was teaching biology, and honors biology. She elected for us to work together with one of her regular biology classes comprising 25 mixed ninth and tenth grade students, describing the 5 ninth graders as high achieving and the 20 tenth graders as "less so". All but two of the students are Filipino or Hispanic but none are classified as English language learners. In this initial meeting Vanessa told me "students just want to pass, they do the work but they have no motivation." Vanessa's biology curriculum is split into four units, one for each quarter, and the class was part way through an eight-week long Evolution and Ecology unit. They were studying classification and students had been learning how to produce cladograms.

Vanessa's Pre-Coaching Approach to Scaffolding Assessments

Vanessa immediately comes across as sensitive to her students' needs and throughout our time working together she talked often about adjusting her existing

assignments specifically for her current group of students. For example, when describing the cladogram assignment:

> I've changed [the assignment] a little bit because I realized last semester that they're, um they just have different skill sets to kids I've had previously…. There's a little more scaffolding going on, there's the hand out with the specific traits… so the problem always comes in in the paragraphs, and so I've rewritten the directions, fingers crossed...
> (Vanessa, initial interview)

It's clear Vanessa is a scaffolding believer. The scaffolded hand out she mentions is a graphic organizer (GO) with a data table for students to complete using provided information, followed by space where they write an analysis of the data. She explains that by using the graphic organizer, students are guided through the process without missing any important information. Students are to write an explanation in the analysis section, but Vanessa expresses concern about ongoing problems in this area. Even though she has rewritten the assignment directions she is cautious about what to expect because her students struggle to produce good explanations. With her current group Vanessa feels that 10% of the class are able produce good scientific explanations, 10% "struggle immensely", and the remaining 80% are "okay but lack detail and don't seem to be getting any better." Together we set a professional growth target to improve the scaffolding in her GOs, to lead to improved written explanations in a C-E-R format for all students. I also hope that Vanessa can move to a place where she incorporates scaffolding techniques into her assessments as a norm.

Vanessa's Coaching Cycle

As we talk more about her students' scientific explanation writing it becomes evident that they are making claims and collecting data but generally do not use

scientific principles as reasoning to link their claim and evidence. This is not uncommon, students often find it difficult to articulate and defend their claims and to recognize and use appropriate and sufficient supporting evidence (McNeill & Krajcik, 2006; 2008a; 2008b; Sandoval & Millwood, 2005).

I began by showing Vanessa a variety of graphic organizers specifically for writing scientific explanations and she selected a design suitable for her students. I also gave her a list of transition words designed to help students with writing in science, and a generic C-E-R scientific explanation scoring rubric. Vanessa planned some time to teach her students about C-E-R during her next lesson, and for the lesson following this I joined the class to observe Vanessa and her students as they began to incorporate C-E-R into their curriculum. During the observation, I notice that Vanessa's students each had a copy of the C-E-R rubric and also the scientific explanation graphic organizer. One student asked "are we doing C-E-R again?" Vanessa explained that she would like the students to redo their latest assignment, together as a class, using the graphic organizer and rubric to build on what they'd learnt about writing scientific explanations. Towards the end of the lesson she tells the students that they have two upcoming labs on natural selection which will both have a focus on improved written scientific explanations.

Implementation. The first lab is the Bird Beak lab, Vanessa shared both her old and newly scaffolded assignments with me. She recognized that her students have been doing well with C-E-R in class so she felt justified to incorporate minimal scaffolding into the assignment with just a simple reminder to use C-E-R format,

supported by showing the points available in an accompanying grading rubric, and

providing access to the generic scientific explanation graphic organizer and the

linking words document they had been using in class. Vanessa's *reminder* scaffolds:

> *Assignment:*
> [write a] conclusion based on the data. Your conclusion should include an
> explanation of whether or not evolution occurred in this lab. Use C-E-R
> format in your explanation.
>
> *Grading rubric:*
> (7 pts) Conclusion in C-E-R format: explain whether or not evolution
> occurred
> Claim – 1 pt
> Evidence – 3 pts
> Reasoning – 3 pts

Reflection. Vanessa and I met again in the middle of February after she had

collected and graded the bird beak lab. I am impressed by her skills in analyzing their

explanations and identifying where students did a good job and where some still need

more help.

Coach:	What did you notice? When they turned it in what did you think, did you notice anything that stood out to you?
Vanessa:	I think because I had it separated as claim, evidence, reasoning, the claim part was distinct as opposed to sometimes where they kind of... I'm not really sure what their claim is and I think maybe they're not either. So I think having that there felt like it helped them do that part.
Coach:	To answer the question, did evolution occur?
Vanessa:	Right, it either did or it didn't. The evidence part is a little…, I think some of them did a good job, some of them just have a hard time using numbers as data in their evidence that they feel like maybe they're being redundant maybe or I'm not sure exactly what but there's still hesitation to use numbers or to look at the data and say specifically we started with a population of 25 and we ended with a population of 7000.

Coach: You had some kids talk about numbers.

Vanessa: Right but there was a lot more of "it increased, or it didn't increase as much" as opposed to using concreted, specific data. So that's something we've got to work on.
And then on the reasoning part I think there were some kids that showed improvement I think just because it was laid out that way. This was still a trouble spot I think. Even though I felt like I'd done a lot of scaffolding for that, because trying to include a scientific principle like what was the principle we were studying in the lab and on the lab, like natural selection somewhere... or "I calculated allele frequency and that's an indication of this and that". So that piece is still missing.

As she reflects Vanessa has realized that the basic reminder to just use C-E-R might not be sufficient for some students, and we plan our next course of action. Since we are not ready to move this coaching cycle into stage 6, fading, we discuss how she can adjust her scaffolding to help her students write with the detail she is looking for in a second coaching cycle. Vanessa is very keen to keep going with having her students write good scientific explanations. She plans to model the process as she is teaching and to incorporate reasoned explanations into her class discussions:

Well I feel like I need to maybe explicitly model it more, just even in the smallest things, like in any claim. Why am I saying this, this is my evidence, this is my reasoning. This is my therefore statement. I think that will help because seeing it done over and over and over again. There are those that will pick up on that, more than just, I mean I feel like I just did more on this lab than I did before but I think just need to do it with everything that I present to them. Like it's just to... identify like this was my claim, this was my evidence, this was my reasoning. And then help them when we're having a class discussion, maybe to have the same kind of format when they discuss. Perhaps in their group and just say, ok remember that you need these three pieces. (Vanessa, interview).

Vanessa's observations about missing scientific principles inspire the development of SET4CER (chapter 1, and appendix B), a combination of a graphic organizer for writing scientific explanations, with spaces to guide students through

the process, and also the model that Vanessa was hoping for with exemplar explanations on the back page. As she reflected about what her students needed to guide them towards incorporating scientific principles and data to support their claims, SET4CER was expanded to ask questions like *where will you look for your reasoning?* The teacher can use this to check in with students before they actually write their explanation, like a pre-thinking activity to encourage students to look at the bigger picture before they begin writing.

Second Cycle

We began our second coaching cycle using SET4CER as the scaffolding tool for her students as they complete two further summative assessments, the Frog and the GO assignments. Vanessa also incorporated a vocabulary list to help students to "incorporate essential vocabulary into their explanations". In addition, Vanessa continued to model all parts of an explanation in class discussions. Throughout this cycle our communication was mostly electronic, exchanging weekly email conversations over a four and half week period, which included her sharing her thoughts on how students were doing, the effectiveness of the scaffolds she was providing and sharing her graded student work, which I also scored with the SET4CER scoring rubric (appendix C). We discussed improvements in her students' written explanations, the whole time Vanessa remained student focused, always edging them to be better and better, always looking for the little bit extra, always trying to move them into their own ZPDs.

In her reflection at the end of the school year she wrote:

The reasoning part is still a struggle. Most students found the organizer helpful, especially the examples included. The *where will you look for evidence* section was used variously - some students were very specific, others super vague. The *where will you look for reasoning* was very puzzling for some students. I had to reinforce the idea that their notes, labs, textbook (!) might be helpful. Some students thought that as long as they began their final sentence with "therefore" it was all good.

Vanessa's Post-Coaching Approach to Scaffolding Assessments

Vanessa found the coaching process helpful and was especially grateful for the opportunity to reflect, something that she knows is good but has no time to do. She felt that her "assignments have been better for it". With regard to anticipated professional growth, Vanessa was incredibly cooperative and sits high on the spectrum of responsiveness to coaching. She also embraced scaffolding with her students as they produced written scientific explanations. This was especially noticeable in her students' achievements when the SET4CER scaffold was provided, with a clear shift to higher overall scores (chapter one). PD can be considered successful when teachers actually use the new ideas and the results produce an increase in student achievement (Guskey, 1986). A year after coaching, Vanessa self-reported to be continuing to scaffold her students' written explanations: "I am still using the graphic organizer, as a reference mostly, and then providing the essential vocab, which I guess is assignment specific". For all of these reasons Vanessa was classified as *high* in terms of responsiveness to coaching, scaffolding scientific explanations, and achieving anticipated professional growth.

Cycle One	• Plan: Improve CERs with improved GO scaffold • After some success, Vanessa reduces scaffold to a written reminder • Not all students show improved CERs • Reflection: Too soon for some students
Cycle Two	• SET4CER developed • Vanessa models with students • Students use for two assignments and CERs show improvement
Post Coaching	• Vanessa continues to use SET4CER to support students when writing CERs

Figure 3. Schematic representation of Vanessa's coaching cycles

Coaching Henry

I began meeting with Henry approximately one month after I started meeting

with teachers at the school, so that by the time we began our formal coaching cycle

four of the other science teachers were regularly scaffolding assessments and had

incorporated C-E-R into their curricula. Henry, a Health Science graduate, has been a

teacher for fifteen years and has been teaching science at S.B. High for the past

twelve years. He has five sections of Freshman dominated conceptual physics classes,

which includes two small classes of low-track Freshmen students who are part of an

extra time program, where they start the school year eight weeks before the other

students. Henry elected to focus on these early start classes for this research project.

Henry's Pre-Coaching Approach to Scaffolding Assessments

Henry comes across as keen to make sure his students produce their best

work. He allows students to prewrite and also will often allow, or even encourage,

them to redo work: "[For] some of them what I do is I look at it and say are you sure you want to turn that in?" While he presented this to me as him trying to be supportive, this type of feedback can be frustrating for students, especially when he adds "and I tell them do you want to fix that but I don't specifically tell them what it is they need to fix." He also reinforces an environment with a focus on grade rather than learning: "And before you turn it in make sure you just review and think am I satisfied with that or do I want to get a better grade?"

Henry is typical of a teacher who thinks a certain PD is a great idea but has no idea of how to incorporate it into his own curriculum effectively, something that is not unusual (Feiman-Nemser, 2001). He had heard the buzz in school about the C-E-R explanations the other teachers were implementing as some of his students also take biology and they had been talking about writing C-E-Rs, but Henry hadn't taken this any further until our meeting. He is incredibly proud and caring of his students, with intent to support them as much as he can, but in his keenness to encourage the best from his students he frames his support in terms of grades. He seems open to using scaffolds with assessments and was excited to come on board with C-E-R. In our first meeting, we set Henry's anticipated professional growth target to incorporate C-E-R into his assessments beginning with replacing a full lab conclusion with a written scientific explanation.

Henry's Coaching Cycle

In our first meeting Henry explained that his conceptual physics classes had recently begun a unit on Light by conducting a polarization investigation. The next

activity scheduled was a Sunglasses lab where students are required to come up with

an answer to the question *which sunglasses provide the best protection?* Students

bring in sunglasses from home and compare them using a short wave ultraviolet B

(UVB) sensor connected to a data logging program. This activity lends itself as a

great place to incorporate a scientific explanation so I encouraged Henry to focus his

assessment on this rather than a full lab report as he had planned. Henry quickly came

on board with the idea, but he was keen to have students write a procedure before

they start because he wants them to be responsible for coming up with their own

variables to measure in the lab. Henry planned to guide them through this process but

wanted his students to start thinking about the data they are collecting in terms of

measurements and number of data points. We agreed that this would enable him to

talk with his students about how data can provide the evidence to support a claim. In

preparation for implementation of the scaffolded assessment, Henry was tasked with

spending time teaching his students about C-E-R. We discussed how he might do this:

Coach: This is something we did with Roxanne's students but with Football, you could introduce it [C-E-R] using a discussion about the NBA championships. If you ask them who will win, they'll say Golden State will win, right?

Henry: Of course!

Coach: And then you ask them to justify – what's your evidence to support that claim? and they'll likely say because the Warriors have Stef Curry. But that's not enough, it might be true but just having him isn't enough, you need to reason that, you need to explain what it is that makes him the best player and then something along the lines of stats that show teams who have players that can do whatever he does have won the most Championships.

Henry: Oh I see, yes I like that! They would like that!

Coach: It's a discussion, a class discussion, you might find they'll even start arguing with each other, but it introduces that notion that more is needed, you can't just state opinion, and it does it with something they are already very familiar with and current, like basketball, or another sport, or I don't know, you know your kids.

Teachers are more likely to adopt new programs and resources when they are explained thoroughly and the teachers are helped to make them appropriate for their own classrooms and students (Feinman-Nemser, 2001; Knight, 2007; 2009). Since Henry had no prior experience of C-E-R, I also shared with him some resources that we thought would help with these early start students. One of these was some examples of written scientific explanations that he could use in a class discussion with the intention that students could talk about what was a good example and why. I also shared with him the same generic rubric of how a written scientific explanation might be scored on a 4-point scale that I'd given to Vanessa, again with the purpose of students discussing the criterion and also having a point of reference for when they write their own C-E-R explanations. I hadn't been able to observe Henry when he introduced his students to C-E-R so in our next meeting I asked him to talk me through how it went, he told me:

> So, some kids began to talk about it [C-E-R], I think from Bio. And they used the same CER [pronounced surr] and they, it was like they had a better buy in on it. And they made a better connection of this is just how it is and they got on with it. I told them we've been doing this all year, we just didn't call it CER and they were like "oh, alright!". So, I said, look back at your [past] labs and tell me if you've covered these three items in there. And some of them come back and said "no, it was all a claim. We wrote a lot but it was all a claim". Exactly! And you had some reasoning but you didn't put any evidence in there. I told them numbers is everything. And they started paying more attention to it.

