

Writing in Times of Deixis:

A Validation Study of a Large-Scale Assessment of New Literacies

Julie A. Corrigan

A dissertation submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Education

Faculty of Education, University of Ottawa

Thesis Advisors: Dr. David Slomp

Dr. Stéphane Lévesque

Committee Members: Dr. Michel Laurier

Dr. Cynthia Morawski

Dr. David Trumpower

External Examiner: Dr. Nadia Behizadeh

Faculty of Education, University of Ottawa

© Julie A. Corrigan, Ottawa, Canada, 2018

Acknowledgements

Without a number of important people in my life—both personally and professionally—this dissertation would not have been possible. During the course of my doctorate, I experienced a number of highs (birth of my twin daughters) and lows (death of my mother, debilitating accident of my spouse) that made the road to my Ph.D. all the more challenging. Without the love and support of family, friends, colleagues, thesis committee, and thesis supervisors, quitting would have seemed a legitimate option. But, with your unwavering support, I persevered. Thank you for standing by my side always.

First, I would like to acknowledge my family. Mom (I feel your presence with me now, even though you have since passed from this earth) and dad: you taught me from a very young age to never give up and to value the transformative power of education. Thank you for always believing in me and for being there when I needed a safe place to land. Also, to my husband Sean: you were always skeptical about my choice to leave behind a career I loved teaching high school. Despite this, I knew I could always count on you to make sure our family had a roof over our head. I know this journey has not been easy for you, so I thank you for staying by my side through it all. Thank you also to my in-laws, Bob and Cathie, for your support and love of our family during this journey. I will be ever grateful for weekend visits where I could sneak away for a few hours of work while the girls enjoyed some treasured time with their beloved grandparents. Thanks also to my brother David and his family, to my aunt Sally and Uncle Phil, and to all of my extended family who prayed for and loved me throughout this time.

Next, I would like to thank my thesis committee and external examiner. Without your expertise and wisdom, my thesis never would have evolved and improved as much as it has. To my external examiner, Dr. Nadia Behizadeh: your work has and continues to inspire my thinking

with regards to how we can do right by our students and do assessment better. Thank you for being a woman and a scholar whom I endeavour to emulate. In a similar way, Dr. Cynthia Morawski, you have always challenged me to think divergently and to put first the needs of those most disenfranchised by large-scale assessment. Your pursuit of socially just ways of teaching and assessing literacy have deeply formed my thinking. Thank you also to Dr. Michel Laurier for inviting me to present on a panel with you and for your tremendous insights regarding effective assessment. Dr. David Trumpower: I will never forget all of the guidance and support I've received from you through not one, but two degrees (my master's and my doctorate)! I've always appreciated how you were able to help me and all of your students to approach statistics without being intimidated. Your guidance has been absolutely instrumental to my growth as a scholar. Et toi, Dr. Stéphane Lévesque, un grand merci! You believed in me and entrusted me to become your research assistant, allowing me to bloom into an emerging scholar under your expert guidance. You treated me as a peer, not just a student, and that meant a great deal to the development of my confidence. You've been much more than a mentor—you've become a friend. I do hope our paths cross again many times in the years to come!

Dr. David Slomp... what could I possibly say to thank you for making this journey with me as my supervisor? You have been endlessly patient, immensely compassionate, extraordinarily wise, and always my biggest cheerleader. At those times when even I doubted what I was doing, you never stopped believing. Like Stéphane, you also treated me as a peer. Doubtless, the endless opportunities, conversations, and experiences you have afforded me have been foundational not only to my development as a scholar, but also as a person. From ice cream cones in Paris to tearful conversations when life threw me curve balls, you were there through it all. Anything I write here is completely inadequate to express my sincerest thanks, so I hope to

honour your commitment to me by always striving to undertake the most rigorous, transformative, and socially just research that I can. I can't wait to see what exciting collaborations lie ahead for us!

I would also like to acknowledge the following friends and colleagues for your support: Annie Koppens, Jordan Norris, Annette Furo, Jennifer Cogswell, Kelsey Schmitz, Rebecca Hogue, Maria Bastien, Danya, Caroline Doran, Terri Selle, Michelle Schira-Hagerman, Megan Cotnam Kappel, Shari Orders, Linda Radford, Trista Halloweck, Nicholas Ng-A-Fook, Ruth Kane, Cheryl Duquette, Christian Blanchette, Cheryl Maykel, Elena Forzani, and the many, many others who have inspired and continue to inspire my research and life!

A special thank you also goes out to Dr. Don Leu at the University of Connecticut's New Literacies Research Lab for allowing me to use the ORCA and of course for your most incredible mentorship. And, speaking of mentors, I would be remiss if I did not thank my dear friend and colleague Dr. Tony Onwuegbuzie for all of his mentorship and support.

Finally, I would like to thank my participants, without whom this research would not have been possible—my sincerest thanks to you!

Dedication

I would like to dedicate this dissertation to my mom, who has been my inspiration, and to my daughters, whom I hope to inspire.

In loving memory of my mother, Anne, and in honour of my daughters, Clara and Chelsea.

Abstract

This dissertation involves a holistic and interconnected examination of the validity, reliability, and fairness of the Online Research and Comprehension Assessment (ORCA). The ORCA is a large-scale assessment of New Literacies that challenges Grade 7 students to solve research problems (e.g., Does playing video games harm your eyes?) by locating, critically evaluating, and synthesizing online information in order to communicate their results in online genres such as email and wiki. My goal in this study was to understand how ORCA test score data should be used and interpreted, in what context, and for what purposes. A secondary goal was to examine the cognitive and metacognitive processes required to support research writing in online contexts.

The study involved three interconnected phases. The first was a systematic, mixed methods literature review of 101 peer-reviewed texts from the last 50 years in order to articulate the construct underlying the ORCA. Finding no construct in the literature that considered the important ways in which the Internet has changed the construct of writing, I opted to conceptualize one of my own. This construct also serves as the theoretical framework for the rest of the dissertation.

In the next phase of the study, I again explored the construct underlying the ORCA, but this time via a mixed methods investigation of the response processes—both cognitive and metacognitive—elicited by the ORCA. By observing both expert and novice participants' response processes, I analyzed the extent to which the tasks and types of responses elicited by the ORCA fit the intended construct. Further, by observing response processes, I was also able to analyze construct underrepresentation and construct-irrelevant variance, which are fundamental to the ORCA's appraisal. The results suggested that there are complex and

sophisticated cognitive and metacognitive processes underlying the ORCA and online research writing more generally, many of which are unique to online contexts. Further, both quantitative and qualitative results suggest significant differences between novice and expert groups.

The third phase of this research concludes with an integrated consideration of the ORCA's validity, reliability, and fairness. Here, I analyzed data collected from the previous two phases; previous validation work done on the ORCA by my colleagues; and new forms of validation evidence collected for this study. I did so in order to build a comprehensive validity argument to demonstrate the ways in which ORCA test scores should be used and interpreted, and the consequences which follow. I used cued retrospective reporting, semi-structured interviews, Venn diagrams, surveys, and writing artefacts to investigate the response processes elicited by the ORCA and to compare and contrast those to the writing practices that participants used in their school, work, and/or personal lives. I also completed an extensive analysis of the sample of observations permitted by the ORCA juxtaposing those with the target domain. Results of this study indicate that the ORCA provides an important form of assessment data regarding 21st century literacies previously neglected on traditional assessments. Limitations of the ORCA such as construct-irrelevant variance and construct underrepresentation are also explored. The results of the study suggest how the ORCA could be re-designed to improve the validity of inferences made.

Table of Contents

Acknowledgements	ii
Dedication.....	v
Abstract.....	vi
Table of Contents.....	viii
List of Tables.....	xiii
List of Figures	xv
<i>Chapter 1 - Introduction</i>	<i>1</i>
Introduction	2
The Online Research Comprehension Assessment	4
Statement of the Problem	10
Overview of the Study and Manuscript Structure	11
<i>Chapter 2 - Methodology</i>	<i>18</i>
Methodology	19
Embracing a Mixed Methods Way of Thinking	21
The Current State of Mixed Research in Writing	29
Embracing messiness, not sloppiness	31
Embracing a Commitment to Bettering the World	33
Conclusion	33
<i>Chapter 3 - Articulating an Integrated Contemporary Construct for the Digital Age</i>	<i>35</i>
Abstract	37

Articulating an Integrated Contemporary Writing Construct for the Digital Age	38
The Value of a Contemporary Construct.....	39
Situating an Integrated Construct in a Bioecological Model.....	41
Method.....	47
Sources of Data.....	48
Limitations.....	52
Mixed Methods Analyses	52
Results.....	55
Qualitative Results.....	55
Quantitative Results	75
Discussion	79
Implications for Curriculum, Instruction, and Assessment	80
Conclusion.....	87
Appendix	90
<i>Chapter 4 - Exploring the Differences between Student Writers and Knowledge Workers in</i>	
<i>Online Research Writing</i>	<i>108</i>
Exploring the Differences between Student Writers and Knowledge Workers in Online Research	
Writing	109
Abstract.....	110
Exploring the Differences between Student Writers and Knowledge Workers in Online Research	
Writing	111
Online Research Writing	112
Past Research on ORW	114
Models of Writing and Writing Expertise	116

A Bioecological Model of Writing	121
The Current Study	125
Method.....	126
Overview	126
Participants	126
Online Research Comprehension Assessment (ORCA).....	131
Procedures	135
Data Sources and Analyses	136
Results.....	140
Quantitative Results.....	141
Qualitative Results.....	146
Discussion	182
Cognitive Processes Enacted during ORW	183
Metacognitive Differences	193
Implications for Theory, Research, and Practice.....	196
Appendices	202
Appendix A Interview Protocol	203
Semi-structured Interviews (~20-30 minutes).....	203
Appendix B Rubrics Used for Scoring ORCAs	205
Table 2	205
Appendix C Interview and CRR Excerpts from Cross-Case Analysis of Metacognitive Processes in ORW	
206	
Appendix D Interview and CRR Excerpts Regarding Cognitive Processes Involved in Locating Online	
Information 212	
 <i>Chapter 5 - Using the Integrated Design and Appraisal Framework to Appraise a Large-Scale</i>	
<i>Assessment of New Literacies.....</i>	<i>245</i>

Abstract	247
Using the Integrated Design and Appraisal Framework to Appraise a Large-Scale Assessment of 21st Century Literacies	248
What is the ORCA?.....	249
Previous Validity, Reliability, and Fairness Research on the ORCA	260
The Need for Large-Scale Assessments of New Literacies	261
The Need for Validity Studies of Large-Scale Assessments of New Literacies	263
Literature Review and Conceptual Framework	265
A Framework for Examining Validity Evidence from the ORCA	265
A Bioecological Model of Writing	268
Online Research Writing	272
The Current Study	273
Method	274
Overview	274
Stimulus	275
Participants	275
Analysis	276
Results	276
Defining Assessment Context, Aims, and Principles of Fairness	276
Identifying Elements Foundational in the Assessment Design	280
Evaluating the Assessment Program	294
Appraising the Scoring System	301
Appraising the Assessment Results	304
Appraising the Assessment Consequences	310
Conclusion	314
Appendix A – Sample ORCA Report to Schools	320