Implement. Henry implemented the sunglasses lab with a scaffolded

scientific explanation requirement in place of post lab questions. Students had access

to a generic C-E-R graphic organizer but he elected not to use SET4CER for this

assignment because he "didn't have the opportunity to go over it with them". When

we met to reflect on the assignment we shared excitement at the progress his students

had made overall, and we discussed some of the changes we had noticed in their

writing of a scientific explanation and a way to move forward:

Henry:	I actually think it was one of the better labs that they had done. And I didn't tell them that this was something I was going to give you, which is one reason I didn't get all the copies back. You can probably tell that some still need a lot of help but overall, I think it's good.
Coach:	Can you tell me a bit more about what exactly you thought was good, and why you think it's the best they've done?
Henry:	[In] the shadow lab they just answered some post lab questions. The lab prior to this they did a prewrite and I gave it back to them. This one I decided not to because I just wanted to see the raw data. I was going do the first and last draft with this one then I figured not to do it so I could see true what they had. The fact that they are writing so much is great, and it's all science, there's no opinion in here. Sometimes it's a little off, but the overall is really good.
Coach:	I think for me there were a few [students] that were trying to use some of the data. They were pulling numbers from the data table which is really good. And I know that talking about lux and polarization is not an easy science to explain, but some of them were really trying to talk about it, and you could clearly see that they were thinking about giving reasoning. So even if the reasoning was a little off, you can tell that they were thinking there's a next step, I've got to take a next step. And this is what is exciting for me, I think it's fantastic and a great improvement. Yes, some of them missed a little bit of science but it doesn't matter here because they're looking for that step.

87

> Now we can see that we need to better help them know how to find that step.

Henry had clearly thought about incorporating C-E-R in the assignment and about scaffolding for his students. He had adopted the language and embraced providing them with support to write in a C-E-R format without fear of giving away answers:

> For this class I actually scaffolded it where I put in the title claim, and then evidence, and then reasoning so that just to remind them….before I actually put in a lot of questions for post lab, but now I don't need guided questions, they know that their claim is just a regular or null hypothesis and they just have to follow up with if it was right or wrong and then include the evidence to prove it. And it's no longer about them saying what I thought…. This one [student] is talking about some of the reasons why polarization is good.

It was exciting for me in this conversation to see how impressed he was with how his students had responded to writing their own scientific explanation rather than answering a list of post lab questions. I was encouraged that we were making good progress along the professional growth pathway regarding scientific explanations, although Henry was still presenting a focus on grade achievement to his students:

> I gave them the rubric that you sent. And I told them make sure that you use that it kind of guides you and you can easily tell yourself what grade you're going to get based on where you match in the rubric. And I tell them don't forget make sure you just review it and ask am I satisfied or do I want to get better grade?

Reflection. Coaching Henry has definitely helped him to implement new assessment techniques into his curriculum. He had already heard about the writing scientific explanations professional development, or C-E-Rs as the department had begun to call them, and he was very responsive to the coaching support provided to help him teach his students about them and also about how to incorporate scaffolds

for students to write C-E-Rs as a summative assessment. Also, the work that his students produced in the Sunglasses assignment showed that they too were beginning to embrace the concept of writing an explanation rather than a conclusion or answering some post lab questions. However, when Henry talked about student overall progress he referred to the rubric he'd shared with his students, and his focus remained on the grade his students could earn rather than the level of learning they were demonstrating. Henry clearly has set beliefs centered around the importance of grades, and these beliefs remained intact even as he adopted both a new assessment method and scaffolding into his curriculum. S.B. High school is in the process of transitioning to proficiency scales to replace their grading system, when they finally make the step to proficiency scales Henry will need some continued coaching in this area. Beliefs are difficult to change. A challenge to a belief may result in that belief being proven unsatisfactory and this may lead to replacement, but this is not easy and a number of conditions must be met before an alternative belief can be accommodated (Pajares, 1992). Consequently, Henry may need some help in shifting to a place where he sees his focus on grades as a challenge.

Henry's Post-Coaching Approach to Scaffolding Assessments

Henry was also classified as someone who sits high on the spectrum of responsiveness to coaching. He was easy to work with, and we followed a textbook coaching cycle working collaboratively regarding tools, solutions, and suggested methods that fit into an upcoming assignment. He implemented the scaffolded C-E-R assessment and scored it for his gradebook, he shared student work with me and we

met to reflect. With regard to coaching, the most successful strategy for him was seeing for himself the improvements that his students made. Working together, talking and reflecting on what his students did gave him encouragement to make changes to the way he assessed their understanding during lab write ups. As mentioned earlier, when teachers actually use the PD and the PD results in an increase in student achievement then the likelihood of changing practices are increased (Guskey, 1986).

In other areas of the spectrum Henry was classified as being in the mid-range because although he committed to the notions of scaffolding and scientific explanations, he remained more focused on student achievement rather than learning, and he might have rated higher if he talked less about the grade outcome and more about the depth of learning. Although in this cycle we hadn't the opportunity for him to use SET4CER with his students, in a one year follow-up post coaching meeting Henry reported that he had used it at the start of the year with his current group of students and planned to continue to use it.

Cycle One

- Plan: Replace post-lab questions with scaffolded scientific explanation
- Henry teaches his students about CER and provides scaffolded GO and rubric
- Students embrace written CERs. Henry impressed with their progress.
- Change successfully implemented but Henry remains grade focused.

Post Coaching

- Henry continues to request scientific explanations in place of post-lab questions
- Henry has incorporated the SET4CER scaffold into his curriculum

Figure 4. Schematic representation of coaching Henry

Coaching Roxanne

Roxanne has a Bachelor's degree in biology and a Master's degree in Education and has been teaching biology for twelve years. Currently in her eighth year at S.B. High School, she is teaching regular and honors biology to classes comprising mostly sophomores and freshmen students. For this study Roxanne elected to work with two small classes (n=10 and n=15) of low tracked students. In the previous school year most of these students had been part of the early start program described above. Like Vanessa, she was part way through an eight-week long Evolution and Ecology unit and her students were also learning how to produce cladograms. This is the first year that Roxanne has been assigned to teach the lower track students and she had been finding it very challenging. She felt frustrated and unsupported, and consequently was keen to begin meeting right away.

Roxanne's Pre-Coaching Approach to Scaffolding Assessments

Roxanne's frustration with these students was immediately evident. She describes them as unmotivated and often producing little or no work such that most of them earned Fs in the first semester. It is her understanding that they are expected (by the school) to get to the same place as her other biology students by the end of the year, but she is concerned that applying scaffolding techniques during assessments for just her low track students would be unfair to her other classes. Currently Roxanne does not use any scaffolding techniques and all students study the same curriculum and take the same assessments: "I can't give them different tests or anything like that because I can't modify, it's the same class. I'm supposed to give them differentiated instruction."

Roxanne's Coaching Cycle

Roxanne and I attempted three coaching cycles. In terms of producing anticipated professional growth each one was considered a failed cycle, but with each one the potential for success increased.

Cycle one – failed cycle. My first formal meeting with Roxanne revealed more about her frustrations. She comes across as passionate about her subject but was becoming increasingly unhappy with the effort from this group: "I try to make it fun and it's disappointing to me when they say "this is too hard". You know one kid today was like "I don't want to highlight", and I was like really? you're complaining about highlighting?" She had tried several motivation techniques from "trying to be super

positive and encouraging" and "just mama bearing them" with little success. Now she had resorted to offering easy participation points:

> I told them to bring in candy [for the classification lab] … I said next class bring in, for the first period, ten pieces of candy. I was like for you guys, you got a small class you could bring in full size candy bars. They didn't like that idea. But then I said, because they're buying something I'd give them extra credit. Five points extra credit, that's a lot. And you know I had kids going "Five points that's all!!" I was just like, are you joking? (Roxanne, interview)

She also awards her students five points for completing a *check-in*, which she described as a formative type of assessment she gives them about every two weeks to check on their understanding of concepts during a unit. Roxanne decided to offer points for check-ins to encourage her students to do them: "any kind of activity they do I try to give them a little something to wrap up, at least a journal or something along with it so that they can 1) have points and 2) otherwise they won't do it."

We talked a lot in this first meeting about scaffolding techniques, in particular ones that support writing. I showed her samples of connecting and focus sentence frames and a generic C-E-R rubric, and we discussed how to incorporate these to scaffold her students in writing a conclusion for the upcoming candy classification lab. As Roxanne was new to scientific explanations and was struggling to get her students to write anything at all we decided to focus only on a short summary conclusion for the candy lab. Roxanne was tasked with either creating a new assessment or modifying her current assessment to include a conclusion section incorporating sentence frames in an attempt to support them into writing. Roxanne did not feel confident teaching her students about C-E-R so we agreed that I would do that during an upcoming lesson that she would observe (see cycle 2 below). However,

Roxanne did not implement the assessment with scaffolding. When I asked her about this she was vague and it became apparent that instead she had used the assessment that Vanessa had created for her students which required a full written C-E-R conclusion and offered no sentence frames.

As a coach, I struggled to get Roxanne to collaborate with me in this first cycle, and I had failed to persuade her of the benefits of scaffolding such that she had not tried it with her students. Changes in teacher attitudes are thought to come about after experiencing improvement in student learning (Guskey, 2002) and I had hoped that beginning with sentence frames would start the path to student improvement which would lead to Roxanne's buy-in to scaffolding. Although I was aware that Roxanne was cautious about implementing scaffolding techniques into her assessments it was disappointing that I had not been able to foster any change in her approach at all.

Cycle two – failed cycle – but with a recovery plan. As we reflected on the failed first cycle and discussed how best to move forward, Roxanne and I determined that I would introduce her students to C-E-R and she promised she would create a scientific explanation assessment with sentence frames for her next assignment. Her students responded well to the lesson, and although many were initially reluctant to participate in a class discussion, eventually all had demonstrated some understanding of the concept of a scientific explanation with claim, evidence, and reasoning. In this second cycle I asked Roxanne to incorporate scientific explanations into several assessments, both formative and summative; and although she still did not incorporate

sentence frames or any other scaffolding techniques into the actual assessments, Roxanne did begin to say to her students "use C-E-R like Ms. C. taught you."

Roxanne clearly finds it difficult to scaffold her assessments yet she feels comfortable in offering participation points because "they [students] are still doing something for the points". At this point I was still unsure if Roxanne is resistant to scaffolding or if there are other reasons why she says she will use them but when it comes down to it, she backs out completely or she uses assessments created by someone else, which often are not appropriate for her own students. Much has been studied and written about teacher beliefs and while it is not the intention of this study to frame analysis within teacher beliefs, Roxanne's behavior and the way she responds to coaching are likely to be driven by her innate beliefs about teaching and assessment. According to Pajares, (1992) beliefs can be descriptive, evaluative or prescriptive but are usually somewhat of a combination of all of these. This is evident with Roxanne: she is not enjoying teaching these low achieving students (evaluative), and she feels under pressure by administration to get them to the same level as the other students (prescriptive), but scaffolding an assessment is not appropriate (descriptive). Guskey (1986) says that PD is likely to be unsuccessful in changing teacher beliefs unless the teacher can be persuaded to use the PD and of course, the PD can be shown to improve achievement for the teacher's own students. With this in mind, I decided to begin to group email Roxanne and Vanessa regarding scaffolding for the approaching bird beak lab. By being included in conversations back and forth between Vanessa and myself, Roxanne, although she did not contribute, was able to

see that Vanessa was embracing the concept of scaffolding, was incorporating it into her scientific explanation assessments, and was experiencing positive results in her students' achievement. This allowed us to move into a place where, although she still would not create scaffolded assessments for her students, Roxanne did begin to regularly use Vanessa's scaffolded assessments with her own students, which included access to the C-E-R specific graphic organizer, SET4CER. We met again after students had completed the bird beak lab and had also taken a quiz using data from the lab. I asked her if she thought the scaffolding had helped for the lab, she told me "It did, at least they wrote a claim. Didn't everybody write a claim?" In fact, all of her students who submitted work produced some kind of explanation that included a claim. Most had attempted to use some evidence and a few had attempted a reasoning statement as well. For the quiz, instead of using SET4CER she had reduced the scaffolding to just a written reminder for them to use C-E-R, and was disappointed when students had not performed well:

> And this time they said they couldn't remember what C-E-R meant… some of them wrote claim, results…..but nobody said a word during the quiz. Because that's the way they are, if they don't know or don't understand they think that they should have known and should have understood so they won't ask a question. (Roxanne, interview)

In this meeting I was very careful to point out improvements in her students' written scientific explanations when they had access to SET4CER and aimed to persuade her to continue to use it. But while Roxanne had begun to incorporate some very basic scaffold reminders into her assessments, her resistance to fully buy in made achieving a successful coaching cycle very difficult. So far, I had been unable

to effect any collaboration from her in creating scaffolds. In our meetings, she agrees to create a new scaffolded assessment, or to modify a scaffold I give her to be appropriate for her students but then she doesn't actually do it. Collaboration is not only a key phase of the coaching cycle, it has been described as central to successful professional development (Chan & Pang, 2006; Wilson & Berne, 1999), and collaborative practices provide the social interactions that facilitate knowledge production (Lave & Wenger, 1991). Roxanne's reluctance to take productive steps in the collaboration relationship were puzzling since she had expressed such eagerness to work with me. Her reluctance might be attributed to an innate view of professional development being delivered by the *expert* such that she either expected to be handed a complete set of ready scaffolded assessments to administer or a fear that engagement in the collaborative process might reveal deficiencies in her abilities as a teacher. Such beliefs are instilled from a long tradition of the institutional conditions of schools and teaching (Scott, 2008). Teachers have historically taught to Standards and are held accountable for student achievement, by asking Roxanne to collaborate with me on scaffolding assessments the tradition became disrupted which may have left her feeling exposed and vulnerable.

Cycle three – an almost successful cycle. Throughout the remainder of the school year, during which time Roxanne and I met on several more occasions, she remained frustrated with the output from this group of students. She did not feel the same about her other classes but she had almost given up on these and was in a place where "I just want the year to be over as much as they do". By this point, Roxanne

had started to talk about scaffolding "I could scaffold it for them even more…" but the level of support she was providing was minimal and much less than these students needed to make the improvements we were looking for. Aiming for one good C-E-R by the end of the year we moved into a food web assignment. Roxanne had again agreed to rewrite the assessment so that students would produce an explanation and she would provide them with the SET4CER scaffold. In actuality, the explanation requirement was added to the voluntary extra credit part of the assignment and the only scaffold was a written reminder to "use the C-E-R format to write your explanation", SET4CER was not made available. Consequently, scores were poor. We met to discuss, after which Roxanne then gave her students SET4CER and they re-answered a question from the packet. The quality of the students' explanations was much better but Roxanne wouldn't look at their scores, she just handed me their completed work and had totally lost interest in their progress. As Guskey (2002) says, the key to successful PD is when teachers experience improvements in student outcomes. "They believe it works because they have seen it work, and that experience shapes their attitudes and beliefs" (Guskey, 2002, p.383) so I was a little disheartened when I was unable to get her to look at the student work with me, even when I told her that there were marked improvements for many students.

Roxanne's Post-Coaching Approach to Scaffolding Assessments

Coaching Roxanne was difficult. In our discussions she presented as engaged and ready to collaborate but often this translated into little or no action on promises she made. Consequently, progress towards anticipated professional growth through

coaching was slow and somewhat ineffective. As someone who scored low on the

Teacher Take Up Rubric (TTUR) and who remained at the low end of the spectrum

throughout the coaching process, coaching cannot be viewed as a successful tool for

PD integration with Roxanne. However, she did eventually make SET4CER available

to her students, and it was helpful in developing their scientific explanations. She also

started to use scaffolding language after being included in group emails with Vanessa.

About a year and a half after our coaching sessions I asked her how she feels about

scaffolding. She told me:

> It is always a great idea. It is often times more work for me of course. It really
> ends up helping ensure everyone can accomplish the same tasks with closer to
> the same level of comprehension in the end. (Roxanne, post coaching
> reflection, summer 2017)

Her *scaffolding* still only consists largely of written reminders which she thinks is

sufficient: "Items are written in the agenda, in the instructions for assignments, in the

rubric for the assignment. The info is everywhere". But she is still using SET4CER

and the transition word document with her students "I use them as they were given to

me, I haven't touched any of them" (Roxanne, post coaching reflection). When asked

about the scientific explanations that her students produce she said "as far as

scientific explanation go they still struggle. I still struggle when reading/grading them

also."

Cycle one	• Plan: Incorporate simple sentence frame scaffolds into upcoming assessment • Not implemented.
Cycle two	• Plan: Incorporate sentence frames inot upcoming assessment that incorporates a CER response • Coach teaches students about CER • Scaffolds not implemented into the assessment by Roxanne but she does begin to use CER language with her students
Cycle three	• Plan: Aim for one good written CER by end of school year • Continued coaching support • Students show improvement using SET4CER • Roxanne reluctant to participate in reflective pratice
Post coaching	• Self reports as continuing to use scaffolding tools • Roxanne acknowledges that she still struggles with the concept of written scientific explanations

Figure 5. Schematic representation of Roxanne's coaching cycles

Discussion

Teachers participate in professional development for many reasons ranging from a self-recognized need or interest in a particular area to school or district mandated requirements. Yet, even if teachers see value in the PD they receive, many often fail to translate the new ideas into their own classrooms to use with their own students. The larger study that these case studies were drawn from investigated the potential of coaching as a method of guiding successful adoption and integration of a certain PD with a group of science teachers. The aim was for teachers to adopt the practice of scaffolding their students during assessments that require a scientific explanation after receiving specific and individualized coaching in how to create and

implement such scaffolds, and in teaching their students about C-E-R style explanations.

Traditionally PD is introduced during staff meetings or workshops and teachers are expected to take away and self-implement what they have learned into their own classrooms, often without any support. However, teachers are unlikely to adopt a new practice if they have only experienced it once without any form of follow up (Knight & Cornett, 2008; Knight, 2009). In fact, as Knight goes on to say "(t)he preliminary research on coaching suggests that teachers rarely implement without sufficient support involving precise explanations, modeling, and encouraging feedback" (Knight, 2009, p.512). The teachers in this study received detailed, specific, and individually tailored coaching. While the use of coaching to implement PD was found to be helpful for the teachers in this study, their response to, participation in the coaching process, and their degree of implementation and adoption of the PD remained varied.

Measuring Successful PD

It is important to keep in mind that the ultimate goal of teacher professional development is an improvement in student learning (Guskey, 2003). Chapter one looked in detail at changes in student achievement in written scientific explanations with and without the SET4CER scaffold. Student achievement was shown to increase when the scaffolding tool was used which indicates successful PD. But the measure of success is surely more complex? If the ultimate goal is improvement in student learning, what other goals must be met before this can be achieved? Roxanne's

students showed improvement when scaffolded but she was reluctant to scaffold their learning and only did so when the tools were provided for her. One could argue that coaching was effective because she did eventually use the target PD, and her students did show improvement, but has the issue of her struggle to incorporate scaffolding and scientific explanations into her own curriculum been resolved? Also, what about her own professional learning? In her end of year reflection, she said that she still finds interpreting C-E-Rs difficult. Likewise, for Henry gaps remain in his professional learning. He is still using SET4CER and has transitioned to incorporate more C-E-R scientific explanations into his curriculum and assessments but he still drives his teaching by grades. Clearly in order to achieve that ultimate goal of student learning then first must come effective teacher learning, and also a better understanding of connections between them.

Another factor worthy of consideration is the extended picture of scaffolding during other assessments. There are many other types of scaffold and many other practices to demonstrate learning besides writing a scientific explanation. If the S.B. High teachers have only adopted SET4CER is that sufficient to say that they have fully taken on board the notion of scaffolding in assessments? For example, while coaching produced a spectrum of success with adopting scaffolded scientific explanations, none of these three teachers have explored using any other assessment scaffolding techniques. Chapter three examines coaching as a method of delivering extended assessment scaffolding PD in more detail with Jessica, another teacher at S.B. High who demonstrates full adoption of incorporating assessment scaffolds.

What is interesting for discussion is how differently the teachers responded to the coaching, and consequently the PD adoption even when students were shown to have made improvements. The guiding questions I asked as I analyzed the data from these teachers were:

1. Was the coaching successful in producing anticipated professional growth?

2. How do teachers respond to individualized coaching for a specific PD?

3. Are there particular coaching strategies that are more or less effective in ensuring successful adoption of said PD?

Anticipated Professional Growth

The teachers received very individualized coaching in how to scaffold students as they wrote scientific explanations for a summative assessment, tailored to their own curriculum. Embracing Jim Knight's comment above, it is reasonable to expect therefore that the anticipated professional growth for all three case study teachers would be successful implementation and continued adoption of the PD: scaffolding their students' written scientific explanations. However, as this study has shown, the level of adoption and the degree of responsiveness to coaching, and thus the success of anticipated professional growth lies on a spectrum. I investigated coaching as a method of helping teachers adopt new practices that traditionally would have been introduced at a PD session. Even though the teachers that I worked with clearly had different degrees of responsiveness to the coaching, I am confident that without it, Henry and Roxanne would not have attempted to implement scaffolding

techniques into their assessments. Also, they were both unfamiliar with C-E-R written scientific explanations even though *constructing explanations (for science)* is one of the eight science and engineering practices identified by the Next Generation Science Standards (NGSS) Framework as essential for all students (NGSS lead states, 2013), and at grades 9-12 the Framework states that students should "Apply scientific reasoning, theory, and/or models to link evidence to the claims to assess the extent to which the reasoning and data support the explanation or conclusion" (NGSS lead states, 2013, Appendix F, p.11). At the time of this study, the new science standards had been available to these teachers for over two years, along with many resources both online and in print, yet two of these S.B. High teachers and their students were unfamiliar with the practice of a written scientific explanation. Even though their responsiveness to coaching was lower on the spectrum than other science teachers at this school, the coaching that Henry and Roxanne received served to introduce the concept and facilitate integration into their curriculum.

Coaching Strategies and Teacher Responsiveness

Adopting new Standards is not always easy, especially if teachers are required to make significant adjustments to the details of their curriculum and also perhaps their teaching style. Viewed through the lens of the sociocultural learning theory, when a teacher is finding it difficult to make adjustments and embrace the new focus, like Roxanne and Henry, it is possible that the new learning is out of reach of their ZPD. When this happens the resources (artifacts) designed for semiotic mediation, i.e.

NGSS texts and online resources, will not facilitate the intended assistance (Wells, 1999).

Within the sociocultural theory, learning and the ZPD are usually discussed in terms of working with a more capable other. For the teachers in this study, the most obvious "other" is the coach, but for the coach to be successful in this role, there also has to be some level of participation from the teacher. Looking at the experience with Roxanne in this way indicates that although we met many times, we were unable to establish an environment where she would follow through on promises she made because we were not working within her ZPD. This situation became frustrating and confusing such that the working relationship became one where neither were learning from the other. As Wells puts it: "…for learning to occur in the ZPD, it is not so much a more capable other that is required as a willingness on the part of all participants to learn from each other" (Wells, 1999, p.324). When successful, a coaching relationship will provide appropriation for both parties, with both learning something from the other through joint productive activity (Moschovich, 2004). But when the joint activity becomes unproductive the goals of both are separated and the partnership becomes ineffective (Ash & Levitt, 2003). This was evident in the early cycles with Roxanne, we struggled to reach a place of mutual appropriation so our joint activity required more meetings and smaller steps than the other teachers before we were able to make any, albeit small, progress. Working within the ZPD is clearly nuanced and very individualized. The ZPD is an area close in proximity to existing knowledge and/or practice, where the learner is most likely to learn so long as the

new information stretches them just beyond where learning easy without over stretching them into an area where learning is too difficult. This area, or zone, is different for individuals and also changes with different situations (Vygotsky, 1978), which accounts for the variations experienced in this study regarding the teachers' responsiveness to coaching.

According to sociocultural theory, learners construct meaning through shared interactions. To be effective the coaching cycle requires collaboration, reflection, and the desire to participate from both parties, but the concept of the individuality of the ZPD is something that should be kept to the forefront in coaching relationships. Consequently, finding a place of joint productive activity within the teacher's ZPD, while not uniquely or easily established with all teachers, is the key to ensuring a pathway towards anticipated professional growth (Ash & Levitt, 2003).

Although unsurprising, the varied response and engagement with the coaching that I experienced did, however, lead me to ask deeper question about the effectiveness of coaching strategies. I asked: *are there particular coaching strategies that are more or less effective in ensuring successful adoption of said PD?*

Henry responded extremely well to feedback, particularly encouragement that he was doing a good job with communicating C-E-R. Discussing improvements in the depth of his students' scaffolded explanations was also an effective coaching strategy for him. Reflection is an important part of the coaching cycle, and while reflective discussion was also extremely effective for Vanessa, Roxanne switched off during these discussions and avoided looking at student work. Dewey (1933) argued that

teacher reflective practices are essential due to the complexity of the work that they do; and studies of the importance of reflection for teachers has found it to be a critical component in facilitating change in practice that can provide insights which may be unobtainable any other way (Baird, Fensham, Gunstone, & White, 1991).

How teachers engage in reflection is influenced by their existing beliefs (Abell, Bryan, & Anderson, 1998), so to encourage successful reflection an understanding of their personal theories about teaching and learning is required, an area that would provide a good extension to this study. Prior research in this area has mainly been carried out with preservice teachers, but arguably examining the reflective practices of in-service teachers is equally important especially when one is trying to implement change. The process of reflection can be an effective tool to identify and confront personal theories, and personal histories can have a big influence on how teachers approach the reflective process (Abell et al, 1998). The reflection phase of the coaching cycle provides space where the coach hears the teachers voice, such that the teacher's specific needs can be identified and incorporated into forward cycles, ensuring the PD remains considerate of their needs and focused on their development (Anderson & Olson, 2006). While reflection was successful for the other teachers in this study, with Roxanne reflection was largely ineffective. My responsibility as a coach was to find/tweak the reflective process to find a way that engaged her, or to incorporate tools that better help her where discussion failed. For example, making use of a journal or perhaps including peers in the discussion process.

Just like a one-size-fits-all approach to working in the ZPD is futile, likewise a narrow approach to creating a list of effective coaching strategies is also unproductive. The strategies of successful coaching need to be differentiated to the needs of the individual teacher and need to adjust as the teachers' needs adjust (Stover, Kissel, Haag, & Shoniker, 2011). The key is differentiation, strategies that are successful for one may be unsuccessful for another. Coaching is a social practice, and successful coaching is influenced by the coach's ability to work out how to inspire the teachers they are coaching to make changes in their thinking and teaching, while simultaneously drawing out and developing the best in them (Rainville & Jones, 2008). Rainvile and Jones (2008) posit that this involves the adoption of various identities by the coach, based upon the individual situations they find themselves in. Responding to a social situation by assuming an appropriate identity is described by Gee (1999) as "situated identity" and can be used to explore the nuances of working in the ZPD. If coaching is scaffolding in the ZPD the essential first step to reach and remain there is the relationship that is created and developed between coach and teacher. If looking for "effective strategies" the most important one has to be creating and maintaining this individualized yet dynamic relationship. Can the coach make the teacher feel at ease, be responsive, be engaging, embrace the PD? The coach needs to assess and react to each situation in a manner that maintains a creative and trusting environment with each teacher (Knight, 2007; Rainville & Jones, 2008). This is made more complex because every meeting may require a new identity, or even multiple new identities with and between teachers. For example,

Henry is also the cross-country coach so our meetings often began with me asking him about the latest race before we discussed student responses to an assessment.

While I may have started out looking for a tool box of effective coaching strategies, as a coach it is important to remember that the phases of coaching applied individually to each teacher are sufficient tools to build the trusting collaborative relationship required to allow the coaching process to be successful. This study also found that while it clearly helped with PD delivery, the coaching process is complex. Teacher attitudes, abilities, knowledge and beliefs are all involved in how they receive and process information. Consequently, it is critical to attend to all of these aspects when considering coaching as a tool for ensuring PD adoption.

Case Study Limitations

The teachers presented in this case study were selected because they represented different parts of a spectrum of responsiveness to the coaching they received as PD to implement scaffolded scientific explanation assessments. I have undertaken thorough qualitative analysis but it is important to keep in mind that the observations made are from only three teachers at one school. While a case study approach provides an effective way to reveal and examine the complexities of an educational situation, a small sample size and the absence of numerical data as supporting evidence mean the probability that the data is representative of or transferable to the larger population cannot be claimed. Another recognized limitation of case study research is that the large amounts of data generated can make analysis difficult, and it is often difficult to present the complexities and depth of the situation

in writing, which makes summarizing the findings difficult. I have attempted to address these limitations with careful thought about what data to include and how to present and discuss the issues I experienced while remaining objective and being considerate of potential biases from researcher effect (Miles & Huberman,1994).

It was not the intention to interpret analysis based upon teacher beliefs and how they guide teacher decisions, however, the tensions experienced during the coaching cycles means that the influences of teacher beliefs and also the institutionalization of schools cannot be ignored. Multiple sources of data were analyzed and triangulated by data source and method (Miles & Huberman, 1994), but a longer, more detailed analysis investigating strategies that might help resolve some of these tensions might be worthwhile. As mentioned above, the information gleaned from case studies is not intended to be generalizable, and that is also the intention with this study. However, with regard to the much-needed research on the potential for coaching as a method of PD delivery the richness of detail that case study analysis provides is invaluable.