Appendix B – Profile of Participants	328
<i>Chapter 6 - Conclusion</i>	<i>332</i>
Conclusion	333
Implications	343
Future Directions for Research	343
<i>Chapter 7 - References</i>	<i>347</i>
References	348
The Statement of Contribution of Co-Author	382

List of Tables

Table 1-1	The Four LESC Scenarios by Topic	6
Table 1-2	Overview of Study	12
Table 3-1	Means and Standard Deviations for the Transformed Values of the Coding Density Scores of Texts by Decade.....	78
Table 3-2	Writing Journals Identified for Hand Search and their Ratings.....	91
Table 3-3	Texts Identified in Literature Review and their Associated Coding Cumulative and Domain Densities.....	92
Table 3-4	<i>Texts with the Top 10 Cumulative Coding Densities</i>	104
Table 3-5	Means (and Standard Deviations) for Transformed Values of Coding Densities by Decade	106
Table 3-6	<i>Intercorrelations among Knowledge Domains</i>	107
Table 4-1	Demographic and ORW Characteristics of Expert Group.....	129
Table 4-2	Demographic and ORW Characteristics of Novice Group.....	130
Table 4-3	Correlations among ORW Measures (N = 20)	142
Table 4-4	Differences in metacognitive domains between novices and experts.....	143
Table 4-5	Case summaries from cross-case analysis of processes in ORW	148
Table 4-6	Case summaries from cross-case analysis of cognitive processes in ORW	151
Table 5-1	The Four LESC Scenarios by Topic	250
Table 5-2	Locate Scoring Criteria for the Cosmetic Contact Lenses and Eye Health Task ...	256
Table 5-3	Evaluate Scoring Criteria for the Cosmetic Contact Lenses and Eye Health Task	257
Table 5-4	Synthesize Scoring Criteria for the Cosmetic Contact Lenses and Eye Health Task	258

Table 5-5 Communicate Scoring Criteria for the Cosmetic Contact Lenses and Eye Health Task	259
Table 5-6 Locating Online Information: Target Domain vs. Sample of Observations	283
Table 5-7 Evaluating Online Information: Target Domain vs. Sample of Observations	286
Table 5-8 Synthesizing Online Information: Target Domain vs. Sample of Observations	289
Table 5-9 Communicating Online Information: Target Domain vs. Sample of Observations .	291
Table 5-10 Relationships among Sources of Variance, Fairness, and Validity	295
Table 5-11 ORCA Constructs and Score Points	301

List of Figures

Figure 1-1.	ORCA-Closed format with ‘Gloogle’ search engine	8
Figure 1-2.	Energy Drinks, Introduction: The introduction portion of an email LESC, along with the email the student clicks on to learn about her research task.	8
Figure 1-3.	The final task on the ORCA challenges students to communicate the results of their research findings either in an email or wiki format.	9
Figure 1-4.	A student edits a wiki page in the cosmetic contact lens LESC scenario.	9
Figure 1-5.	Video link to a high-performing student taking the ORCA-Virtual “Do cosmetic contact lenses harm your eyes?”	10
Figure 3-1.	A bioecological model of writing	42
Figure 3-2.	Flowchart of literature review process.....	49
Figure 3-3.	Integrated writing construct	56
Figure 3-4.	Number of articles coded per writing knowledge domain	76
Figure 3-5.	Coding density of knowledge domains by decade	77
Figure 4-1.	A Bioecological Model of Writing	122
Figure 4-2.	Facebook-like platform used in ORCA.....	132
Figure 4-3.	"Gloogle" Search Engine used in ORCA	133
Figure 4-4.	Participants communicate their research findings via wiki.....	134
Figure 4-5.	Differences between novices and experts on three ORW measures	142
Figure 4-6.	Metacognitive domain differences between novices and experts	144
Figure 4-7.	Cognitive Domain Differences between Novices and Experts	145

Figure 4-8. Excerpt from Catherine’s blog post in which she anticipates the objections of her audience	159
Figure 4-9. Image used by Darryl on the title page of his school project.....	161
Figure 5-1. ORCA-Closed format with ‘Gloogle’ search engine.....	253
Figure 5-2. Energy Drinks, Introduction: The introduction portion of an email LESC, along with the email the student clicks on to learn about her research task.	253
Figure 5-3. The final task on the ORCA challenges students to communicate the results of their research findings either in an email or wiki format.....	254
Figure 5-4. A student edits a wiki page in the cosmetic contact lens LESC scenario.	254
Figure 5-5. Video link to a high-performing student taking the ORCA-Virtual “Do cosmetic contact lenses harm your eyes?”	255
Figure 5-6. The bioecological model of writing.....	270

CHAPTER 1 - INTRODUCTION

Introduction

As I sit here writing this, I reflect on the ways in which the Internet and its associated technologies have changed the way I write. How specifically? The technologies I use have changed, that is for sure. Open on my computer at this very moment are a word processor; an online dictionary and thesaurus; no fewer than a dozen websites; [Mendeley](#) to manage and annotate my references; two social networking sites (Academic mums and Virtual SUAW [Shut Up And Write] Parents Edition) to inspire me, give me a sense of belonging, and keep me on track; Twitter; the library's online database; email; an online calendar; a [Pomodoro](#) application to help my productivity; and a cloud sharing application so that I can share my work with my thesis co-advisors. But aside from the technologies, what else has changed? I see how the Internet has changed my social practices, my strategies, and my dispositions too. When it comes to changing social practices, I think about how, this semester, I wrote a conference proposal and am working on a journal article with a colleague in another country. We have not seen each other in many months, yet we regularly write together. Technologies have made it so much simpler to collaborate across time and space. My strategies have changed too. I rarely go to the library to do research anymore, as I did during my undergraduate degree. Now when I go to the library, it is more to socialize. Ironically, while libraries were traditionally quiet work spaces, they have become community hubs. Instead, I find myself doing research mainly in online databases, but also using Twitter, blogs, and social networking sites from organizations and individuals renowned in my field. As far as dispositions go, I am trying to be more adaptable, for example. I frequently encounter new technologies, of course, but these technologies also generate new genres and new discourse communities to which I must adapt. I see my students trying to adapt too. This semester, I have witnessed an increasing number of

students opt for a multimodal assignment over a traditional essay to demonstrate their learning. They embed these objects into digital portfolios to share with their colleagues and potential employers.

Indeed, writing in the 21st century looks very different now—different from a century ago or a decade ago, but even from a year ago. We are writing in times of deixis, a term that linguists use to describe words that change as rapidly as their contexts (Leu, Kinzer, Coiro, Castek, & Henry, 2013). Take the word “new,” for example; by tomorrow, what is new is old. If writing has changed so much, why do we continue to use old technologies to assess new writing? For the most part, we continue to use pen and paper for large-scale assessments. Or, marginally better, we give students computers without Internet. More and more, argued Stagg Peterson, McClay, and Main (2011), “large-scale assessments are likely to become increasingly removed from the actual literacy practices of literate people” (p. 440). Because assessment drives instruction—what is known as washback in the assessment community (Hamp-Lyons, 1997; Messick, 1996; Morrow, 1986)—it is likely that choosing to ignore particular aspects of the writing construct on these assessments will lead to their not being taught in the classroom. In fact, the research has already shown this (Hillocks, 2002; Slomp, Corrigan, & Sugimoto, 2014; Stagg Peterson et al., 2011). For example, a systematic review of the literature of large-scale writing assessments across Canada that I co-conducted revealed exactly this sort of construct underrepresentation. Among other negative consequences, the results revealed that these tests ignored important elements of the writing construct such as writing process knowledge, critical literacies, digital literacies, and multiliteracies (Slomp, Corrigan, and Sugimoto, 2014). The disjuncture between real world and schooled literacy “leads to the irrelevance of school literacy in the eyes of the young” (Stagg Peterson et al., 2011, p. 440).

Those who design large-scale writing and literacy assessments have begun to respond to this disjuncture. For example, the Organisation for Economic Co-operation and Development's (OECD's) Programme for International Student Assessment (PISA); the Ontario Secondary School Literacy Test (OSSLT) in Canada; and the National Assessment of Educational Progress (NAEP) in the United States have moved from pen and paper to computerized testing formats. One large-scale assessment—the subject of this dissertation—has been innovated that challenges students to solve inquiry-based research problems in a fully online environment. The Online Research and Comprehension Assessment, or the ORCA as it is known, was developed by the team at the University of Connecticut's [New Literacies Research Lab](#) to assess students not only in a digital context, but in an online context.

The Online Research Comprehension Assessment

The ORCA emerged out of a recognition that 21st century literacy practices go beyond what Lankshear and Knobel (2007) have termed literacy, or traditional reading and writing. Empirical research from this dissertation and from others in the field (Coiro & Dobler, 2007; Forzani, 2016; Hull & Katz, 2006; Kiili, Laurinen, & Marttunen, 2008; Myers & Beach, 2001; Vasudevan, Schultz, & Bateman, 2010; Wiley et al., 2009) has shown that literacy in a digital age requires skills, strategies, social practices, and dispositions beyond those required for traditional literacy. These findings do not diminish the importance of literacy; they are simply a recognition that literacy on its own is insufficient for full participation in today's global information age. This realization led Leu, Kinzer, Coiro, and Cammack (2004) to develop New Literacies, a theory that recognizes the epochal changes brought about by the Internet and its associated technologies. The ORCA was developed as a way to collect data on a large-scale level regarding students' competencies in New Literacies. More specifically, the ORCA was

designed to be a valid, reliable, and practical measure of seventh-grade students' ability to use the Internet to conduct and communicate research (Leu, Kulikowich, Sedransk, & Coiro, n.d.). Like other literacy assessments, the ORCA is a hybrid reading-writing assessment. It should also be noted that the ORCA is not a high-stakes test. Rather, the ORCA is unique in that it is a large-scale formative assessment.

The current iteration of the ORCA contains four research scenarios (one per assessment version) that challenge students to investigate different life science topics through doing online research and communicating their results using online writing tools, particularly email or wiki. In each of these scenarios, students were presented with a problem that focused on the domains of health and human body systems, an area common to many seventh-grade science curricula. An example of a research question presented to participants from one of the ORCA scenarios is, "How do energy drinks affect heart health?" The scenarios were framed around two types of research: Learn More About (LMA) and Investigate Conflicting Claims (ICC). Half of the scenarios presented the research problem to students via an email message from the school board president (LMA scenarios) and half via a class wiki with a message from the teacher (ICC scenarios). LMA scenarios ask students to learn more about the research topic and to form a main idea about what they learn. ICC scenarios, on the other hand, ask students to investigate two sides of an issue and to take a position.