References

Abell, S. K., Bryan, L. A., & Anderson, M. A. (1998). Investigating preservice elementary science teacher reflective thinking using integrated media case-based instruction in elementary science teacher preparation. *Science Teacher Education, 82*(4), 491-509.

Anderson, L., & Olsen, B. (2006). Investigating early career urban teachers' perspectives on and experiences in professional development. *Journal of Teacher Education, 57*(4), 359-377.

Ash, D., & Levitt, K. (2003). Working within the zone of proximal development: Formative assessment as professional development. *Journal of Science Teacher Education, 14*(1), 23-48. Downloaded from: springer.com.

Baird, J. R., Fensham, P. J., Gunstone, R. F., & White, R. T. (1991). The importance of reflection in improving science teaching and learning. *Journal of Research in Science Teaching*, *28*(2), 163-182. doi: 10.1002/tea.3660280207. Downloaded from researchgate.net.

BSCS.org. (2012). Scientific Explanation Tool. Retrieved from: https://bscs.org/tools/scientificexplanationtool.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (2000*).* Learning: From speculation to science. In *How people learn: Brain, mind, experience, and school*. Washington, DC: National Academy Press.

Chan, C.K.K., & Pang, M. F. (2006). Teacher collaboration in learning communities. *Teaching Education*, *17*(1), 1-5.

Cornett, J., & Knight, J. (2009). Research on Coaching. *Coaching: Approaches and Perspectives*, 192–216. Downloaded from: https://resources.corwin.com/sites/default/files/Cornett_Knight_2008.pdf.

Dewey, J. (1933). *How we think: A restatement of the relation of reflective thinking to the educative process*. Chicago: D.C. Heath.

Feiman-Nemser, S. (2001). From preparation to practice: Designing a continuum to strengthen and sustain teaching. *Teachers College Record, 103*(6), 1013-1055.

Fletcher, S. J. (2012). *Coaching: An Overview*. In Fletcher, S. & Mullen, C. (Eds.) The SAGE Handbook of Mentoring and Coaching in Education (pp.24-40). Thousand Oaks, CA: Sage Publications, Inc.

Gee, J. P. (1999). Critical issues: Reading and the new literacy studies: Reframing the National Academy of Sciences report on reading. *Journal of Literacy Research, 31*(3), 355-374.

Guskey, T. R. (1986). Staff development and the process of teacher change. *Educational Researcher. 15*(5), 5-12.

Guskey, T. R. (2002). Professional development and teacher change. *Teachers and teaching, 8*(3), 381-39. doi: 10.1080/135406002100000512.

Guskey, T. R. (2003). What makes professional development effective? *Phi Delta Kappan, 84*(10), 748-750. doi: 10.1177/003172170308401007.

Knight, J. (2007). *Instructional Coaching: A partnership approach to improving instruction*. Thousand Oaks, CA: Corwin Press.

Knight, J. (2009). What can we do about teacher resistance? *Phi Delta Kappan*, *90*(7), 508-513. doi:10.1177/003172170909000711.

Knight, J., & Cornett, J. (2008). Studying the impact of instructional coaching. Downloaded from http://instructionalcoach.org/research/tools/paper-studying-the-impact-of-instructional-coaching.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge University Press, Cambridge.

Lichtman, M. (2010). *Qualitative research in education: A user's guide*. Los Angeles, CA: Sage Publications Inc.

McNeill, K. L., & Krajcik, J. (2008a). Scientific explanations: Characterizing and evaluating the effects of teachers' instructional practices on student learning. *Journal of Research in Science Teaching, 45*(1), 53-78. doi:10.1002/tea.20201.

McNeill, K. L., & Krajcik, J. (2008b). Inquiry and scientific explanations: Helping students use evidence and reasoning. Downloaded from www.katherinelmcneill.com

McNeill, K.L., Lizotte, D. J., Krajcik, J., & Marx, R.W. (2006). Supporting students' construction of scientific explanations by fading scaffolds in instructional materials. *The Journal of Learning Sciences, 15*(2), 153-191.doi: 10.1207/ s15327809jls1502_1

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook*. Thousand Oaks, CA: Sage Publications Inc.

Moschkovich, J. N. (2004). Appropriating mathematical practices: A case study of learning to use and explore functions through interactions with a tutor. *Educational Studies in Mathematics, 55*(1/3), 49-80.

NGSS Lead States. (2013). *Next Generation Science Standards: For States, By States*. Washington, DC: The National Academies Press.

Pajares, M. F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct, *Review of Educational Research, 62*(3), 307-332.

Patton, M. Q. (2002). *Qualitative research & evaluation methods*. Thousand Oaks, CA: Sage Publications, Inc.

Rainville, K. N., & Jones, S. (2008). Situated identities: Power and positioning in the work of a literacy coach. *The Reading Teacher, 61*(6), 440-448. doi: 0.1598/RT.61.6.1.

Sandoval, W. A., & Millwood, K. A. (2005) The quality of students' use of evidence in written scientific explanations, *Cognition and Instruction*, *23*(1), 23-55. doi:10.1207/s1532690xci2301_2.

Scott, W. R. (2008). *Institutions and organizations*. Los Angeles, CA: Sage.

Stover, K., Kissel, B., Haag, K., & Shoniker, R. (2011). Differentiated coaching: Fostering reflection with teachers. *The Reading Teacher, 64*(7), 498-509. doi: 10.1598/RT.64.7.3.

Tharp, R. G. & Gallimore, R. (1988). *Rousing Minds to Life: Teaching, Learning and Schooling in Social Context*. New York. Cambridge University Press.

Vygotsky, L. S. (1978*). Mind in Society: The development of higher psychological processes*. London: Cambridge University Press.

Wells, G. (1999)*. Dialogic Inquiry: Towards a sociocultural practice and theory of education*. New York: Cambridge University Press.

Wilson, S. M., & Berne, J. (1999). Teacher learning and the acquisition of professional knowledge: An examination of research on contemporary professional development. *Review of Research in Education*, *24*(1), 173–209. doi:10.3102/0091732X024001173.

Zhang, Z., & Burry-Stock, J. A. (2003). Classroom assessment practices and teachers' self- perceived assessment skills. *Applied Measurement in Education, 16*(4), 323-342. doi: 10.1207/S15324818AME1604.

Appendix A
Teacher Take Up Rubric

Component	Levels			
	0	**1**	**2**	**3**
Coaching Level of take up by teacher. Interest, response, action	No participation. Does not participate in coaching. May have participated in baseline data collection but no interviews or further meetings	Minimal take up. Participates in initial interview. Difficult to schedule. Delayed responses. Blames external reasons. Unprepared	Moderate take up. Participates in at least one follow-up meeting. Somewhat difficult to schedule meetings and/or delayed response from teacher. Partial implementation	Full take up. Several meetings with focused agenda and good progress. Implementation of items discussed. Is prepared for meeting.
Scaffolding Level of take up by teacher. Support/belief or views of scaffolding.	No participation. Does not value use of scaffolds in assessments involving scientific explanations. May view scaffolds as 'helping' or 'cheating'. Unable to differentiate assessments.	Minimal take up. Somewhat values scaffolds as useful. Small amount of use if provided by researcher. No co-design. T reports no real gain from using and/or may not continue.	Moderate take up. Values scaffolds as supports. Tries to implement into CER assignments when guided by researcher. Participates in co-creation or chooses own scaffolds.	Full take up. Incorporates a variety of scaffolds into a variety of CER type assessments. Tries multiple scaffolds Reports positively.
Scientific Explanations Level of take up by teacher. Using CER with students. Importance of SEs	Few or no assessments that require students to give written scientific explanations or no focus on CER.	Minimal. Includes SEs but minimal focus on CER. e.g. might remind students to use CER but no further scaffolding provided.	Moderate. Good CER focus but only in few summative assessments, eg only lab reports [and/or may provide rubric or go]	Full. Requires CER in all or almost all summative assessments. [and provides scaffolds such as rubric or go]

Appendix B

Developing a Scientific Explanation using CER

What is the question that you want to answer?
What evidence do you need to answer this question?
Where will you look to find the scientific principles to provide reasoning?

Support for your explanation

Evidence	Scientific Principles

Scientific explanation = Claim + Evidence + Scientific Reasoning (CER)
My claim is __(answer to the question)__ because __(Evidence followed by scientific reasoning)__

Question:	Examples: - Did a chemical reaction occur? - How would the hawk population change if there were no seeds available? - Can animals survive in the desert? - Do levers make work easier?
Claim: *A statement that answers the original question.*	Examples: - A chemical reaction occurred. - The hawk population would decrease if there were no seeds. - Only animals that are adapted to desert conditions are able to survive in the desert. - Levers can sometimes make work easier.

Evidence: _Specific_ _data that supports the_ _claim, such as numbers from a_ _data table or observations._ _Data needs to be appropriate_ _and you need_ _enough_ _data to_ _support your claim (e.g. high_ _and_ _low numbers from the data_ _table)_	Examples: - The density of butanol is 0.81g/cm^3, whereas the density of layer A is 0.87 g/cm^3 - The hawk eats squirrels, rabbits and sparrows. Squirrels only eat seeds. Sparrows eat seeds and grasshoppers. - Camels have humps to store fat, they don't sweat much, they have really big feet and they have fur. - When the effort distance was 60 cm the effort force was 2.5N, but when the effort distance was 20 cm the effort force was 8N.
Reasoning: _Justification that links the claim_ _and the evidence. Include_ _appropriate and sufficient_ _scientific principles._ _This is where you include the_ _science you have learned. Look_ _in your class notes, resources_ _provided by the teacher,_ _introductory or background text_ _for a lab, or other available_ _texts._	Examples: - Density is a property of a substance along with melting point and color. In a chemical reaction new substances are formed with different properties to the original substances. The differences in densities between butanol and layer A indicate that a chemical reaction occurred and that layer A is a new substance. - Organisms in a food web are affected by other organisms in the same food web even if they are not directly linked to them. Without seeds the squirrels would have no food so they would die out and the sparrows would have less food so their numbers would decrease. This means there will be fewer hawks because they will have less food to eat. - Camels have adaptations that help them survive in desert conditions. By storing fat in their humps they are able to go longer without food and because they don't sweat much they lose less water from their bodies and can go longer without water. Their big feet make it easier to walk on the sand and their fur keeps them warm at night as it can get extremely cold in the desert at night. - Doing work is the ability to move an object. If it takes less force the work feels easier. A lever can help the work feel easier depending on the position of the fulcrum, effort, and load.

Adapted from BSCS.org (2012)

Appendix C

Scientific Explanation Rubric

Component	Levels			
	0	**1**	**2**	**3**
Claim A statement or conclusion that answers the original question.	Does not make a claim.	Makes an inaccurate claim.	Makes an accurate but vague or incomplete claim.	Makes an accurate <u>and</u> complete claim.
Evidence Scientific data that supports the claim. The data needs to be appropriate and sufficient to support the claim.	Does not provide evidence.	Provides inappropriate evidence (evidence that does not support the claim).	Provides appropriate, but insufficient evidence to support the claim. May include some inappropriate evidence.	Provides appropriate <u>and</u> sufficient evidence to support the claim. E.g. - specific values from a data table and/or graph. - observations - supporting text from a scientific reading
Reasoning A justification that links the evidence to the claim. It shows why the data counts as evidence to support the claim by using the appropriate and sufficient scientific principles.	Does not provide reasoning.	Only provides reasoning that does not link evidence to the claim. Or repeats claim or evidence.	Provides some reasoning that links the claim and evidence. E.g. Includes some scientific principles, but not sufficient and/or detailed enough to fully support claim.	Provides reasoning that links evidence to the claim. Includes appropriate <u>and</u> sufficient scientific principles. E.g. - theory behind observed values or changes in values during an investigation - uses notes, text book, handout or other suitable sources to provide scientific principles that explain how the evidence supports the claim.

Adapted from McNeill, Lizotte, Krajcik, & Marx (2006)

CHAPTER THREE

TEACHER DESIGNED SUMMATIVE ASSESSMENT SCAFFOLDS IN SCIENCE: STUDENT INTERACTIONS AND ACHIEVEMENTS

Abstract

Research of scaffolding in education is most often presented through best practices that highlight the success stories of a particular scaffold, see for example, Robinson's (1997) review of graphic organizer research. Yet much can be learned about scaffold design if the focus is not only on the students who do well when supported by a particular scaffold, but also includes those who do not behave as expected when the scaffold is present (Pea, 2004). This is a study of how all 27 students of a high school kinesiology class interact with teacher designed assessment scaffolds intended to support them as they demonstrate their learning in a complex and challenging biomechanics unit. Three summative quiz questions were supported by checklist scaffolds designed and implemented by the classroom teacher. The results illuminate the difficulties of both assessment and scaffold design and indicate that checklist scaffolds add value to assessments by ensuring they evaluate students' conceptual understanding rather than their memorization or organizational skills.

Introduction

Designing classroom assessments is a task that teachers carry out continuously, but many do not feel confident in this required practice due to insufficient or inadequate training in assessment techniques (Mertler, 2004; 2009) that adequately helps them to create and implement sound and relevant assessments (Zhang & Burry-Stock, 2003). Well-designed assessments allow students to successfully show what they know and can do, and they provide teachers with information about how their students are managing the content and where there might be gaps in their understanding. Consequently, the quality of teacher designed assessments have a significant effect on how students interact with the assessments as well as how they perform (Shepard, 2005). Scaffolding the learning process has been well researched and the benefits to students well documented (van de Pol, Volman, & Beishuizen, 2010). Scaffolding assessment tasks has been shown to provide improved access to assessments for students (Abedi, Hofstetter, & Lord, 2004; McNeill & Krajcik, 2006) yet there is scarce research of teacher designed assessments that include scaffolding. Consequently, little is known about the types of scaffolding that teachers might be using during assessments, or how students interact with these scaffolds and how learning is supported (Kang, Thompson, & Windschitl, 2014).

In this study Jessica, an experienced high school science teacher, implements a variety of self-designed checklist scaffolds into two mid-unit summative assessments for her class of kinesiology students. Kang and colleagues (2014) studied the impact of a variety of scaffolds in teacher-designed written science assessment

tasks and found that checklists helped students to focus on the requirements of the task. A checklist is a list that is used as a reminder, which could be for things not to be forgotten, or points to consider, or things to be collected. For example:

> **Idea Checklist:** These ideas need to be included in your response. When using an idea, be sure to *explain what it means and why you are using it*.
>
> - o Diffusion
> - o Osmosis
> - o Concentration Gradient
> - o Hypertonic
> - o Hypotonic
> - o Molecules
> - o Energy Required
> - o Solute
> - o Semipermeable Membrane
>
> **Answer Checklist**: *Be sure to check and make sure your explanation addresses and answers the following concepts:*
>
> - o Explain how paramecium gets water to survive.
> - o Explain how paramecium gets oxygen to survive.
> - o Explain what would happen to paramecium if salt water is added

Figure 1. Examples of checklists used by Kang et al, 2014.