Table 1-1 presents the four LESC (i.e., Locate, Evaluate, Synthesize, and Communicate) scenarios used in the ORCA.

Table 1-1 The Four LESC Scenarios by Topic

Topic	Student Research Question	Type of Research	Communication Tool
Energy Drinks	How do energy drinks affect heart health?	Learn more about	Email
Heart-Healthy Snacks	How do snacks affect heart health?	Learn more about	Email
Contact Lenses	Do decorative contact lenses harm your eyes?	Investigate conflicting claims	Wiki
Video Games	Do video games harm your eyes?	Investigate conflicting claims	Wiki

This study uses a format of the ORCA known as the ORCA-Virtual, while past iterations of the ORCA have included formats called the ORCA-Open and ORCA-Multiple Choice (Leu, Coiro, Kulikowich, & Cui, 2012). The ORCA-Open represented the least restricted information space as assessment tasks used an open Internet space and the responses were in an open format (i.e., the task was a performance-based assessment, as opposed to a multiple choice one). The ORCA-Open format was eventually discontinued due to the instability of the assessment context. Unlike print-based texts, the form, content, and structure of online texts is in a constant state of flux. This presented a problem for the ORCA-Open because target websites used in the assessment were altered or would disappear altogether. In response to this challenge, the ORCA-Virtual was developed using a restricted Internet environment (such as ones used in online banking in the sense that, once you are on the bank's website, you cannot link to outside websites without logging off or opening a new window). This assessment format was designed to provide a more stable assessment context in that the while it takes place on the Internet, it was

not open to the Internet at large and was stable over time. Another current format used, although not investigated in this study, is the ORCA-Multiple Choice. The ORCA-MC is the most restricted format in terms of its assessment context (it used screenshots of websites, but was not open to the Internet at large) and its response format (multiple choice).

In the ORCA-Virtual, students conducted their research and selected information from pre-determined websites from the project's search engine named "Gloogle" (Figure 1-1). As the tasks were in a constructed-response format, the ORCA-Virtual is a performance-based measure. The LESC research problems appear within a Facebook-like environment via avatars named Brianna and Jordan (Figure 1-2) who are introduced as students from another school. The questions did not appear in a linear sequence according to their domain (i.e., Locate, Evaluate, Synthesize, and Communicate). Rather, a more natural and logical sequence was used according to the nature of the research task. Students were guided to engage in their online research and writing tasks via requests and questions from Brianna and Jordan. The ORCA culminates in a task requiring students to communicate their finding regarding the research problem via either writing an email or editing a wiki entry (Figure 1-3 and Figure 1-4). A video of a high-performing student taking the ORCA-Virtual, "Do decorative contact lenses harm your eyes?", may be viewed by clicking on the video link below (Figure 1-5). ORCA responses are assessed using a multidimensional scoring model for a total ORCA score of 16 points with 4 points per dimension each of locate, evaluate, synthesize, and communicate (LESC). Scoring procedures and will be explained in greater detail in Chapter V.

Figure 1-1. ORCA-Closed format with 'Gloogle' search engine

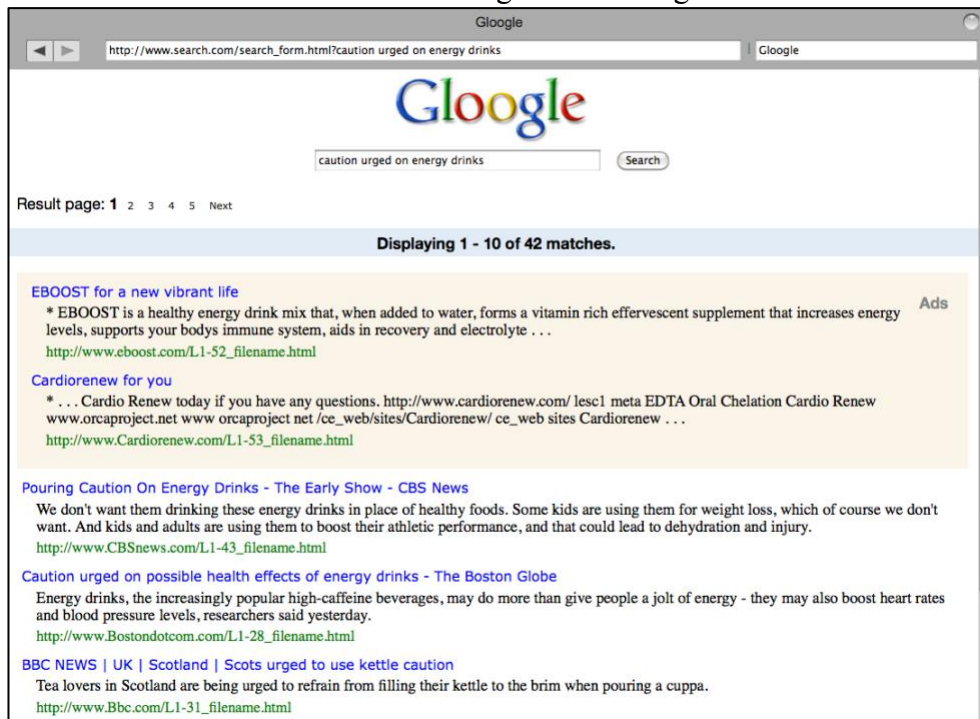


Figure 1-2. Energy Drinks, Introduction: The introduction portion of an email LESC, along with the email the student clicks on to learn about her research task.

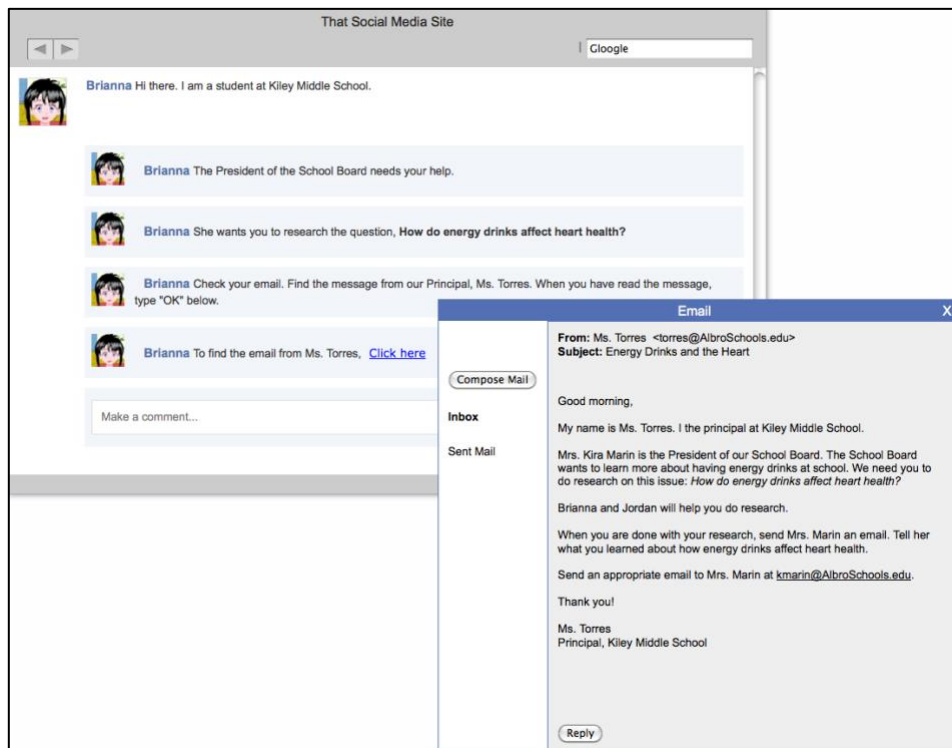


Figure 1-3. The final task on the ORCA challenges students to communicate the results of their research findings either in an email or wiki format.



Figure 1-4. A student edits a wiki page in the cosmetic contact lens LESC scenario.

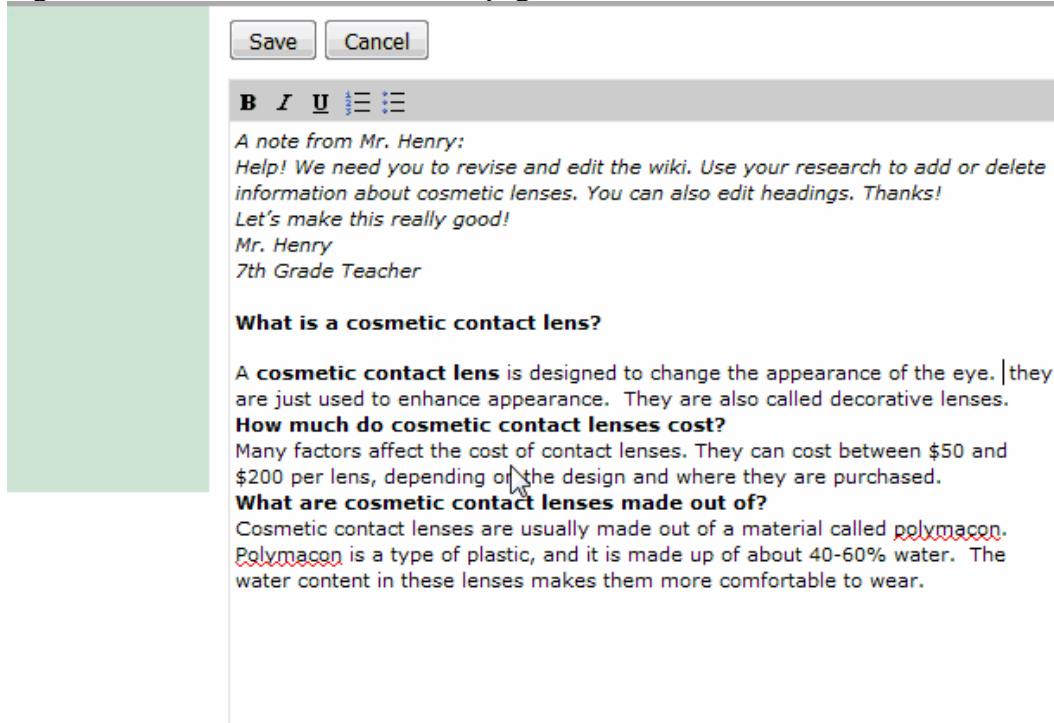
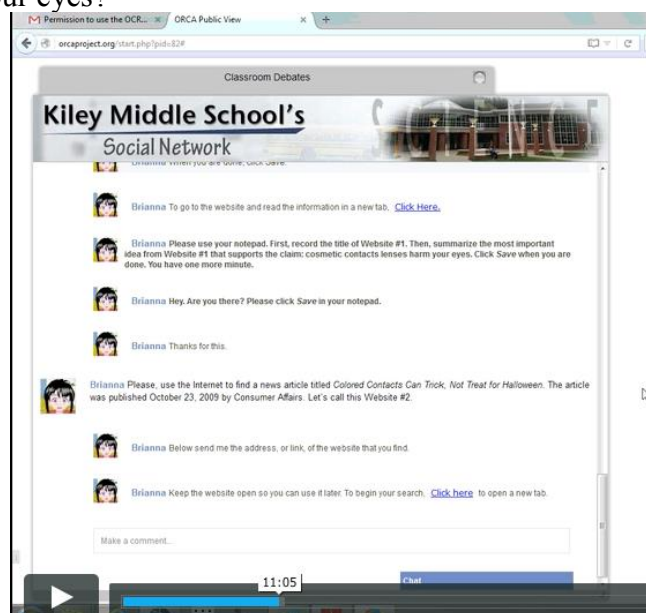


Figure 1-5. Video link to a high-performing student taking the ORCA-Virtual “Do cosmetic contact lenses harm your eyes?”



Statement of the Problem

As large-scale, online writing assessments continue to be innovated, there is an increasing need for validation studies to accompany them. While test designers typically carry out objective measures of a test’s psychometric properties, there is now wide-spread recognition among those in the assessment community that validation work needs to take a more holistic approach (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 2014; Kane, 2006, 2013; Slomp et al., 2014). Specifically, it is imperative for validation researchers to analyze (in addition to objective psychometric properties) how test score data are used and interpreted, in what context, and the consequences (intended and unintended) that follow. Therefore, this study used the Integrated Design and Appraisal Framework (IDAF; Slomp, 2016) in order to approach validation from a more holistic perspective, considering validity, reliability, and fairness in one model. A strength of IDAF is that it highlights the interdependent nature of validity, reliability, and fairness. That

is, each of those facets on its own is insufficient for the ethical design and validation of assessment. Evidence from each of those facets considered together, however, presents a more comprehensive and ethical validation argument. The IDAF is explained in greater depth in Chapter V.

Previous validation work regarding the ORCA completed by the New Literacies Research Team included the following: cognitive labs over two years with approximately 300 students; pilot testing with approximately 1600 students; and, a panel review of the ORCA by measurement and online research and comprehension experts. Previous reliability estimates by measurement experts determined that the KR-20 for the ORCA ranged from .86 to .90, depending on the version of the ORCA (Leu et al., 2014). I will expand the scope of the preceding validation work to include a holistic and interconnected examination of validity, reliability, and fairness as analyzed through the IDAF in order to better understand how ORCA test score data should be used and interpreted, in what context, and what the intended and unintended consequences of this might be.

Overview of the Study and Manuscript Structure

This doctoral dissertation is a three-phase study centred on this overarching research question: How should test score data from the ORCA be used and interpreted? The research questions for each of the phases were as follows:

1. Based on a systematic review of the literature, what is the construct underlying the ORCA?
2. Based on an empirical investigation of response processes, what is the construct underlying the ORCA?

3. Using the IDAF, how should ORCA test scores be used and interpreted, considering the results of the preceding literature review and empirical study, and additional research?

I have chosen to present my doctoral dissertation in article format comprised of three interrelated articles contextualized within a larger study. Each research question above coincides with its own study and article. In the current chapter, Chapter I, I situate the study and provide an overview of the study, organization of the manuscript, and overview of the methodology. Next, in Chapter II, I position myself in terms of my methodology. From there, I transition into the three articles (Chapters III - V). As I will explain, the first article also serves as the theoretical framework for my dissertation. Below, I present an overview of all three studies and their methods (Table 1-2).

Table 1-2 Overview of Study

Study	Research Questions	Method	Sample and Sources of Data	Analysis
<i>Dissertation</i> Writing in times of deaxis: A validation study of a large-scale assessment of New Literacies	How should test score data from the ORCA be used and interpreted?	Mixed and multiple methods, as listed below	All those listed below	All those listed below
<i>Study 1</i> New Writing: Articulating a Contemporary Construct for the Digital Age	Based on a systematic review of the literature, what is the construct underlying the ORCA?	A systematic, mixed methods review of the literature	<i>N</i> = 101 peer-reviewed texts written over the last 50 years	Qualitative: provisional coding, pattern coding, operational diagramming *Mixed methods: frequency counts, coding densities, one-way between groups ANOVAs

<i>Study 2</i>	Based on an empirical investigation of response processes, what is the construct underlying the ORCA?	Cued retrospective reporting (a retrospective, think-aloud method) and semi-structure interviews	$N = 20$ novice and expert participants ORCA test scores, cued retrospective reporting and interview transcripts, questionnaires, and writing artefacts from school and work	Qualitative: provisional coding, analytical coding *Mixed methods: magnitude coding, Mann Whitney U tests, MANOVA
<i>Study 3</i>	Using the IDAF, how should ORCA test scores be used and interpreted, considering the results of the preceding literature review and empirical study, and additional research?	Using IDAF to conduct an integrated analysis of validity, reliability, and fairness	Validity and reliability data from previous studies; data from studies 1 & 2	Qualitative analysis of the ORCA's validity, reliability, and fairness

* I refer to these as mixed methods (as opposed to quantitative) analyses because they involve a fully integrated form of mixed methods analysis. This involves first quantizing data (transforming qualitative data into a quantitative format) and then subsequently analyzing it using quantitative techniques.

Chapter III presents the first article, co-written with my thesis co-advisor Dr. David Slomp, entitled “New Writing: Articulating a Contemporary Construct for the Digital Age.” This first study involved a systematic literature review of writing research over the last 50 years and resulted in the articulation of construct we called new writing. This construct is timely in that it acknowledges the profound ways in which the Internet has revolutionized the construct of writing. Robust constructs are an integral part of IDAF. Not finding examples of robust constructs in the literature that met our needs, we opted to articulate one of our own. This construct then formed the basis for analysis in the subsequent two studies. Our mixed methods review of $n = 101$ texts revealed that the following writing knowledge domains have

predominated the literature: metacognitive, critical discourse, discourse, rhetorical, genre, writing process, subject matter, and digital literacies knowledge. Our quantitative results showed how the emphasis on these domains has changed over the decades, with the most recent decades developing a stronger focus on digital literacies knowledge. We bring these domains together to form a construct we call the integrated writing construct, which describes the knowledge domains necessary to develop expertise. We conclude by situating this construct within a bioecological model that elucidates the implications this model has for writing curriculum, instruction, and assessment. This construct then forms the basis for analysis in the subsequent two studies. As such, it also serves as the theoretical framework of this dissertation. We submitted this article to the *Review of Educational Research*, and the article went on to a second round of revision. We made substantial revisions, but unfortunately it was rejected at this point, despite two of the three reviewers requesting that the article be revised and resubmitted and commenting on the important contribution this article could make. We did receive some very positive feedback from this journal. Reviewer one said, “This manuscript has the potential for making a wonderful contribution to writing research and the way in which writing development, assessment, and curriculum development are viewed.” The third reviewer was also highly positive: “In this manuscript, the author(s) argues for a new construct, which they consider to be a contemporary and holistic model, for writing. The authors argue a new construct for writing is necessary because all of the old constructs are outdated. The authors do a nice job of laying out the main constructs that exist within the field of writing studies within the first pages of the manuscript. I have not seen these so clearly laid out and summarized before. This is a nice contribution providing a macro lens for writing studies and educational research. The manuscript adds something new and valuable to the field, providing a useful construct and a powerful

theoretical lens for writing research. I recommend this article be revised and resubmitted.” After making some revisions, we went on to submit the article to the *Journal of Writing Research*. Our article was once again rejected, but we were encouraged to rework the article and resubmit it again to this same journal. We look forward to building on this experience towards an even stronger article.

The second article, Chapter IV, is entitled “Exploring the Differences between Student Writers and Knowledge Workers in Online Research Writing.” In this second study I explored the differences in response processes between novices and experts during online research writing (ORW), and particularly during the ORCA. By observing participants’ response processes—including participants cognitive and metacognitive processes—I was able to analyze the extent to which the tasks and types of responses elicited by the ORCA fit the intended construct (American Educational Research Association et al., 2014; Goodwin & Leech, 2003). Further, by observing response processes, I was also able to analyze construct underrepresentation and construct-irrelevant variance, which are necessary steps in the IDAF (Slomp, 2016). Finally, as this study was a novice-expert study, I was also able to observe and analyze a continuum of processes and how these differed between the two groups. In this exploratory study, I used a mixed methods approach to investigate the differences between student writers (ranging from sixth to ninth grade) and knowledge workers (including a technical writer, lawyer, professor, registered dietician, financial management analyst, and nuclear safety analyst, among others) who were invited to complete the ORCA. Data sources for this study included test scores, but also cued retrospective reporting and interview transcripts, questionnaires, and writing artefacts from school and work. The results suggest that there are complex and sophisticated cognitive and metacognitive processes underlying online research writing and, tentatively, that some of

these processes are unique to online contexts. Secondly, both quantitative and qualitative results suggest significant differences between novice and expert groups. Implications for research, theory, and practice are discussed. I have presented part of this research on metacognitive writing processes at the Canadian Society for Studies in Education conference this past spring (2018) and I will be co-presenting a paper, specifically dealing with evaluating online sources, at the Literacy Research Association annual convention later this year. I plan on submitting this article to the *Journal of Literacy Research* because it is a highly respected journal with international reach, it is open access, and it has no maximum length (important, because this is a lengthy article) due to it only being published online.

The third and final article, Chapter V, is entitled “Using the Integrated Design and Appraisal Framework to Appraise a Large-Scale Assessment of New Literacies.” This study used a holistic approach known as the IDAF—wherein the interconnected issues of validity, reliability, and fairness are considered within one model—to appraise the ORCA. The study included data collected from the previous two phases; previous validation work done on the ORCA by my colleagues; and new forms of validation evidence collected for this study. Included in this were the expert-novice data from the previous study. I used Cued Retrospective Reporting, semi-structured interviews, Venn diagrams, surveys, and writing artefacts to investigate the response processes elicited by the ORCA and to compare and contrast those to the writing practices that participants used in their school, work, and/or personal lives. I also completed an extensive analysis of the sample of observations permitted by the ORCA juxtaposing those with the target domain. Results of this study indicate that the ORCA provides an important form of assessment data regarding 21st century literacies previously neglected on traditional assessments. Limitations of the ORCA such as construct-irrelevant variance and

construct underrepresentation are also explored. I have been considering the journals *Assessing Writing* and the *Journal of Writing Assessment* for this article.

This dissertation concludes with Chapter VI, which intertwines the results of all three studies and discusses the study's implications as well as future directions for research. In this chapter, I also include important recommendations for the re-design of the ORCA.