Jessica designed three checklist scaffolds to support different types of questions in two summative mid-unit quizzes. This study examines how her students respond to and interact with those scaffolds and the effect(s) on their performance. Jessica's students had been studying a biomechanics unit where they had learned about the lever systems of the human arm and leg. This is an eight-week long unit that has traditionally been difficult for her kinesiology students. Jessica has recently received coaching in assessment design and incorporation of scaffolds into the assessment process. As part of that coaching relationship, Jessica and I worked collaboratively to create a graphic organizer combined rubric scaffold for her students to use as they produce scientific explanations for summative assessments submitted in place of traditional laboratory reports. As a result, Jessica's students are familiar with

being scaffolded during assessment. But there are many forms of summative assessment; some are completed by students at home over a period of days or longer and are submitted to the teacher on a preset due date. Others are taken in class under test conditions where students work independently, in silence, and complete the assessment in a fixed time period. The graphic organizer/rubric scaffold was designed for use in the first scenario and is discussed in detail in chapter one. Jessica's scaffolded quiz questions were taken in the second scenario, under test conditions. Both quizzes comprised four short answer questions. A total of 78 student responses were examined.

Theoretical Framework

Scaffolding Theory

The notion of scaffolding in education was first introduced by Wood, Bruner and Ross (1976) and is most often described as providing support from a more knowledgeable other such that a learner is able to proceed further than they could alone (Wells, 1999). Scaffolding as a means to assist students to go further draws upon the sociocultural theory of learning and Vygotsy's (1978) zone of proximal development (ZPD) (Kang et al, 2014; Shepard, 2005). The ZPD is often described as the place beyond a student's current level, a place where students are supported to go further. Vygotsky described the ZPD as:

> The distance between the actual developmental level as determined by
> independent problem solving and the level of potential development as

determined through problem solving under adult guidance or in collaboration

with more capable peers (Vygotsky, 1978, p.86).

During the administration of assessment, rather than a more knowledgeable adult or

peer, that guidance most often comes from prompts or tools that are provided to

students with the intention of supporting them to work in their ZPD throughout the

assessment process (Poehner & Lantolf, 2005). A well-designed scaffold will provide

a high level of support that allows students to be successful at high challenge tasks

(Reiser, 2004; Wilson & Devereux, 2004). A high challenge task is one which is

slightly beyond a student's current achievement, a task that stretches them to apply

their learning and knowledge, to work in their ZPD. Teachers need a good

understanding of their students' abilities to provide high level challenge assessment

tasks that are positioned well within student ZPDs and importantly are supported with

well-designed, appropriate scaffolds that ensure the task is reachable. The scaffold

should provide just enough information for the learner to make progress on his or her

own (McNeill, Lizotte, Krajcik, & Marx, 2006) but during assessment teachers may

be tempted to offer low challenge tasks that do not take students into their ZPD, or for

fear of giving away answers, they may provide low level scaffolds which have little

value. If the scaffold is too simple or it provides too much information, the student

will not be challenged to learn more (Wilson & Devereux, 2004).

Optimal scaffolding. Most definitions of scaffolding in education are framed

as temporary supports that promote learning or achievement with assistance usually

provided by another person (teacher) who intervenes when appropriate to guide and

support the student further than they might go alone (McNeill & Krajcik, 2006;

Siegel, Wissehr, & Halverson, 2008). When used during assessment, the design and

nature of the scaffold is crucial such that their value comes from not only scaffolding

student responses in the assessment but that they also support student learning,

leading to development and eventually outgrowth of the support (Reiser, 2004;

Wilson & Devereux, 2004). Often during science assessments students are somewhat

conditioned to providing the *right* answer such that they focus on achieving that

rather than the scientific principles behind their results or the learning goals of the

assignment. But if students are scaffolded just to get to the right answer and do not

engage in or learn from the experience then the scaffold is not optimal (Reiser, 2004).

Reiser describes scaffolding as a dual aspect process: accomplishing the task, i.e.

getting to the right answer, and learning from the experience such that performance is

increased in future tasks. Consequently, the structure of a scaffolding tool is

important. It can shape how the learner interacts with the tool and how the tool can

impact the learner. For example, when designed to reduce what is required of the

learner, as in the graphic organizer/rubric scaffold mentioned above (see chapter one)

where students have to remember less about the structure of a scientific explanation

and instead can concentrate on the details, the student is then able to focus in a more

conceptual way and the scaffold has provided the potential to learn from the

assignment. When this happens, the tool is scaffolding both the assignment and the

learning: it is an optimal scaffold. Reiser (2004) has highlighted and categorized

successful mechanisms of scaffolding that support learners in accomplishing complex

tasks. He categorizes this reduction of what is required as "offloading" because the overall complexity of the assignment is reduced.

Needs of Learners

When examining the effects of scaffolding tools, it important to keep student needs at the forefront. Teachers give summative assessments that students must be ready for, which means they need to be ready to show what they know and can do. To be successful they must have conceptual knowledge mastery, be able to demonstrate domain-specific skills and strategies, and cognition. But as mentioned above, students often concentrate on what the teacher is looking for rather than cognitive learning processes which can produce weak understanding (Reiser, 2004) and consequently, unsuccessful assessment results. Clearly students are in need of tools that support them in these requirements and not tools to just get to the right answer.

Measurement Issues

Measuring the extent of the effects of scaffolding is hard and consequently less studied than one might hope. In a review of research on scaffolding via teacher-student interactions, van de Pol and colleagues (2010) found many reported scaffolding success stories but concluded that the challenge in scaffolding research is effective measurement. However, Pea (2004) sees a great need for new research within scaffolding theory that goes beyond measuring student performance to investigate how the scaffolding process relates to student learning. He asks for "thickly textured empirical accounts" that not only highlight successful scaffolding designs but more importantly detail "the troubles that arise when learners turn out not

to act in the ways that designers hoped…" (Pea, 2004, p.446). This would include the very difficult task of documenting the comparison of what a student can do while working in their ZPD because of using the scaffold and what they can do without the scaffold. Most importantly, observations should be recorded for every learner, not just a selected representation of learners (Pea, 2004), especially as different learners will perceive and interact in different ways with different scaffolds (Lin, Hsu, Lin, Changlai, Yan, & Lai, 2012).

Context of Study

This study was conducted in a culturally diverse urban high school in Northern California, called S.B. High school (pseudonym). The student body comprises 43% Asian/Filipino, 16% European American, 15% Hispanic and 6% African American, with 20% of students reported as multi-racial. Jessica, a veteran science teacher, implements a variety of self-designed summative assessment checklist scaffolds for her kinesiology students throughout their study of an eight-week long complex and difficult biomechanics unit. Students had been learning about the biomechanics of the human leg and arm. Throughout the unit they completed two quizzes and a series of written scientific explanation assessments. The purpose of this study is to document student interaction with the assessment scaffolds included with the quiz questions and to examine the effectiveness of the scaffolds. The participants are a mixture of 28 junior and senior kinesiology students. Students took Biomechanics quiz #1 half way through the unit. In this 4-question quiz Jessica designed and incorporated checklist scaffolds to support two of the four questions.

Quiz #2, also a 4-question quiz, was taken by the students two weeks later and included a rubric scaffold to support one of the questions.

Assessment Scaffolds: Checklists, Rubrics, and Checklist-like Rubrics

In a study of the impact of teacher designed scaffolds in written science assessment tasks, Kang and colleagues (2014) found that checklists helped students to consider dimensions of a task that they may not otherwise have considered. Inspired by this finding, and after a coaching session where she was introduced to checklists and rubrics as assessment scaffolds, Jessica designed and employed three checklist scaffolds into her two summative mid-unit quizzes. Each quiz was worth a total of 15 points and together make up 33 percent of the total points available for this unit. Biomechanics quiz #1 incorporated a *solving equations checklist* and a *terms checklist,* while biomechanics quiz #2 included what Jessica called a *problem checklist*. The type and purpose of each scaffold is described below. Full details of each scaffold are included in the analysis section.

Table 1
Style and purpose of assessment scaffolds used

Quiz #1	Purpose of scaffold	Quiz#2	Purpose of scaffold
Solving equations checklist	Requirements. Reminder of steps to take and how to present work	Problem checklist	Reminder of C-E-R components and points available
Terms checklist	Offloading. Provides terms to be labeled		

Checklists have many variations and purposes. Operational checklists list sequential steps in a task and offer assistance with organization or ordering, they are

often used during instruction (Rowlands, 2007). For example, a science teacher might include an operational checklist with a lab report writing task that reminds students to begin with a suitable title, followed by a short introduction, then a list of equipment, and so on. While having some characteristics of an operational checklist, Jessica's *solving equation checklist* is more about requirements. It reminds students of the order in which to approach the problem but also offers an insight into what she is looking for in terms of what the students are showing her they can do. The correct answer is only worth half a point, Jessica is looking for students to recall and manipulate an equation while retaining variables and units throughout. It is not uncommon for students to have conceptual understanding but still be unsuccessful in an assessment because of poor organizational skills (Rowlands, 2007). For Jessica, being able to rearrange an equation to isolate a variable when solving and giving an answer that includes units is part of what she terms "being a science student" (Jessica, initial interview). This is the skill that she wants her students to demonstrate and the scaffold reminds them of this. The purpose of her equation checklist is not to support students in calculating the correct answer but rather to scaffold the inclusion of her required components. Including scaffolds like this during the assessment process forces students to think about the criteria of the assignment and so provides a learning opportunity that can aid in developing metacognitive awareness (Rowlands, 2007).

Jessica's *terms checklist* is a word bank checklist. Described as a simple checklist (Kang et al, 2014) this type of scaffold reduces the need to memorize and recall as students label the provided terms onto a diagram provided by Jessica. Most

students in the class scored poorly on this question but by providing the terms checklist Jessica is able to glean that when students did not answer correctly it was because they did not know the answer and not because they could not remember the terms.

Biomechanics quiz #2 made use of a scaffold that Jessica called a *problem checklist*. This one falls somewhat out of the category of requirements checklist and is really a simple rubric scaffold. Rubrics are assessment tools that present a predetermined points breakdown either before or at the time of assessment with the aim of indicating to students what is important during the assessment (Jackson & Larkin, 2002). Often, they are detailed, comprising several levels of scaled breakdown that not only helps students know what is expected but also helps with the learning process. Sheppard (2000) has argued the need for assessments that capture the learning process and which have visible expectations. Rubrics are an excellent tool for this. Kang et al (2014) found that rubrics provide transparency in what is required and the points available for an assignment, or parts of an assignment, which they say makes the activity a learning opportunity.

While they share many characteristics, rubrics and checklists are not the same. They can both identify expected performance but rubrics go further than checklists because they are descriptive and scaled to identify level of performance. Jessica's *problem checklist* is not a full rubric, there are no degrees of achievement, and as such it is acting more as a reminder of what to include rather than supporting depth of response. It is also important to note that as mentioned above, Jessica's students have

been writing C-E-R style scientific explanations throughout this unit in place of traditional laboratory reports, supported by a detailed rubric/graphic organizer scaffold. When Jessica provides this simplified *problem checklist* scaffold she is essentially fading the rubric scaffold that students have been used to using. Fading is when a scaffold is gradually simplified or removed, and prior research has suggested that fading encourages student independence (McNeill et al, 2006).

Research Design

To examine the effectiveness of the assessment scaffolds designed and employed by Jessica and the corresponding student interaction, I use a descriptive and exploratory research design. In descriptive studies the functions and characteristics of the study focus are observed and described in an effort to explore and explain a situation or phenomenon (Creswell, 2014). The exploratory influence comes from a more flexible analysis than is usual with descriptive studies and also the objective of capturing thoughts and ideas highlighted during analysis relating to Jessica's self-designed scaffolds. While framed strongly within the sociocultural theory of learning, the purpose of my analysis is to describe the behavior of all students regarding scaffold interaction to add to the limited literature in this area rather than report just the accounts where the scaffold is successful. I collected and analyzed student responses to each scaffolded quiz question together with an analysis of their interaction with the scaffolds and the scores they earned. I also collected, transcribed, and analyzed three audio interviews with Jessica throughout the eight-week long unit. Several rounds of analysis were carried out guided by patterns that emerged from the

data. Much of the analysis is qualitative, with a focus on how students behave in terms of their response to and interaction with the provided scaffolds. However, student achievement scores are also considered, providing a quantitative view of the data to ensure a complete and valid analysis (Creswell & Plano Clark, 2007).

When a scaffold is provided it is reasonable to expect that students will make use of it. The intended purpose is that the scaffold helps students to show what they know without giving away anything that might provide false information to the teacher about what the student knows and can do. The scaffold will also support students to move away from just getting to the right answer and instead will provide a learning component. My focus of analysis regarding student interaction with the scaffolds follows Pea's (2004) advice in that it examines the responses from all students in the class rather than just the students who interacted with the scaffold and scored well. Analysis of student responses illuminated multiple scenarios of behavior. The responses were grouped by student behavior regarding the scaffold and their score for each question. The multiple response scenarios are shown in figure 2 below:
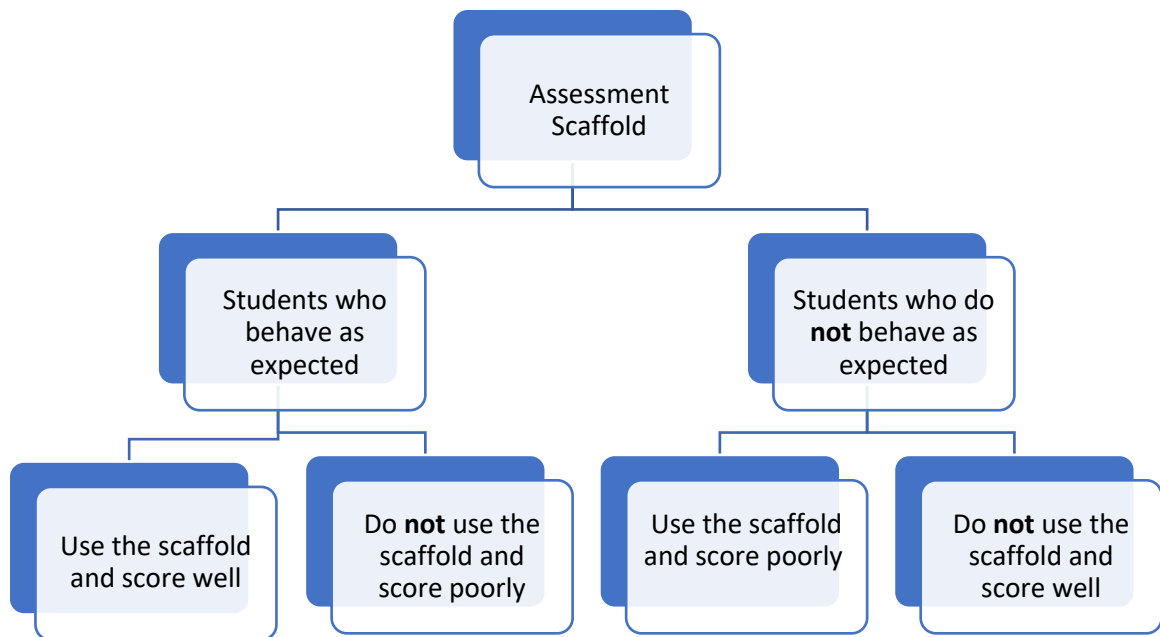
Figure 2. Scenarios of student responses to scaffold

Results and Analysis

Each scaffolded question was analyzed separately and is reported and discussed below. Student responses were grouped by how they interacted with the scaffold and the score they achieved. The analysis addresses one main guiding question:

What are the different scenarios of student interaction and achievement with teacher designed assessment scaffolds that are provided to support summative assessments in a science unit?

Solving Equations Checklist

This checklist was the first scaffold that students were exposed to during a summative assessment under test conditions. Students worked independently in

silence and turned in their completed work to the teacher at the end of the quiz

without consulting with other students or the teacher. The question asked:

> Solve the following problem: A teeter-totter was arranged so that the board is
> 5 meters long on one side and 6 meters long on the other. A 30-kg child sits
> at the end of the 5 meter side and exactly balances another child sitting at the
> end of the 6 meter side. How big (mass) is the child on the 6 meter side?
> Remember to show all your work as a science student would when using
> equations and solving problems. 3 pts.

When teaching her students how to solve math problems in science Jessica had spent

time talking to them about *being science students* and she often verbally reminded

them to "be a science student during lessons and before giving quizzes or tests

involving calculations" (Jessica, initial interview). For Jessica, being a science

student means "…doing science; and also showing clearly how they solve equations,

in a set way; they isolate the variable and not forgetting to include units…" (Jessica,

initial interview). She decided to include the reminder into the question as an

additional scaffold for them as well as the *solving equations checklist*, below:

Solving Equation Checklist	
Equation rearranged - Variable solving for is on the left of equation (1.5 pts)	
Correct units (1 pt)	
Boxed Correct Answer (.5 pt)	

Figure 3. Solving equation checklist

Because Jessica has included a score breakdown into this checklist, it doubles

somewhat as a grading rubric because it indicates to students what is important during

the assessment (Jackson & Larkin, 2002). However, Jessica's rubric here is very

simple, there is no scale and it serves to focus students on the fact that the correct

answer is only worth half a point such that she is more interested in the process than the product.

Solving Equations Checklist: Results

All 27 students who took the quiz attempted this quiz question. The intent of the scaffold is to support students in setting out and showing their work in the way the teacher requires while performing calculations. The expectation of the scaffold is that students will use the checklist as a reminder of the order in which to proceed and how to present their work such that they are able to complete the calculations without losing *housekeeping* points. There are two scenarios where students might behave as expected: students mark each box of the checklist, carry out the required calculations and set out their work corresponding to the requirement of each box and score well, or they will not make use of the checklist and score poorly. Likewise, there are similar scenarios where students do not behave as expected: those who check the boxes but score poorly, or those who do not check the boxes but score well. In addition, some unexpected scenarios were observed: students who check the boxes but do not perform the action they checked, and students who neither check the boxes or perform the action, both of whom earn a reduced score because of not performing a required action, are also not behaving as expected.

Three points were available for this question. To score well was categorized as earning full marks (3 points). A reduced score was categorized as greater than 2 points but fewer than 3 points; and scoring poorly was categorized as earning 2 or fewer points.

Table 2

Solving equations checklist: student behavior and score groupings

Equations checklist: Students who behave as expected	
Boxes checked and score well (3 points)	Boxes not checked and score poorly (≤ 2 points)
7 students	7 students

Equations checklist: Students who do not behave as expected			
Boxes not checked and score well (3 points)	Boxes checked but function not executed, reduced score (> 2 points)	Boxes not checked and function not executed, reduced score (>2 points)	Boxes checked and score poorly (≤ 2 points)
7 students	2 students	4 students	0 students

Solving Equations Checklist: Discussion

14 students earned full marks for this assignment, half of the class. One can hypothesize that of these, the 7 who appeared to ignore the scaffold did so because they were skilled in the requirements and were able to successfully complete the assignment without the scaffold. During instruction scaffolds are faded as students develop their own understanding (McNeill et al, 2006), students who ignored the scaffold may have already developed a deeper understanding and therefore had outgrown the need for the scaffold. Likewise, because it was expected that by using the checklist students would be scaffolded to score well, it could also be claimed that the 7 who did check the boxes and scored well were helped to a successful full score because the scaffold was available, helpful, and they made use of it. These students were helped to go further than they might have without support, so for them the scaffold could be deemed to have been successful. Much of the existing scaffolding research presents analysis in this way, with a focus on the success of the scaffold, and in particular those students who performed better when scaffolded (Lin et al, 2012;

Pea, 2004; Robinson, 1997). But there are two potential problems with focusing the argument in this way, on just the successful students. The first, and possibly most important observation is that we don't know for certain if students made use of the scaffold or not just because they checked or did not check the boxes. It is critical to recognize that just because students did not mark the checklist does not automatically mean that they did not make use of the scaffold and vice versa. Students who did not check the boxes may have still used the scaffold, and students who did check them, might, for example, have checked them after they had completed the question, as a double-checking gesture. This is information that we cannot know without some kind of post-test survey with each and every student after each and every assessment. Jessica's students were not directly interviewed about their interactions with the scaffold but Jessica did comment in a reflection meeting that her students were familiar with checklist use and during a class discussion most students had said they felt it was useful to have the scaffold. Secondly, as Pea (2004) reminds us, the behavior from the other students in the class, the ones who did not score full marks, tells much more regarding the success (or not) of the scaffold. Since Jessica had previously discussed use of checklists with her students, for my analysis I have treated students who marked the checklist as making use of the scaffold and those who did not mark the checklist as not making use of the scaffold.

For students behaving as expected. The 7 students (25% of the class) who did not physically make use of the checklist and scored poorly all earned a score of 2 points or below and one student earned a score of zero. Students who scored 2 points

had either missed a key equation or used incorrect values in their calculations. If Jessica is certain that this group of students were ready for a summative assessment it seems they might benefit from more detailed scaffolding than the equation checklist provided, such as help recalling the equations. The gap between what they can do alone and what they can do when Jessica scaffolds them with her equations checklist is too large for it to support them into their individual ZPDs and help them be successful. If students are not engaging with the scaffold then it is unlikely to be providing support or the opportunity to learn. What seems like a simple reminder checklist has highlighted the extent of the different ZPDs in the class and Jessica is informed that a quarter of her class are unable to perform the required calculations even when supported by a simple scaffold.

For students not behaving as expected. There were 6 students (almost 25%) who did not rearrange the equations, as requested, to isolate the variable being solved for. Instead, they carried out a cross-multiplication method. Of these students, 2 marked the checklist indicating that they had rearranged the equation, but had not, and 4 did not mark the checklist. The cross-multiplication method of solving equations is taught and used by the math department at S.B. High and it appears these students are reluctant to change to Jessica's method of solving equations in science. They were mostly successful in recalling the correct equations and selecting the correct values from the text to solve the problem, but because they did not show work as the question asked, each of them earned only 2.25 of the available 3 points. This is the group of students that Jessica had in mind when she designed the checklist, i.e.

those who are able to carry out the calculation but are not showing their work clearly. During her debrief interview she explained that it matters to her how students show work in science calculations, particularly because she feels that this is helping them to prepare for college science. As we discussed the quiz scores she said "we talked about this in class, we learned how to isolate variables and we asked how is math and science different? So, I wrote in here *like a science student would"* (Jessica, debrief interview). While one quarter of the class still did not rearrange the equation, it is clear from our discussion that this is fewer than usual: "And it helped, they did good. They showed their work they have their units, they boxed their answer. Only a few kids didn't do it [isolate variable]". However, one quarter of her students are either unable or uncomfortable enough with solving equations *like a science student* that they continue to use the math department method, ignoring the scaffold, even though they were made aware that they would, and did, lose partial credit.

To know for certain if the equations checklist was an optimal scaffold for her students Jessica would need to compare their performance in this quiz not only with past quizzes but also with performance in a similar future assessment. Without this information the measure of the success of the scaffold can only be taken from how students behaved regarding the scaffold. We do know that half of the students behaved as expected and half did not and can conclude that

the solving equations scaffold was potentially optimal for 7 of 27 students (25%), was unnecessary for 7 of 27 students (25%), was ignored by 6 of 27 students, and was beyond 7 of 27 students.

Terms Checklist

In the same quiz Jessica scaffolded a second question with a checklist. This was another simple checklist that provided a list of terms students had been using in class. The question asked students to:

Mark and label the common biomechanics terms on the human model below. 5pts

Term Checklist	
COG	
RA	
EA	
EF	
RF	

Figure 4. Terms quiz question and scaffold

As in the first question, not all students made use of the checklist by physically checking the terms but all students in the class did attempt to label all five terms which indicates that the checklist may have been helpful by providing the terms. Jessica was not testing her students on their memorization of the terms, she was asking them to indicate their location on a diagram. By providing the terms, her students were given full opportunity to answer this question.

Terms Checklist: Results

All 27 students attempted this question which asked students to label 5 common biomechanics terms onto a provided diagram of the human leg. Jessica provided the terms in a checkbox to reduce the cognitive load on her students. She was not testing their recall of the term names, instead the question examined if her

students were able to identify where on the human leg these key labels would be. The expectation is that students will use the checklist as a reminder of the 5 terms such that they will attempt to label all 5. The scaffold should help Jessica to know that if a label was missed it was because the student did not know where to label it and not because they had forgotten the term. This scaffold, unlike the equations checklist, has no direct effect on how students score. Again, there are two scenarios where students might behave as expected and four where they do not. Students who check the boxes and attempt to label all terms, and students who do not check the boxes and do not label all terms are behaving as expected in terms of the scaffold. Students who check the boxes but do not label all terms, and students who attempt to label all the terms but do not mark the checklist are not behaving as expected. Five points were available for this question, one point for every correctly labeled term.

Table 3
Terms checklist: student behavior and score groupings

Terms checklist: Students who behave as expected			
Boxes checked and all labels attempted		Boxes not checked and labels not attempted	
21 students		0 students	

Terms checklist: Students who do not behave as expected			
Boxes not checked and all labels attempted	Boxes checked but all labels not attempted	Boxes not checked and some labels attempted	Some labels attempted but do not correspond to boxes checked.
6 students	0 students	0 students	0 students

Terms Checklist: Discussion

Close to three quarters of the students checked the terms and attempted to label them onto the diagram. Of the 6 who did not mark the checklist, 2 earned the

full 5 points that were available. As mentioned above, this scaffold has no direct

relationship with student achievement, but since they did not visibly interact with the

scaffold it can be hypothesized that, as with the solving equations checklist, students

who did not make use of the checklist but scored well did not need the scaffold.

Overall though, the class did poorly on this question, with the mean score being 2.6

points. There were no obvious common errors other than many students mislabeled

the center of gravity (COG) term. Only one other student scored full marks for this

question and only 6 students correctly labeled 4 or more terms. Jessica was

disappointed with the results and during her debrief interview she began to question

her skill in creating the assessment:

> I was really surprised. I thought they were going to do ok. I don't know where
> to go. They clearly as a class did not do well on this particular question, and
> we've spent more time on this than anything else. Did they remember all the
> labels? Yes. But they weren't often in the correct place, especially COG. I
> need to find a different model. [In class] we stood up, we modeled it, COG.
> Heel, toe, connecting teeter totter and the functional model and then we'd
> have a discussion, so I don't know why they can't make that connection.
> Maybe this picture wasn't clear enough.

Jessica was disappointed that during class discussions her students had shown that

they were able to identify key biomechanics terms, but the quiz shows that many of

them are not proficient in this skill. She acknowledges that not all students are

participating in those discussions: "I have 30 kids in the class so it's easy to hide" and

decides that she will "change this diagram so that it's not a walking person but both

legs are straight. They don't like the biomechanical models. I think my next

assessment should have this again [picture of leg] and a question that's about

identifying the three levers and have them give an explanation" (Jessica, debrief interview).

The terms checklist was intended as an offloading scaffold, where students are able to focus in a more conceptual way because the scaffold has reduced the complexity of the assignment (Reiser, 2004). Once again, without surveying students we do not know for certain if they found the scaffold useful or if they did not need the scaffold, but by providing the terms that she wanted them to use Jessica gave her students more opportunity to be successful. Most students were unable to correctly label all 5 terms and Jessica is provided with clearer information about why that happened than she would have if the checklist were not provided. Simple word checklists reduce cognitive load and allow students to use information. For example, a vocabulary list reduces the anxiety to recall the correct terminology and instead students can concentrate on incorporating that vocabulary into their response (Kang et al, 2014). This was not quite what happened with this quiz question because instead of using the vocabulary to produce a written answer, such as a scientific explanation, the checklist was used to label a diagram. The terms scaffold was potentially optimal for 21 of 27 students (78%); was potentially unnecessary for 2 of 27 students ($\approx$7%); was ignored by 6 of 27 students; and was beyond 0 of 27 students.

Quiz CER Checklist: Results

In quiz #2 Jessica incorporated a *problem checklist* for a question that required a scientific explanation. During instruction, students had learned the characteristics of first, second, and third-class levers that allow the class of lever to be identified. This

142

question required students to identify the lever system and write their answer as a C-E-R style scientific explanation.

> A person is lying on a yoga mat prone (on their stomach) and performs knee flexion. The primary muscle involved in this movement is the hamstring muscle group. This muscle group originates from the ischial tuberosity and crosses the knee joint and attaches on the tibial condyle and the head of the femur approximately 2.5cm from the knee joint. The COG for this lever system is about 13cm from the knee joint. State your claim in identifying the class of lever system for this movement. Also include evidence and scientific reasoning to support your claim. Finally, include a labeled drawing to support your answer. 4.5 pts.