CHAPTER 2 - METHODOLOGY

Methodology

The intent of this chapter is to explain my methodology and thereby justify my use of a mixed methods approach. It is important to note that I will focus on methodology and not methods. As I used distinct methods for each study, these will be best explained each in their own chapter.

For many writing researchers, there is but one way of knowing and understanding the phenomenon of writing. Whether those researchers strictly adhere to quantitative or qualitative ways of thinking, “neither dogmatic adherence to the positivistic pipe dream nor chaotic methodological relativism [...] promises to advance research” (Howe, 1988, p. 14). I believe that in order to understand the complex, dynamic, social phenomenon that is writing, there is generative and synergistic value to combining words and numbers, messiness and measurement, inductive and deductive reasoning, emic and etic perspectives, exploration and confirmation, description and prediction, art and science. This combining will lead to convergence and corroboration, but also divergence and dissonance, and seemingly unending paradoxes; it will unsettle the settled, probe the contested, challenge the given, and engage the multiple, often from discordant perspectives and lenses (Greene, 2007).

I consider myself a xenophilous researcher (Onwuegbuzie, 2012). Be it through schizoanalysis (Deleuze & Guattari, 1983) or foreign statistical symbols, I am drawn to a mixed methods way of thinking. In the eloquent words of Greene (2007),

A mixed methods way of thinking involves an openness to multiple ways of seeing and hearing, multiple ways of making sense of the social world, and multiple standpoints on what is important and to be valued and cherished. A mixed methods way of thinking

rests on assumptions that there are multiple legitimate approaches to social inquiry, that any given approach to social inquiry is inevitably partial, and that thereby multiple approaches can generate more complete and meaningful understanding of complex social phenomena. A mixed methods way of thinking means genuine acceptance of other ways of seeing and knowing as legitimate. A mixed methods way of thinking involves an active engagement with difference and diversity. (p. xii)

I will argue that this dialectic stance on mixed research articulates a fertile approach to writing research, particularly for those who study writing assessment. While mixed methods research is the most commonly used term for research that integrates qualitative and quantitative approaches, I prefer the term mixed research (Johnson, Onwuegbuzie, & Turner, 2007), as it connotes the value of mixing beyond merely methods, but opens up the potential for mixing methodologies, stances, and philosophical approaches. I will also show that, although mixed research offers great potential to the field of writing, it is being underutilized.

Despite mixed researchers' epistemologically xenophilous approach to other forms of research inquiries and philosophies, sentiments from those outside this approach are not always reciprocated. Generally speaking, philosophers eschew mixed research for challenging the seemingly inviolate nature of paradigms¹, post-positivists for mixed research being messy, and

¹ Morgan (2007) noted four ways in which the term paradigm is being used: as a worldview, as an epistemological stance, as shared beliefs in the research field, and as model examples. In this paper I am referring to paradigms as metaphysical/epistemological stances, as per Lincoln et al.'s (2011) conceptualization.

constructivists for mixed research not being messy enough. I wish to take up these critiques in turn by challenging the holistic fallacy (Greene, 2007) of research paradigms, and then by justifying an approach to writing research that embraces a mixed research way of thinking; an approach that embraces messiness, not sloppiness; and finally, an approach that embraces a commitment to bettering the world.

Embracing a Mixed Research Way of Thinking

From the late 19th century until the 1980s, discussions surrounding positivism, and subsequently post-positivism, as opposed to the fledgling field of constructivism or interpretivism, were dominated by "trite clichés" based on "simple and rigid polarities," and there had been a "continual fixation upon what is 'good' about one approach or 'bad' about another" (Rist, 1977, p. 42, as qtd. in Smith & Heshusius, 1986, p. 5). This paradigm-deficit thinking (Onwuegbuzie, 2012) was subsequently applied to mixed research when it formally came into being some two decades ago. This occurred despite field sociologists, applied psychologists, and anthropologists having used mixed and multi- methods for quite some time (Johnson & Gray, 2010). Some would argue to 'stay put' within our own paradigm and 'not to talk to strangers' from foreign paradigms. They use the term *paradigm wars* (Onwuegbuzie, 2012) to connote that traversing paradigms is to traverse hostile territory. They prefer to stay confined to one paradigm even when approaches that engage and dialogue with multiple paradigms could greatly enrich writing research.

This paradigmatic xenophobia is justified by talk of paradigm incommensurability (Howe, 1988), which fails to acknowledge that "paradigms are social constructions, historically and culturally embedded discourse practices, and therefore [they are] neither inviolate nor fixed" (Greene & Hall, 2010, p. 121). Is meeting paradigmatic strangers easy? Certainly not. When

we focus on our similarities as opposed to our differences, the results can be lasting and meaningful. I am not advocating that we completely ignore differences, but rather that we come to understand them so that we can know the limitations and affordances of various paradigms in order to “compensate for inherent method weaknesses, capitalize on inherent method strengths, and offset inherent method biases” (Greene, 2007, p. xiii). Yanchar and Williams (2006) declared that mixed researchers have “little regard for challenging issues pertaining to the nature of reality, knowledge, the good, and so on” (p. 3), yet it behooves them to provide justification for this claim. This straw-man argument is easy to counter. As noted by Mertens (2012), “the [mixed research] community is awash in discussions about philosophical frameworks” (p. 255). The engagement with philosophy is evident when one looks at Tashakkori and Teddlie's (2010) *SAGE Handbook of mixed methods in social & behavioral research*, which devotes the first 11 chapters to philosophical, conceptual, and theoretical issues. To date, Onwuegbuzie and Frels (2013) have identified at least 13 philosophical stances associated with mixed methods research, with the most popular being pragmatism in its various forms (e.g., pragmatism-of-the-middle, pragmatism-of-the-right, pragmatism-of-the-left), followed by the transformative-emancipatory stance, and the dialectic stances (e.g., dialectical pluralism, critical dialectical pluralism). As is required of any attentive researcher, being a mixed researcher means knowing how epistemology, ontology, axiology, and methodology shape—and are shaped by—our research.

A *dialectic stance* towards mixed research views paradigms as “constituted by sets of interconnected philosophical assumptions regarding reality, knowledge, methodology, and values,” yet acknowledges that although these “assumptive sets of different paradigms are different in important ways,” as mentioned, paradigms are not sacrosanct (Greene, 2007, p. 69).

A dialectic stance in writing research could promote an open source approach wherein all writing

researchers are invited to contribute intellectual capital and benefit from the contributions of others—whatever their paradigmatic persuasion. I see the appeal to this, yet many do not. Perhaps many researchers feel “bound within a net of epistemological and ontological premises which—regardless of ultimate truth or falsity—become partially self-validating” (Bateson, 1972, p. 314; as qtd. Denzin & Lincoln, 2011, p. 13). Unless we become conscious of these silos, we cannot escape them in order to gain some distance from our own mental model² to see it for what it truly is. Also, if we do not free ourselves of these silos, how are we to “traverse the epistemological hiatus which opens up between the research traditions”? (Bryman, 1984, p. 80)

For many, the borders to other paradigms are closed, just as solid lines form borders around paradigms in popular research textbooks (e.g., Creswell's [2009] table of worldviews and Guba and Lincoln's [2009] Basic Beliefs of Alternative Inquiry Paradigms table). As Guba (1987) once said, "the one [paradigm] precludes the other just as surely as belief in a round world precludes belief in a flat one" (p. 31). Fortunately, in my opinion, Guba's stance on paradigm commensurability is (nearly) as fluid as the borders drawn around paradigms. In a more recent statement, Guba admitted that “there is great potential for interweaving of viewpoints, for the incorporation of multiple perspectives, and for the borrowing, or *bricolage*, where borrowing seems useful, richness-enhancing, or theoretically heuristic” (Lincoln, Lynham, & Guba, 2011, p. 115).

² According to Greene (2007), a “mental model is the set of assumptions, understanding, predispositions, and values and beliefs with which all social inquirers approach their work. Mental models influence how we craft our work in terms of what we choose to study and how we frame, design, and implement a given inquiry. Mental models also influence how we observe and listen, what we see and hear, what we interpret as salient and important, and indeed what we learn from our empirical work” (p. 12).

To understand what is preventing researchers—notably social science philosophers—from embracing a mixed methods way of thinking, one needs simply to look at the numerous philosophical debates that continue to brew at conferences, in books, and in multiple journal articles (Creswell, 2011; Ellingson, 2011; Ercikan & Roth, 2006; Guba & Lincoln, 2009; Howe, 1988; Yanchar & Williams, 2006) discussing paradigm incommensurability:

One of the main problems with the notion of paradigm is that it tends to bring under one heading a range of different ideas and assumptions that do not *necessarily* have to go together. This tends to make the notion of paradigm into a container concept and leads to a situation in which paradigms have to be embraced or rejected in a whole sale manner rather than letting the discussion focus on smaller elements, such ontological, epistemological, or methodological views and assumptions. The so-called “incompatibility thesis” (see Howe, 1988) is an example of how positions can become ossified when articulated in terms of the parts that make up a particular point of view. (Biesta, 2010, p. 98)

Rather than seeing paradigms as incompatible, taking a dialectic stance acknowledges that any one paradigmatic stance inevitably gives us only a partial understanding of human phenomena. Thus, by deliberately incorporating multiple stances we might gain a more complex, holistic perspective (Greene & Hall, 2010). Most social science philosophers agree that there is no issue with mixing data or even methods (see Biesta, [2010] for a discussion of mixing at various levels). However, philosophers disagree on when and where mixing can or cannot occur at a metaphysical level. This especially happens when we see the world in terms of binaries, rather than continua (Niglas, 2010). By taking a dialectic stance, I believe that mixing at a metaphysical level is legitimized by the diversity of perspectives that is elicited. Instead of

using one lens, using multiple lenses can give us new ways of seeing, hearing, knowing, and understanding complex phenomena.

But what might this look like in practice? I will now demonstrate how I used a dialectic stance in this dissertation. In particular, I will focus on my second study regarding cognitive and metacognitive differences between novices and experts. In this study, I combined a realist ontology with a constructivist epistemology (often called critical realism; Maxwell & Mittapalli, 2010). By having a realist ontology, I took the perspective that there is a real world that exists independently of my or others' perception of it. In this study, I took the stance that cognitive and metacognitive processes are real phenomena, rather than merely constructions. For example, I believe that triaging (i.e., deciding how much time and resources to devote to a particular task based on its importance to self and/or others), a metacognitive process I discovered primarily among expert writers, truly exists. I coupled this realist ontology with a constructivist epistemology. My understanding and my participants' understanding of triaging is inevitably a construction from our unique perspectives and collective histories. I also mixed emic and etic perspectives to gain a better understanding of these processes. For example, to understand the emic perspective on metacognitive processes used in writing, I used cued retrospective reporting (CRR) and semi-structured interviews wherein I invited participants to construct meaning around these processes. During CRR I would pause during portions of the video capture playback and say, "explain to me what you are doing here." For an etic perspective, I would analyze the video capture as an outsider. What did I observe the participants doing that perhaps they did not reveal during CRR or the interview? Also, I took an etic perspective by examining their responses on the ORCA tasks, conducting within and cross-case analyses between novices and experts. While analyzing data, I noticed that data from the etic and emic perspectives sometimes contradicted

one another—which is the whole point of taking a dialectic stance. This was more often the case working with data from the novice, emic perspective. For example, novices often portrayed a naïve sense of self-efficacy regarding their cognitive and metacognitive processes. Experts, conversely, had a wealth of experience and education to draw upon to quite accurately judge their self-efficacy.