Problem Checklist	
Claim stated (1 pt)	
Evidence included (1 pt)	
Scientific reasoning given (1pt)	
Labeled drawing included (1.5pts)	

Figure 5. C-E-R problem checklist scaffold

24 students took this quiz. All 24 attempted to answer the question that asked them to write a scientific explanation regarding the class of lever system in knee flexion. The scaffold is a simple reminder checklist of what is required for a scientific explanation, i.e. a claim, supported by evidence and scientific reasoning. Students are not provided with any reminders around what might constitute as evidence in this scenario, nor are they scaffolded in the scientific principles that will provide reasoning. Instead they are reminded that they need to provide a three-part answer with a labeled diagram to earn the full 4.5 available points. If the scaffold serves its purpose, students will mark the checklist and provide a C-E-R style scientific explanation. Their claim will be supported by measurement evidence and their reasoning statement will incorporate at least one scientific principle relating to the

class of lever. For students who behave as expected Jessica should expect to see either the checklist marked and a C-E-R explanation supported by a labeled diagram, or the checklist remains unmarked and students are unable to provide all parts of the explanation. 20 students behaved as expected regarding the scaffold, 5 marked the checklist and scored well, while 15 did not mark the checklist and struggled to produce a scientific explanation. Scores for these 20 students ranged from zero to 3 points, with most earning between 1 and 2 points. There were 4 students who did not behave as expected; 2 did not mark the checklist but produced a good explanation and supporting diagram, and 2 scored poorly but marked the checklist.

Table 4
C-E-R problem scaffold: Student behavior and score groupings

CER quiz checklist: Students who behave as expected	
Boxes checked with good CER explanation	Boxes not checked and missing parts of explanation
5 students	15 students

CER quiz checklist: Students who do not behave as expected	
Boxes not checked with good CER explanation	Boxes checked and missing parts of explanation
2 students	2 students

CER Checklist: Discussion

The scaffold provided for this question serves only as a basic reminder to make a claim and to include supporting evidence and reasoning when writing a scientific explanation. It offers no support or reminders to students about what might constitute acceptable evidence or scientific principles for reasoning – information that students have been used to finding available in the graphic organizer/rubric scaffold

144

that they have used for other summative assessments. Jessica needs to be careful that this does not send the wrong message to her students. Is she assuming that they can easily answer the question and easily recall all details of a scientific explanation? The scaffold implicitly assumes that students are able to identify the class of lever from the information provided, i.e. they need to be able to identify the resistance arm (RA) and effort arm (EA) from the text and use the measurements to identify that RA > EA and also know that this scenario indicates a 3rd class lever. There are no scaffolds provided to help with this. Also implicit is that students will produce a correct claim, use correct evidence, and provide correct scientific principles. There is no room in the rubric for incorrect or partially correct answers, although when Jessica scores she does give partial credit. The scaffold has the purpose of helping students take this information and present it as a scientific explanation. This was a very basic scaffold and it did not support students in the way that Jessica had intended. The scaffold gives the appearance that it is supporting the explanation such that the grade on offer is about the explanation. But students make many mistakes in identifying the lever class and also with their scientific principles and demonstrate some confusion regarding how to use measurements. The rubric says 'claim stated (1pt)' but students who provided an incorrect claim, scored 0. For example, student 302 provided a claim but it was incorrect:

> The lever class for this movement is a second-class lever. This is clear because the effort arm is longer than the resistance arm. The load is also in between the force and the pivot. It's easy to conclude this because of the data given in which the quadricep muscle goes all the way through the femur from the hip to the tibial tuberosity and the RA is from the COG to the pivot which is 13 inches.

In providing the simple C-E-R reminder scaffold with this quiz question Jessica is reducing the detail of scaffold that students have been used to when writing scientific explanations such that she may be introducing a form of scaffold fading before students are ready. Fading scaffolds is an important part of the scaffolding process, especially as the ultimate goal is that students are supported to a place where they are able to complete the required task independently. However, if a scaffold is removed or faded too quickly, the task may now be too far beyond a student's ZPD so that the student finds the task too challenging and the reduced scaffold is ineffective (McNeill et al, 2006). By observing how students behave regarding this scaffold Jessica is not only provided with information about her students' understanding but also of her own assessment and scaffold design. While simple scaffolding practices have been found to be sufficient to increase the accessibility of classroom assessments, the quality of the scaffold has been found to be more effective than the quantity of scaffolds used in supporting students (Kang et al, 2014).

Concluding Remarks

The purpose of this study was to document student interaction with teacher designed scaffolds that have been provided to support summative assessment questions in science, and to examine the effectiveness of those scaffolds. The overall analysis was conducted in a way that accommodates recommendations from previous scaffolding research in that it focuses on the behavior and responses of all participating students rather than a select few or on highlighting the success of the scaffold (Pea, 2004). Attention was given to the variety of interactions by students

with the scaffolds and the information this provides to the teacher, and those observations have been discussed and reported in a rich account. The very difficult task of documenting the comparison of what a student can do while working in their ZPD because of using a scaffold and what they can do without scaffolding, is recognized as important but was beyond the scope of this study, which reaffirms the difficulties and issues surrounding measuring the effects of scaffolding. The dynamic nature of scaffolding adds to the complexity of measuring the effects, and so most studies are small exploratory or descriptive studies like this one, or case studies (van de Pol, et al, 2010). Several interesting scenarios were observed when students did not interact with the scaffold as intended, and as a result of focusing on these interactions the teacher was provided with quality information about her students' understanding as well as the effectiveness of her scaffolds. An important outcome of assessment scaffolding is student learning. Jessica's solving equations scaffold was designed to provide a learning opportunity yet she still observed several students who did not respond as she intended. Kang et al (2014) claim that "…with effective use of scaffolding, teachers create better opportunities for students to demonstrate disciplinary proficiency" (2014, p.697) but teachers need specific and detailed training in the creation and use of assessment scaffolds for this to happen.

Two important issues for consideration can be drawn from this study. First, many teachers claim they are not well trained in assessment practices (Campbell, 2013) and so we need better training for preservice and in-service teachers in assessment design, particularly assessment design that incorporates scaffolding, and

also in evaluating and acting upon student responses and interactions with the scaffolds. For example, with regard to solving equations Jessica clearly has four groups of student responses. Analyzing how students responded to this scaffold has highlighted the nuances of data evaluation and assumptions, which when examined with a focus on those who do not behave as expected has shown not only how difficult it is to research the benefits of scaffolding in assessment, but also the difficulties of designing of a good scaffold. Secondly, not all scaffolds are good scaffolds. The quality of a scaffold used to support the assessment process is important, and more valuable than the number of scaffolds provided (Kang et al, 2014). An effective scaffold is an optimal scaffold. If students are scaffolded to just get to the right answer but do not engage in or learn from the experience then the scaffold is not optimal (Reiser, 2004). By engaging with the scaffold, a student should perform better than they would have without it and their performance in future tasks of the same nature should be improved. This is described by Reiser (2004) as the dual aspect process of scaffolding. But designing a scaffold that is optimal for all students is not as easy as it sounds. Jessica's terms scaffold, while full of offloading potential, was only effective for students who knew the correct label location and so turned out to be less effective than she had hoped. One could argue this was still an optimal scaffold, the purpose was to ensure that students did not forget the terms to label and the results show that no students forgot a label. Any lost points were due of lack of content understanding. This was a great formative tool for Jessica, and students said they felt supported with the scaffolds in place (Jessica, debrief interview). So, this

might actually have been more of an optimal scaffold than it appears. Also, Jessica's C-E-R problem scaffold, while appearing somewhat familiar to students actually lacked sufficient detail to support most of them in the way she had intended. Overall, while not all of the scaffolds worked for every student, Jessica was pleased with the results she obtained from including checklists with her assessments: "I think overall this quiz was better than before. I think the middle kid did a little better and I don't have so many outliers" (Jessica, final interview).

When well designed, checklists scaffold students' metacognition. Operational and requirement checklists provide support for students to help them complete complex tasks without missing key parts, which boosts confidence, aids in learning, and maximizes student performance (Rowlands, 2007). Providing scaffolds with assessments can help teachers evaluate their students' conceptual understanding rather than their memorization and/or organizational skills (Reiser, 2004). This study found that incorporating checklists into the assessment process was helpful for both the students and the teacher, which supports findings from previous work on using checklists as an assessment scaffold (Kang et al, 2014). Further, how different learners perceive and interact with scaffolds depends on the design and form of the scaffold provided so by reporting on the interactions of all students rather than just those for whom the scaffold was successful, the opportunity to illuminate important information for future assessment scaffold design is provided (Lin et al, 2012; Pea, 2004).

References

Abedi, J., Hofstetter, C. H., & Lord, C. (2004). Assessment accommodations for English language learners: Implications for policy-based empirical research. *Review of Educational Research, 74*(1), 1-28.

Campbell. C. (2013). Research on teacher competency in classroom assessment. In J. H. McMillan (Ed.), *Sage Handbook of Research on Classroom Assessment* (pp.71-85). Thousand Oaks, CA: Sage Publications, Inc.

Creswell, J. W., & Plano Clark, V. L. Thousand Oaks, CA: Sage Publications, Inc. (2007). *Designing and conducting mixed methods research*. Thousand Oaks, CA: Sage.

Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches.* Thousand Oaks, CA: Sage.

Jackson, C. W., & Larkin, M. J. (2002). Teaching students to use grading rubrics. *Teaching Exceptional Children*, *35*(1), 40-45. Retrieved from https://doi.org/10.1177/004005990203500106.

Kang, H., Thompson, J., & Windschitl, M. (2014). Creating opportunities for students to show what they know: The role of scaffolding in assessment tasks. *Science Education*. doi:10.1002/sce.21123.

Lin, T., Hsu, Y., Lin, S., Changlai, M., Yang, K. & Lai, T. (2012). A review of empirical evidence on scaffolding for science education. *International Journal of Science and Mathematics Education, 10*(2), 437-455. doi:10.1007/s10763-011-9322-z.

McNeill, K. L., & Krajcik, J. (2006). Supporting students' construction of scientific explanation through generic versus context specific written scaffolds. American Educational Research Association Annual Meeting, San Francisco, CA. Retrieved from: http://www.umich.edu/~hiceweb/iqwst/Papers/McNeill_Krajcik_AERA2006.pdf.

McNeill, K. L., Lizotte, D. J., Krajcik, J., & Marx, R. W. (2006). Supporting students' construction of scientific explanations by fading scaffolds in instructional materials. *The Journal of the Learning Sciences,15*(2), 153-191.

Mertler, C. A. (2004). Secondary teachers' assessment literacy: Does classroom experience make a difference? *American Secondary Education, 33*(1), 49-64.

Mertler, C. A. (2009). Teachers' assessment knowledge and their perceptions of the impact of classroom assessment professional development. *Improving Schools, 12(*2), 101-113.

Pea, R. D. (2004). The social and technological dimensions of scaffolding and related theoretical concepts for learning, education, and human activity. *Journal of the Learning Sciences, 13*(3), 423-451. doi:10.1207/s15327809jls1303_6.

Poehner, M. E., & Lantolf, J. P. (2005). Dynamic assessment in the language classroom. *Language Teaching Research, 9*(3), 233-265. doi:10.1191/1362168805Ir166oa.

Reiser, B. J. (2004). Scaffolding complex learning: The mechanisms of structuring and problematizing student work, *Journal of the Learning Sciences*, *13*(3), 273-304. doi:10.1207/s15327809jls1303_2.

Robinson, D. H. (1997). Graphic organizers as aids to text learning. *Reading Research and Instruction, 37*(2), 85-105. doi:10.1080/19388079809558257.

Rowlands, K. D. (2007). Check it out! Using checklists to support student learning. Retrieved from: https://www.siprep.org/uploaded/ProfessionalDevelopment/ Readings/Checklists.pdf.

Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational Researcher, 29*(7), 4-14.

Shepard, L. A. (2005). Linking formative assessment to scaffolding. *Educational Leadership*, *63*(3), 67-70.

Siegel, M. A., Wissehr, C., & Halverson, K. (2008). Sounds like success: A framework for equitable assessment. *The Science Teacher*, *75*(3), 43–46. Retrieved from http://people.uncw.edu/kubaskod/NC_Teach/ Class_5_Assessment/Sounds_Like_Success_Equitable_Assessment.pdf.

Van de Pol, J., Volman, M., & Beishuizen, J. (2010). Scaffolding in teacher—Student interaction. *Educational Psychology Review*, *22*(3), 271-296. Downloaded from http://www.jstor.org/stable/23364144.

Vygotsky, L. S. (1978). *Mind in Society: The development of higher psychological processes.* London: Cambridge University Press.

Wells, G. (1999). *Dialogic Inquiry: Towards a sociocultural practice and theory of education.* New York: Cambridge University Press.

Wilson, K., & Devereux, L. (2004). Scaffolding theory: High challenge, high support in academic language and learning (ALL) contexts. *Journal of Academic Language & Learning*, *8*(3). ISSN 1835-5196. Downloaded from www.aall.org.au.

Wood, D., Bruner, J. S., & Ross, G. (1976) The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry*, *17*, 89-100.

Zhang, Z., & Burry-Stock, J. A. (2003). Classroom assessment practices and teachers' self-perceived assessment skills. *Applied Measurement in Education, 16*(4), 323-342. doi:10.1207/S15324818AME1604_4.

CONCLUSION

In this final section I revisit the research objectives of each chapter to provide a summary of the findings and resulting conclusions. I then discuss how this research might contribute to existing knowledge followed by some recommendations. Finally, I provide a short self-reflection section where I discuss how I have grown as a researcher and how this dissertation has affected me as an academic.

Research Objectives

In chapter one, "Scaffolding Students' Scientific Explanations", the students of three different life science classes were scaffolded with the SET4CER tool, a graphic organizer/rubric developed during coaching sessions with the classroom teachers. The aim was to see if SET4CER could support students to produce well written, persuasive C-E-R style scientific explanations that demonstrated their understanding of the scientific principles they had been learning about. Overall, all three classes showed a good degree of improvement when scaffolded. Vanessa had 72% of her students producing explanations with better use of evidence and reasoning, i.e. a better demonstration of understanding the scientific principles, when scaffolded. The number of students with low scores also decreased, from 4 students to only 1 student. Similarly, Jessica found an upwards shift in students who had earned lower scores when un-scaffolded. When SET4CER was provided 100% of her students attempted to incorporate scientific principles into their explanations and the number of students scoring well in this category (reasoning) raised from 35% to 64%. While Roxanne herself showed less commitment to the notion of scaffolding, the

structure of her students' explanations was improved with SET4CER as they were scaffolded to make attempts to score in each category.

The objective of Chapter two "Exploring the Role of Coaching in Professional Development: A Study of Three Science Teachers and How They Respond to Coaching in Techniques to Scaffold Students' Written Scientific Explanations" was to provide rich details of the response to and interaction with the coaching process from three science teachers, Henry, Vanessa, and Roxanne who were receiving coaching in creating and implementing scaffolds to support their students in summative assessments. The purpose was to evaluate the potential of coaching as a method of professional development delivery. The range of responsiveness and ultimately adoption of assessment scaffolding techniques revealed that while coaching did help with professional development delivery, the coaching process is complex. The narrative case studies in this chapter highlighted the complexities of teacher attitudes, knowledge, and beliefs as critical aspects that must be considered for successful coaching and professional development adoption.

Chapter three, "Teacher Designed Summative Assessment Scaffolds in Science: Student Interactions and Achievements", concentrated on just one teacher, Jessica, as she designed and implemented checklist scaffolds into two Kinesiology summative assessments taken under test conditions. The purpose was to describe the variety of student interaction and scores when these scaffolds were provided, being particularly careful to include responses from all students - those for whom the scaffold was successful and also students who did not behave as expected regarding

their interactions with the scaffolds. While Jessica was pleased with the overall improvements in how her students performed when scaffolded, the variety of interaction discovered serves to illuminate the difficulties that teachers face in designing and analyzing classroom assessments and scaffolds that are appropriate for all students.

As a result of the empirical research presented in each of the three chapters, and the literature guiding each one, several things of note stand out in relation to the objectives outlined above. First, a common theme observed across all three chapters is that scaffolding is beneficial but complex when applied to assessments. The scaffolding process is more commonly used during instruction where in addition to providing scaffolding tools it can be better individualized for students by talking with the teacher or their peers. When administering assessments, effective scaffolding is more difficult because students usually work in isolation. However, these three studies have shown that scaffolding in the assessment process is possible and also very effective, and together have highlighted successes and also challenges facing classroom teachers as they design and implement scaffolds for classroom assessments. Second, the instructional coaching process, when framed within a sociocultural perspective, is an effective driver of teacher professional growth. Teacher and coach engage in cycles of collaborative planning, action, and reflection, where the coach encourages the teacher into their zone of proximal development (ZPD). Instructional coaching has the potential to bring about changes in teacher practice (Knight, 2007; Knight & Cornett, 2008), but the coaching process has a

degree of intricacy that I had not fully considered which emerged through my coaching relationship with Roxanne and led to the conclusion that for coaching to be fully effective as a transport for professional development the thorny issue of teacher beliefs needs full attention. Third, the importance for students in understanding how to create or evaluate a scientific explanation is huge. For example, fifteen of the High school life and physical science Next Generation Science Standards (NGSS) performance expectations require students to apply scientific principles and evidence to produce an explanation; or to construct and revise an explanation; or to evaluate claim, evidence, and reasoning; or to validate the reliability of a claim. This is a big change from the requirements of the outgoing Standards which were more about content than application, so it is important to keep in mind that the move to NGSS is still relatively new and many teachers are still learning how to translate the new Standards into their curricula, particularly in terms of assessments. Chapters one and two demonstrated the benefits of scaffolding for students as they produce written scientific explanations for summative assessments, which supports earlier work in this area (Gotwals & Butler Songer, 2013; Kang, Thompson, & Windschitl, 2014; McNeill & Krajcik, 2006; Sandoval, 2003). The SET4CER scaffold serves as a good example of a tool that can support students of all levels of achievement in their written explanations, but the scaffolding process is more nuanced than just providing a good tool. Teachers need detailed understanding of their students' individual ZPDs and of the intent of the scaffolds they select. This was illuminated in chapter 3 which showed how simplifying a tool like SET4CER too early can inhibit the purpose of the

scaffold for some students. Lastly, the studies that form this dissertation are important because they focus on what actually happens in the classroom. Researchers have consistently found deficits in the knowledge and skills of teachers in their understanding and practices of classroom assessment, i.e. they are not effectively prepared to use assessments or to evaluate the results of assessments (Campbell, 2013).

Contributions to Knowledge

In a review of scaffolding literature, Lin, Hsu, Lin, Changlai, Yang and Kai (2012) found that only 2.33% of the studies they reviewed were concerned with scaffolding in assessments. Most studies investigate the effects of scaffolding related to learning contexts such as curriculum, teacher education, and learning conception, which are worthy and much needed, but the lack of attention to assessment is disappointing. The studies forming this dissertation have focused on scaffolding in the summative assessment process and contribute to the body of knowledge concerning the design and application of scaffolding in assessment and interpreting the associated outcomes. The findings from this research have the potential to help teachers, researchers, teacher educators, and professional development providers to better understand the challenges of scaffold design and the interpretation of student engagement when viewed alongside existing research in this area. Collectively, these studies highlight the importance of scaffolding the assessment process as an essential component of science classroom pedagogy.

The analysis included in chapters one and three of how students' achievement changed when they were scaffolded during the assessment process adds value to our understanding of how elements of classroom assessment contribute to student outcomes and what matters most in assessment with regard to improving student learning (Randel & Clark, 2013), as well as some paving in the way to addressing the question "what impact does coaching have on student achievement?" (Cornett & Knight, 2009). Chapter two provides a contribution to the growing literature on coaching in education in several ways. It "bridges the gap between training and application of new learning in classrooms" (Killion, 2012, p.273), and confirms the instructional coaching model as an effective mode of transporting and retaining professional development into teachers' classrooms. Also, the detail provided in chapter two goes beyond the self-reported anecdotes that make up much of the limited existing literature and adds affirmation to claims of the benefits of coaching on teacher practice (Cornett & Knight, 2009).

Recommendations

Scaffold for success. There are few studies that examine the use of scaffolding in the assessment process (Lin et al, 2012) and more research is needed in this area, in particular studies that are contextualized within teachers own classrooms, working with their own curriculum and their own students. Scaffolding challenges students to perform in their zones of proximal development; well-designed scaffolded assessments provide better opportunity for students to show what they know and can

do and should be "necessary, not optional, when trying to support students in meeting twenty-first century standards" (Kang et al, 2014, p.702).

Situate professional development in the classroom. There remains a disconnect between research and practice which could be addressed by making better use of involving classroom teachers in implementing reforms (Lin et al, 2012). For example, the changes to assessment practices brought about by adoption of NGSS creates professional development opportunities that need to be better managed to avoid the problems highlighted with more traditional methods of in-service teacher training, such as decontextualized material presented during mandatory attendance at a program organized by the district or school (Feiman-Nemser, 2001). This means working directly with classroom teachers in ways that help them implement the professional development by tailoring to suit the needs of the teacher, their content area, and the grade level of their students (Zhang & Burry-Stock, 2003). This dissertation research has found coaching to be a worthy alternative to more traditional methods that needs further exploration. For example, as well as studies that report on the overall results of coaching, research would be beneficial that investigates how many coaching cycles across what time period result in the best teacher achievement, and consequently student achievement.

Link to student achievement. Yoon, Duncan, Lee, Scarloss, and Shapley (2007) conducted a report for the Institute of Education Sciences (IES) on the relationship between teacher professional development and student achievement and concluded that students' achievement can be increased when teachers receive

substantial professional development but that demonstrating the link is difficult. Indeed, this relationship is not a well-studied area, consequently more studies continue to be urgently needed (Johnson, Kahle, & Fargo, 2007; Sleeter, 2014). Education experts have been asking for this for a long time, for example, almost twenty years ago Stiggins (1999) described students as the ultimate and often overlooked user of assessment and recommended that student outcomes should be at the forefront of educational research in teacher preparation and professional development in classroom assessment. In 2014, Sleeter reviewed leading teacher education journals and reported that there continues to be much published research on teacher professional development but still little that directly connects teacher learning with student outcomes that can be used to contribute to a coherent knowledge base that could inform policy.

Student focus. While gains were made in student achievements using SET4CER (chapter one), and Jessica felt that her kinesiology students had performed better on their summative quizzes when scaffolded (chapter three), the student perspective regarding the scaffolding provided is based upon their observed interactions which were provided by the classroom teachers and not the students' themselves, and their scores. A more individual take from each student would provide added value and leads me to recommend that future studies involving student interaction with assessment scaffolds should include a full student perspective, together with a research focus on comparing what students can do within their ZPDs with scaffolds and what they can achieve without scaffolding.

Many researchers of scaffolding have emphasized the importance of scaffolding learners' higher-order thinking skills in complex tasks (Kang et al, 2014; McNeill & Krajcik, 2006; Reiser, 2004) but few have shown how or if scaffolds influence students' thinking. Thus, a need continues for systematic empirical studies that include how the scaffolding processes influences learners' cognitive growth rather than only examining learning outcomes or achievement (Lin et al, 2012; Pea, 2004).

Personal Reflections

As little as five years ago I was a high school science teacher who became driven to find learning and assessment practices that could enable my students to be more successful in their mastery of chemistry. I felt certain that many of my students knew more, understood more, and could do more than was being demonstrated using traditional methods of summative assessment. That drive ultimately led to this dissertation and the journey has developed and reinforced my belief in the benefits to students and teachers of incorporating scaffolds into the assessment process. The notion of students' being assessed in their ZPD is of growing interest to me, particularly how best to compare what students can do while un-scaffolded and what they can do when scaffolded in their ZPD, and this may be an area I would like to develop in the future.

I have also come to realize that classroom assessment as a whole is more nuanced than I had imagined and I remain driven to find better ways to assess student learning and understanding. There is still much to be done to establish a detailed body

161

of knowledge on classroom assessment, more studies are needed in many areas (McMillan, 2013) including examining how teachers interpret and use the information from assessments for enhancing student learning, and of course more studies of the use of scaffolds in the assessment process. As a researcher, I am left feeling the urgency of all that is still left to do, but in my brief experience as a coach I am encouraged when the department chair of S.B. High school reports that her science teachers continue to ask their students for C-E-R explanations in place of full REEPEPA laboratory reports and that they continue to scaffold those explanations with the SET4CER tool. It seems appropriate then to conclude with Jessica's final remarks about her experience with assessment scaffolds:

> I've really liked the idea of using scaffolding during assessments, great idea. I think they helped students to clarify their thinking to enable them to answer questions. They are sort of like a replacement for being able to talk with a student to find out what they know. I found that students who didn't need them didn't use them, but those who needed them to help communicate their learning benefit from them. It's almost like personalizing the question for the student (Jessica, exit interview).

References

Campbell. C. (2013). Research on teacher competency in classroom assessment. In J. H. McMillan (Ed.), *Sage Handbook of Research on Classroom Assessment* (pp.71-85). Thousand Oaks, CA: Sage Publications, Inc.

Cornett, J., & Knight, J. (2009). Research on coaching. *Coaching: Approaches and Perspectives*, 192–216. Downloaded from: https://resources.corwin.com/sites/default/files/Cornett_Knight_2008.pdf.

Feiman-Nemser, S. (2001). From preparation to practice: Designing a continuum to strengthen and sustain teaching. *Teachers College Record, 103*(6), 1013-1055.

Gotwals, A. W., & Butler Songer, N. (2013). Validity evidence for learning progression-based assessment items that fuse core disciplinary ideas and science practices. *Journal of Research in Science Teaching*, *50*(5), 597–626.

Johnson, C. C., Kahle, J. B., & Fargo, J. D. (2007). A study of the effect of sustained, whole- school professional development on student achievement in science. *Journal of Research in Science Teaching, 44*(6), 775-786. doi:10.1002/tea.20149.

Kang, H., Thompson, J., & Windschitl, M. (2014). Creating opportunities for students to show what they know: The role of scaffolding in assessment tasks. *Science Education*, *98*(4), 674-704. doi:10.1002/sce.21123.

Knight, J. (2007). *Instructional Coaching: A Partnership Approach to Improving Instruction*. Thousand Oaks, CA: Corwin Press.

Knight, J., & Cornett, J. (2008). Studying the impact of instructional coaching. Downloaded from http://instructionalcoach.org/research/tools/ paper-studying-the-impact-of-instructional-coaching.

Killion, J. (2012). Coaching in the K-12 context. In S.J. Fletcher & C.A. Mullen (Eds.) *The SAGE Handbook of Mentoring and Coaching in Education* (pp. 273-294). London, UK: Sage Publications, Inc. Accessed online 1 Apr 2018. doi:10.4135/9781446247549.n19.

Lin, T., Hsu, Y., Lin, S., Changlai, M., Yang, K. & Lai, T. (2012). A review of empirical evidence on scaffolding for science education. *International Journal of Science and Mathematics Education, 10*(2), 437-455. doi:10.1007/s10763-011-9322-z.

McMillan, J. H. (2013). Why we need research on classroom assessment. In J. H. McMillan (Ed.), *Sage Handbook of Research on Classroom Assessment* (pp.2-17). Thousand Oaks, CA: Sage Publications, Inc. Accessed online 3 Nov 2014. doi:10.4135/9781452218649.n1.

McNeill, K. L., & Krajcik, J. (2006). Supporting students' construction of scientific explanation through generic versus context- specific written scaffolds. American Educational Research Association Annual Meeting, San Francisco, CA. Retrieved from http://www.umich.edu/~hiceweb/iqwst/Papers/McNeill_Krajcik_AERA2006.pdf.

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, doi:10.1207/s15327809jls1303_6.

Randel, B. & Clark, T. (2013). Measuring classroom assessment practices. In J. H. McMillan (Ed.), *Sage Handbook of Research on Classroom Assessment* (pp.145-164). Thousand Oaks, CA: Sage Publications, Inc.

Reiser, B. J. (2004) Scaffolding complex learning: The mechanisms of structuring and problematizing student work, *The Journal of the Learning Sciences*, *13*(3), 273-304, doi:10.1207/s15327809jls1303_2.

Sandoval, W. A., 2003, Conceptual and epistemic aspects of students' scientific explanations, *The Journal of the Learning Science, 12*(1), 5-51. Downloaded from http://www.jstor/org/stable/1466633. Accessed: 17-08-2016.

Sleeter, C. (2014). Toward teacher education research that informs policy. *Educational Researcher, 43*(3), 146-153. doi:10.3102/0013189X14528752.

Stiggins, R. J. (1999). Evaluating classroom assessment training in teacher education programs. *Educational Measurement: Issues and Practice, Spring 1999,* 23-27.

Yoon, K. S., Duncan, T., Lee, S. W.-Y., Scarloss, B., & Shapley, K. L. (2007). Reviewing the evidence on how teacher professional development affects student achievement (Issues & Answers Report, REL 2007–No. 033). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southwest. Retrieved from http://ies.ed.gov/ncee/edlabs.

Zhang, Z., & Burry-Stock, J. A. (2003). Classroom assessment practices and teachers' self-perceived assessment skills. *Applied Measurement in Education, 16*(4), 323-342. doi:10.1207/S15324818AME1604_4.