To me, it is reductionist thinking to see ontology as either realist or relativist; epistemology as inductive or deductive; and, lenses as emic or etic (Lincoln et al., 2011; Onwuegbuzie, 2012). When it comes to ontology, I see reality as intersubjective. For instance, suppose we are conducting a study exploring whether or not a particular writing intervention is effective in helping students learn to write better. Student *x*'s experience might be that the intervention was not effective. Student *y*'s experience might have been that it was effective. Both students' realities are equally real. One person's reality does not preclude the other's from being true. Or in the words of Biesta (2010), "everyone's experience is equally real" (p. 107). Thus, reality is intersubjective: each person's experience is true, but at the same time, there is no one experience that is true. Additionally, reality varies not only by person, but also group, place, and time. What is true for student *x* today may not be true next year or if she moved to another region or if she belonged to another group. However, when we are investing large sums of money into the implementation of a writing intervention, for example, we need to know what is true about the effectiveness of the intervention for that group and its individuals, at this time, and in this place.

And what type of reasoning might we used to arrive at an answer to such a question? Taking a dialectic stance I believe that I can know through using abductive reasoning, "anchored by inductive reasoning at one end of the continuum and deductive reasoning at the other end of

the continuum, and the midpoint representing the place where inductive and deductive reasoning are most interactive” (Onwuegbuzie, 2012, p. 205). I would use abductive reasoning as “a way to work back from an observed consequence to a probable antecedent to create insights” (Christ, 2010, p. 662). Of course, my knowledge about whether or not the intervention is effective can at best only be partial. This is because humans are dynamic and complex, predictable and at the same time unpredictable when they have sufficient motivation for doing so.

As humans, we both construct reality (relativist ontology, subjectivist epistemology) and are constructed by it (realist ontology, objectivist epistemology). Even plants do this, to a certain extent.³ In other words, variables shape people and people shape variables. A classic example of variables shaping people (deterministic, post-positivist thinking) is the high correlation between socio-economic status (SES) and academic achievement. When we look at a large enough sample size, patterns emerge; this is known in the statistical world as the theory of normal distribution. These patterns can be discovered using statistical analyses. But when people behave with intentionality (say, as my first-generation Polish father did by being the first in his farming family to go on to higher education), they can defy statistics and become outliers (relativistic, constructivist thinking). When enough people become outliers, they are no longer so, and thus the variable is also changed.

As a writing researcher, I see great value in taking a dialectic stance to mixing methods so that I can study how variables shape people while simultaneously studying how people shape

³ Plant growth is, reductively speaking, a function of temperature, light, and moisture. More plants growing (especially if people plant them or if animals do by carrying seeds), however, causes a change in the temperature and CO₂ outputs, thus altering the environment. The environment in turn alters what types of plants can grow in the future. This iterative process is a function of any open system.

variables. In my own experience as a visiting doctoral student in the New Literacies Research Lab at the University of Connecticut, I have had the opportunity to investigate how over 1,100 seventh graders in two states—one with a state-wide one-to-one laptop program and low socio-economic status, and one with a high socio-economic status but no laptop program—perform in their ability to conduct online research and communicate their findings via email or wiki. We know from the literature that low SES correlates with low literacy scores (Aikens & Barbarin, 2008). We also know that females outperform males on literacy tasks by a small margin (Brookhart, 2009) and writing tasks by a significant margin (Cole, 1997). Some of my colleagues have learned that one-to-one laptop programs may compensate for issues of SES (Kennedy, Rhoads, & Leu, 2016). My contribution has been to discover that, while females still outperform males in online environments, the effect size is significantly lower than in traditional forms of writing (Corrigan, 2014a). While the study has to this point solely benefited from a post-positive paradigm (how variables shape people), I can see the complementarity⁴ that would be gained from a constructivist approach (how people shape variables)—so many more and richer insights could be gained. For example, it would be interesting to use the quantitative data to select a sample of outliers, such as those who achieved high scores on the online assessment, despite being from a low SES community. What would using think-aloud protocols, eye-tracking analyses, and interviews reveal that our statistical methods previously left concealed? My guess is, a great deal. It should be noted here that while certain stances towards mixing methods advocate using one paradigm (that of pragmatism, the melting pot of paradigms), a

⁴ Complementarity refers to mixing research for the purpose of “seeking broader, deeper, and more comprehensive social understandings by using methods that tap into different facets or dimensions of [a] complex phenomenon” (Greene, 2007, p. 101).

dialectic stance (the mosaic of paradigms) advocates for the productive tensions that are created by the polyvocality of multiple paradigms.

The Current State of Mixed Research in Writing

Writing researchers are beginning to make a foray into mixed research. However, in my opinion, mixed research is being underutilized. Many from within the field have commented that writing research for the most part has remained siloed (Abbott & Berninger, 1993; Graham, Gillespie, & McKeown, 2013; Hacker, Keener, & Kircher, 2009). Parsons and Gallagher (2016) have urged literacy researchers to honour diverse methods and perspectives in the field:

Engaging in scholarly dialog across thought collectives will help us identify anomalies in our current understanding, which will advance us as a field by compelling additional research and subsequent paradigm shifts to help us move toward advanced understandings of literacy processes, literacy teaching, and literacy learning. (p. 499)

While there are no prevalence rate studies to date on the use of mixed research in the field of writing, a study by Parsons and Gallagher (2016) of 1,238 literacy articles (among which were numerous writing articles) might shed some light. Parsons and Gallagher completed a content analysis of articles from nine influential literacy journals published between 2009-2014. Among other things, they investigated the research designs of the studies published in these articles. They found that across 1,238 articles that they analyzed, 78 articles (or 6.3%) used a mixed methods design. The authors noted that while some journals favoured either qualitative or quantitative research designs, a few journals had a relative balance among qualitative, quantitative, and mixed research approaches. These included the *Journal of Literacy Research*, *Reading Research Quarterly*, and *Literacy Research & Instruction*.

Important advancements in writing research have been made from both quantitative and qualitative camps of writing researchers, and mixed approaches are beginning to make a foray. But, can either one alone suffice to support the complex, dynamic, and interdisciplinary field that is writing? And what of the burgeoning impact of mixed research in social inquiry (Creswell & Plano Clark, 2011; Denzin & Lincoln, 2011; Hesse-Biber & Johnson, 2013; Onwuegbuzie, Johnson, & Collins, 2009; Onwuegbuzie & Leech, 2005)? What role could a dialectic stance (Greene, 2007; Greene & Hall, 2010) play in advancing writing research? Considering the multiple disciplines (e.g., sociology, psychology, linguistics, sociolinguistics, psycholinguistics, semiotics, composition studies, media studies, rhetorical studies, communication studies) and multiple theoretical frameworks (e.g., cognitive theory, sociocultural theory, critical theory, feminist theory, actor-network theories, postcomposition theory, postprocess theory, ecomposition, New Literacies, New Literacy Studies, digital literacies, multiliteracies) that writing research draws upon, choosing but one paradigm may prove too restrictive or too reductive for researchers of writing, especially in an age of emergent technologies, emergent social practices, and emergent research methods. Writing has become far too complex for those from divergent disciplinary and theoretical backgrounds to remain in their intellectual silos; progress will be seen when we dialogue with perspectives other than our own forming new collaborations and new insights along the way.

Where this xenophilous research (Onwuegbuzie, 2012) is especially needed is in the field of writing assessment. Writing researchers and those in educational measurement—the two communities of practice most invested—hold different purposes for assessing writing and “distinct conceptualizations of validity in writing assessment” (Behizadeh & Engelhard, 2015, p. 35). For example, writing researchers are primarily concerned with assessing writing for the

purpose of “informing and improving instruction for culturally and linguistically diverse students” (Behizadeh & Engelhard, 2015, p. 38). For these researchers, sociocultural theory dominates notions of validity because writing is viewed as sociocultural and situated. Thus, validity evidence for those in the writing community is typically qualitative and argument-based. Similarly, those in the assessment community also consider argument-based approaches to validity, specifically, establishing appropriate uses and interpretations of assessment data. However, though this is a concern, the focus tends to be on specific forms of validity evidence (e.g., construct, content, criterion-related, and consequential; Behizadeh & Engelhard, 2015). I would argue that in order to have a holistic approach to validation, both qualitative and quantitative evidence are necessary. Messick emphasized the evidentiary nature of validity (Messick, 1989). Validation, within a unified theory, requires the collection of multiple forms, classes, or types of evidence. This unified theory of validity continues to inform contemporary models of validation (American Educational Research Association et al., 2014; Kane, 2006, 2013; Slomp et al., 2014). Contemporary perspectives point to validation being both a qualitative and a quantitative pursuit after all (Cronbach, 1989; Kane, 2006; Messick, 1992; Moss, 1994).

Arguably, a dialectic mixed methods stance is an attractive option for writing researchers, and particularly those in writing assessment. If so, why are so few writing researchers employing mixed research strategies? As the next sections show, for some, mixed research is too messy, while for others, it is not messy enough.

Embracing messiness, not sloppiness

If the world is complex and messy, then at least some of the time we're going to have to give up on simplicities. But one thing is sure: if we want to think about

the messes of reality at all then we're going to have to teach ourselves to think, to practice, to relate, and to know in new ways.

—John Law (2004, pp. 1-2), *After Method: Mess in Social Science Research*

Unlike hard science, research in the social sciences involves one large, uncontrollable variable: human behaviour, or as Philips (2011) puts it, “cause only becomes effect through the medium of wilful human action” (p. 17). This is messy business for social scientists. That being said, some of the greatest scientific discoveries have been made when scientists were messy—witness Pasteur discovering antibiotics by accidentally leaving out a bacterial culture, or Roentgen discovering X-rays when a screen coated in barium happened to be standing nearby his cathode-ray tube. Although these scientists were messy, there was no sloppiness in the ways in which they subsequently recorded their discoveries. Post-positivists often accuse both mixed and qualitative researchers alike of messiness, when what they really mean is sloppiness (for messiness is unavoidable in social inquiry). If social inquiry is “an awful mess... then would something less messy make a mess of describing it?” (Law, 2004, p. 1) Good research distorts chaos into clarity so that we may gain some approximation of that which is messy, complex, and diffuse. If methods produce the ways of knowing they describe, and those ways of knowing produce knowledge, then a dialectic stance towards mixing methods will be needed to undertake knowing, seeing, hearing, and understanding in multiple ways.

Practically speaking, though, what can social science researchers do to embrace messiness, while avoiding sloppiness? Collins, Onwuegbuzie, and Johnson (2012) advocate for *philosophical clarity* as being one of the legitimization criteria used to evaluate mixed research; this philosophical awareness represents “the degree that the researcher is aware of and articulates her/his philosophical proclivities in terms of philosophical assumptions and stances in relation to

all components, claims, actions, and uses in a mixed research study” (p. 855). Legitimation is mixed research’s counterpart to validity in quantitative research; or in the qualitative world, to credibility, transferability, dependability, and confirmability (Denzin & Lincoln, 2011). It is especially incumbent upon dialectic stance researchers to have philosophical clarity when they blend aspects of mental models, or blend with others from different mental models, and not to take a sloppy approach where “anything goes.” When we are blending across philosophical stances, there is a polyvocality that needs interpretation for those from distant paradigms. Clearly articulating what guides and directs our research, and what we know and value, will provide a guidebook of sorts to our visitors.

Embracing a Commitment to Bettering the World

In this last but most important section, I wish to implore writing researchers to embrace a commitment to bettering the world. Instead of using pejorative terms to describe other ways of knowing, seeing, hearing, and understanding the world, let us get to the matter of addressing complex social problems. Many studies in writing research show that there are glass ceilings to be smashed (Cole, 1997; Newkirk, 2000; Pajares & Valiante, 2001); cycles of poverty to be broken (Brandt, 2003; Neuman & Celano, 2012); digital divides to be bridged (OECD, 2011; van Dijk, 2006); and linguistic, regional, and racial inequities to be leveled (Cole, 1997; Kinloch, 2010). If we strive for the quality criteria that are ascribed to a dialectic stance of mixing methods—difference, understanding, tolerance, holism, integration, explicit values, and social change (Collins et al., 2012)—our research might go a long way in making a difference.

Conclusion

In closing, I would like to add that although I believe in the generative and synergistic potential of taking a dialectic stance to mixing methods (especially in the dynamic and complex

field of writing), I acknowledge that there are many valuable ways to do research—what kind of dialectic mixed researcher would I be if I did not? Denzin (2008) wisely reminds us that “a change in paradigmatic postures involves a personal odyssey; that is, we each have a personal history with our preferred paradigm and this needs to be honoured” (p. 322). Dialectic mixed research is not for everyone. Yet, I believe all researchers can benefit from embracing multiple ways of knowing, embracing messiness not sloppiness, and embracing a commitment to bettering the world.

**CHAPTER 3 - ARTICULATING AN INTEGRATED CONTEMPORARY
CONSTRUCT FOR THE DIGITAL AGE**

Articulating an Integrated Contemporary Writing Construct for the Digital Age

Julie A. Corrigan

University of Ottawa

David H. Slomp

University of Lethbridge

Author Note

Julie A. Corrigan, Faculty of Education, University of Ottawa; David H. Slomp, Faculty of Education, University of Lethbridge.

This research was generously supported by a Social Sciences and Humanities Research Council of Canada Bombardier Canada Graduate Scholarship, Ontario Graduate Scholarships, and internal funding from the University of Ottawa.

Correspondence should be addressed to Julie A. Corrigan, Faculty of Education, University of Ottawa, 145 Jean-Jacques-Lussier Private, Ottawa, ON K1N 6N5, Canada

Abstract

This systematic review of the literature integrates nearly 50 years worth of writing research and theory in order to articulate an updated construct, one that meets the contemporary needs of those who research, teach, and assess writing ability—particularly in a digital age. Our mixed methods review of $n = 101$ texts revealed that the following writing knowledge domains have predominated the literature: metacognitive, critical discourse, discourse, rhetorical, genre, writing process, subject matter, and digital literacies knowledge. Our quantitative results showed how the emphasis on these domains has changed over the decades, with the most recent decades developing a stronger focus on digital literacies knowledge. We bring these domains together to form an integrated writing construct, which describes the knowledge domains necessary to develop expertise. We conclude by situating this construct within a bioecological model that elucidates the implications this construct has for writing curriculum, instruction, and assessment.

Keywords: writing theory; writing assessment; construct; bioecological; New Literacies

Articulating an Integrated Contemporary Writing Construct for the Digital Age

The goal of this systematic, mixed methods review of the literature is to offer a new construct of writing for the digital age, one that better supports curriculum design, pedagogical practices, and assessment techniques tailored to the needs of contemporary students. The advent of the Internet has triggered a concomitant revolution in the construct of writing. As the screen supplanted the page as our primary mode of communication, the practice of writing has profoundly shifted. Where once writing was static, linear, and (mostly) monomodal, it has become hypertextual, a-linear, and multimodal. Further, in our global village, we are challenged to adapt to the influence of multiple Englishes and multiple sociocultural contexts (New London Group, 1996). We are writing in times of deixis—a term linguists use to describe words whose meanings change as rapidly as their contexts (Leu, Kinzer, et al., 2013). Writing using the Internet and its associated technologies has led to the need for new semiotics, new grammars, new discourses, new genres, and new writing processes. Despite the momentous changes that the construct of writing has undergone, we rely on outdated constructs to measure and conceptualize writing. But do these changes necessarily merit a new construct? We would argue that they do indeed. While traditional constructs are foundational to this new writing construct, the skills, strategies, dispositions, and social practices required to write in a digital age are so fundamentally different (as will be argued throughout this paper) that the time has come to articulate a new construct.

Writing theory—through which this construct is defined and subsequently measured—has simply not kept pace. Currently, there exist theories that describe the centrality of the Internet to literacy, but none are specific to writing. These theories are known as digital literacy (Gilster, 1997) or digital literacies (Lankshear & Knobel, 2008), multiliteracies (Cope &

Kalantzis, 2009; New London Group, 1996), New Literacy Studies (Gee, 1990; Street, 1995), Multiple Literacies Theory (Masny, 2011), and New Literacies (Coiro, Knobel, Lankshear, & Leu, 2008). There are also important theories that describe the construct of writing—for example, semiotic (Kress, 2003; Kress & van Leeuwen, 1996), cognitive (Bereiter & Scardamalia, 1987; Flower & Hayes, 1981; Hayes, 2012; Leijten, Van Waes, Schriver, & Hayes, 2014; MacArthur & Graham, 2016), metacognitive (Hacker et al., 2009), sociocultural (Bazerman, 2016; Prior, 2006; Vygotsky, 1997), genre (Derrida & Ronell, 2009; Rose, 2016), rhetorical (Berlin & Inkster, 1980), and process (Emig, 1971; Flower & Hayes, 1981; van den Bergh, Rijlaarsdam, & van Steendam, 2016) theories of writing—though none account for the profound ways in which the Internet has revolutionized it. Thus, we propose an integrated construct in order to enrich and extend current conceptions of writing.

We begin this paper by articulating the need for an updated construct of writing, situated within a bioecological context. Next, we describe the mixed methods, systematic review of the literature that we used to arrive at this construct. Additionally, we present trends in writing research—specifically, what aspects of the writing construct were emphasized in empirical and theoretical research over the last 50 years. Finally, we discuss the implications of this updated construct for designing curriculum, instruction, and assessment.

The Value of a Contemporary Construct

Clearly defined and thoroughly researched constructs are essential for excellence in education because they form the basis of pedagogical models, curriculum standards, and assessment practices and test designs. To teach well and to assess effectively, educators need to have a clear understanding of the knowledge, skills, and dispositions that are core to the outcomes they are attempting to cultivate in students.

For more than 50 years now, research into the construct of writing has been conducted across a range of research traditions (Behizadeh & Engelhard, 2011). For the most part, however, this research has remained siloed and has yet to be integrated into a more holistic model (Abbott & Berninger, 1993; Graham et al., 2013; Hacker et al., 2009). This situation leaves the construct ill-defined, a clear problem for the field of writing assessment (Sainsbury, 2009; Shermis, Burstein, & Leacock, 2016; White, Elliot, & Peckham, 2015). As a consequence, no cognitive or developmental scale used to *directly* measure writing ability has been developed to date (Burdick et al., 2013). Instead, writing assessment programs have depended on the evaluation of narrowly defined, written products to extrapolate a writer's cognitive or metacognitive capacity for writing. For example, in a systematic review of the literature of large-scale writing assessments across Canada, we found that the tests constrained writing as a construct by assessing only the final product and ignoring the writing process, focusing more on procedure (write *x* number of paragraphs) than actual writing, ignoring digital writing and multiliteracies, focusing on narrow/traditional conceptions of writing, and ignoring the importance of multiple assessments over time (Slomp et al., 2014). These narrow, indirect measures of writing ability are problematic because the inevitable gaps between cognitive/metacognitive ability and the written products students produce give a distorted measure of students' writing ability.

The absence of an integrated construct of writing also creates problems for curriculum and instruction because siloed constructs do not fully account for the range of factors that contribute to the development of writing ability. Historically, research into the writing construct has fallen into two distinct camps: the cognitive and the sociocultural. Research in the cognitive camp attempts to understand writers and their cognitive processes while research in socio-

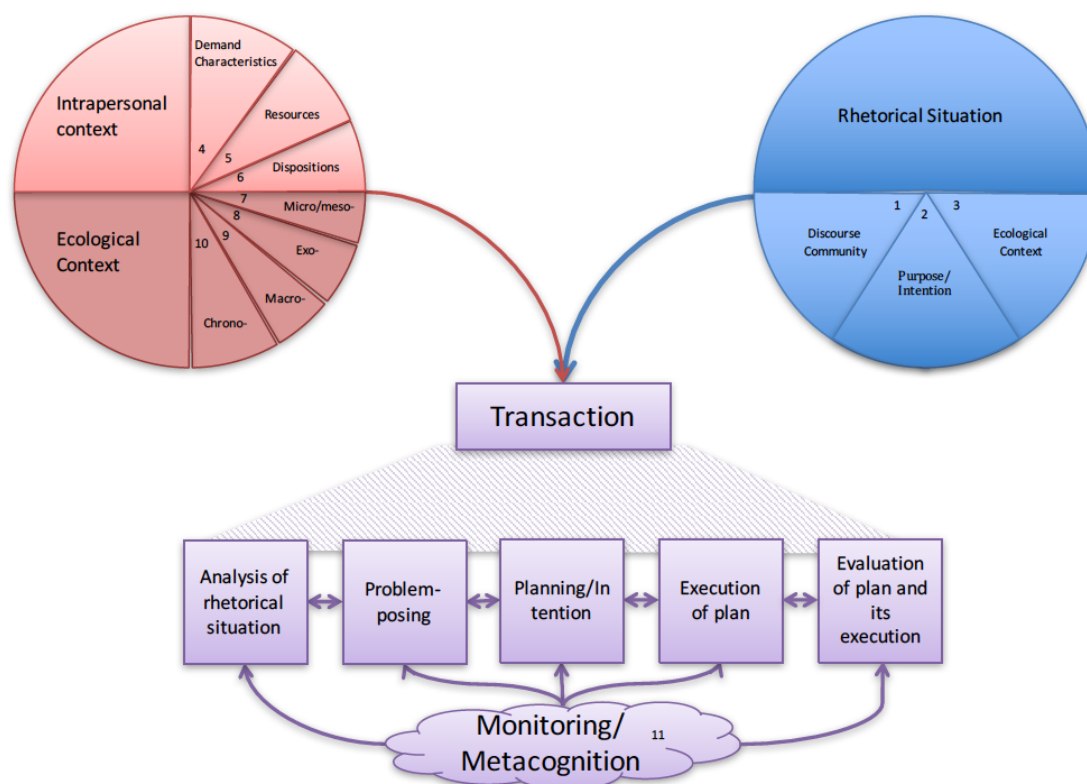
cultural camp attempts to understand how social context shapes the demands and expectations placed on writers. Ivanic's (2004) discourses of writing and learning to write is one of the few exceptions because integrated into his framework are both cognitive and sociocultural approaches to understanding writing. We see parallels between our work and Ivanic's (2004), and intend to build on this earlier attempt to unite the camps. Graham and Harris (2013) point out that bifurcation of the field is a mistake as writing development and instruction will never be adequately understood without considering both perspectives. By articulating an integrated construct, we are attempting to articulate how cognitive processes and sociocultural forces shape writing ability thereby providing a more complete account of the range of factors that contribute to the development of writing ability. A bioecological model offers a mechanism for holding these two perspectives in balance.

Situating an Integrated Construct in a Bioecological Model

We draw on Bronfenbrenner's bioecological model of development (Bronfenbrenner & Morris, 2006) to unite the sociocultural and cognitive sides of the field (Leu, Slomp, Zawilinski, & Corrigan, 2016; Slomp, 2012). According to Bronfenbrenner, neither the environment nor the individual solely drives human development. Rather, development occurs as unique individuals transact within their complex physical and social environments—what Bronfenbrenner describes as a series of nested, dynamic systems (i.e., from the microsystem level all the way to the macrosystem level). These systems act upon the individual, shaping, constraining, and supporting development. At the same time, unique aspects of the individual—their resources, dispositions, and demand characteristics—shape their transactions within this ecological system.

Key to positioning the writing construct within a bioecological model (Figure 3-1) is the concept of transaction. In various domains (Bronfenbrenner & Morris, 2006; Dewey & Bentley, 2008; Rosenblatt, 1988), transaction is a process within which individuals construct meaning as they engage with their social and physical environments. Rosenblatt (1938), for example, observed that when reading, meaning is made as individuals transact with text: Readers bring their purpose, experiences, and prior knowledge into the act of reading, and those purposes, experiences and prior knowledge shape how they make meaning of that text, in that moment. With respect to writing, transactions (purple rectangle, Figure 3-1) are always situated in two ways: through the rhetorical situation (including the writer's discourse community, purpose/intention, and ecological context; blue circle below) and through the bioecological context (including the writer's intrapersonal and ecological contexts; red circle). Like readers, writers construct meaning as they bring their purposes, experiences and prior knowledge both to the process of interpreting the rhetorical problem and to the process of composing text in response to that interpretation. Recognizing transaction as a central aspect of writing is key to unifying the field of writing research because it brings the cognitive dimension and the sociocultural dimensions into conversation with one another.

Figure 3-1. A bioecological model of writing



In the composing process, metacognition is the mechanism that drives transaction (purple rectangle, Figure 3-1). Through the conscious control and monitoring function, it drives the analysis of the rhetorical situation, it formulates the writing problem, it develops the plan to execute that solution, it monitors and guides the execution of that plan, and it evaluates the execution of the plan and the final product (Flower & Hayes, 1981; Hayes, 2012; Kellogg, 2008). Situated within transaction, metacognition is shaped and constrained by a host of factors. These factors can be broken into two sets: those internal to the writer (i.e., intrapersonal context) and those external (i.e., ecological context).

Intrapersonal context. The intrapersonal context of the writer—specifically demand characteristics, resources, and dispositions—also situates the transaction (Bronfenbrenner &

Morris, 2006). Knowledge related to the writing construct forms a small but essential element of a writer's intrapersonal context—a complex web of genetic and ecological factors that shape development and foster or constrain opportunities for success.

One aspect of intrapersonal context is demand characteristics, which invite or discourage reactions from the social environment and serve to either support or impede writing development. Subsumed under demand characteristics are personality (defined in terms of social stimulus value), demographic factors (e.g., age, gender, and ethnicity), and even physical appearance (in the sense that research has shown people's responses to others are often based on physical appearance; Bronfenbrenner & Morris, 2006). Although demographic factors are based on a person's various physical characteristics, they are included as they “place that person in a particular environmental niche that defines his or her position and role in society” (Bronfenbrenner & Morris, 2006, p. 814).

A second important aspect of intrapersonal context is a writer's resources, which refers to the writer's biopsychological liabilities and assets (Bronfenbrenner & Morris, 2006). For example, a resource liability such as a writer's poor fine motor skills can impede the transcription process (i.e., handwriting or keyboarding). Another example of a liability would be a concussive injury that causes impairment to a writer's short and long-term memory, which would constrain one's ability to enact the writing process. A writer's resource assets, conversely, would include the knowledge, abilities, skills, and experiences that evolve over a writer's development allowing a writer to form more dynamic, complex mental schema that facilitate easier adaptation to a variety of rhetorical situations. With respect to successful writing, knowledge about writing (the writing construct)—such as, critical discourse knowledge, discourse community knowledge, rhetorical knowledge, genre knowledge, writing process

knowledge, and subject matter knowledge—is a key resource that writers must draw on. We will expand on this shortly.

The final aspect under the umbrella of intrapersonal context is a writer's dispositions. Dispositions can be characterized as either developmentally generative or disruptive. Classroom teachers and writing instructors are all too familiar with disruptive dispositions, such as distractibility, impulsiveness, apathy, feelings of insecurity, and inability to delay gratification. Conversely, generative dispositions include curiosity, cooperativeness, the ability to delay immediate gratification to pursue long-term goals, and high self-efficacy (Bronfenbrenner & Morris, 2006). White, Elliot and Peckham (2015) identify five further dispositional characteristics as being generative for writing development: openness, conscientiousness, extraversion, agreeableness, and stability (p. 75). As an example of how a generative disposition can foster writing development, research has shown a positive correlation between writing self-efficacy and motivation and resiliency in writing (Bruning & Kauffman, 2016).

Ecological context. In addition to a writer's intrapersonal context, a writer is also situated within an ecological context that enables and constrains possibilities. Microsystems—or an individual's immediate context such as family, school, or work—have immediate and powerful influences on a writer and his or her capacity to write (complexities in family life, for example, can impede a writer's capacity to complete required tasks or experiment with new writing strategies). While the microsystem is the level closest to the writer, the other ecological system levels can also exert considerable influence in shaping a writer's capacity. Next, is the mesosystem, which is where an individual's microsystems interact, such as between a child's parents and teacher at school. The third level is the exosystem, which involves those contexts that have an indirect effect on the writer's life; these broader social, cultural, and economic

factors shape and constrain a writer's capacity. At the outermost level, the macrosystem, are the cultural ideologies and attitudes that indirectly affect the individual. The macrosystem is often described as a "societal blue print for a particular culture or subculture" (Bronfenbrenner, 1994, p. 40). Though society's norms and values around writing (e.g., free speech, funding for the arts) indirectly affect the writer, the effect can nonetheless be profound. Encapsulating these levels is the chronosystem examining how individuals interact with their ecological system over time. A writer's history is fundamental in shaping who that writer is and what that writer might become (Gutierrez, 2016).

Integration of Intrapersonal and Ecological Factors through Metacognitive Processes

Writing and the metacognitive processes that shape the creation of text is dependent on the resources a writer brings to the task (e.g., long-term and working memory, strategic knowledge, subject matter knowledge and writing domain knowledge such as process, rhetorical, discourse, and genre knowledge); the dispositions of the writer (e.g., motivation, sense of self-efficacy), and the writer's demand characteristics (i.e., the physical and psychological attributes that mediate their experiences with the physical and social environment). Writers draw on these attributes as they develop, enact, and monitor their plans for creating a text. Capacity with respect to these resources mediates metacognition. For example, as a writer's capacity with working memory, memory storage, and executive function increases, cognitive energy that can be directed toward metacognition opens up (MacArthur & Graham, 2016). Similarly a writer's level of motivation or sense of self-efficacy can influence the degree to which a writer actively engages metacognitive processes (Bruning & Kauffman, 2016).

For a text to be successful, the writer also needs to attend to the social and physical environment in which the text is being created (blue circle; Figure 3-1). The beliefs, values, and

expectations of the discourse community for whom the text is being written define the rhetorical problem and they shape and constrain the possibilities the writer can enact. Additionally, the physical environment—including the modality and location of publication—shapes and constrains a writer’s design choices. One’s ability to critically analyze these ecological factors is a key element of the problem-posing and monitoring aspects of metacognition in writing. First, the rhetorical situation the writer must navigate shapes possibilities for text production. The rhetorical situation is shaped by three elements: the discourse community, the ecological context, and the writer’s purpose or intention. The discourse community for whom the text is intended situates the transaction by defining the values, genres, content, and purposes relevant to the transaction. The ecological context situates the transaction both with respect to the physical environment (e.g., task environment, location and modality of publication/sharing, task materials, transcription technologies) and the social environment (e.g., educational environment, collaborators, critics, cultural milieu), and their capacity to enable and constrain the work of the writer. Collectively these factors shape possible courses of action available to the writer. The success of the writer is predicated on his or her ability to understand these factors, to account for them, and to respond to them appropriately. The writer’s purposes and intentions for that text also situate the transaction defining how the writer will work within or against the constraints imposed by the discourse community. Success, however, balances on a knife’s edge: misunderstanding any one facet of the rhetorical situation often results in a text not being successful in achieving a writer’s intentions.

Method

When structuring our review of the writing research we took a dialectic stance, respecting paradigmatic differences, juxtaposing these and the tensions and paradoxes they produce to

generate fresh insights and novel understandings (Greene, 2007). With the overarching purpose of developing a contemporary, integrated construct, we conducted a systematic review of the literature. A secondary purpose was to investigate how the construct of writing has evolved over the past five decades, for knowing where the construct has been and is currently situated helps us to understand where it might be headed. With these purposes in mind, we explored the following mixed research questions:

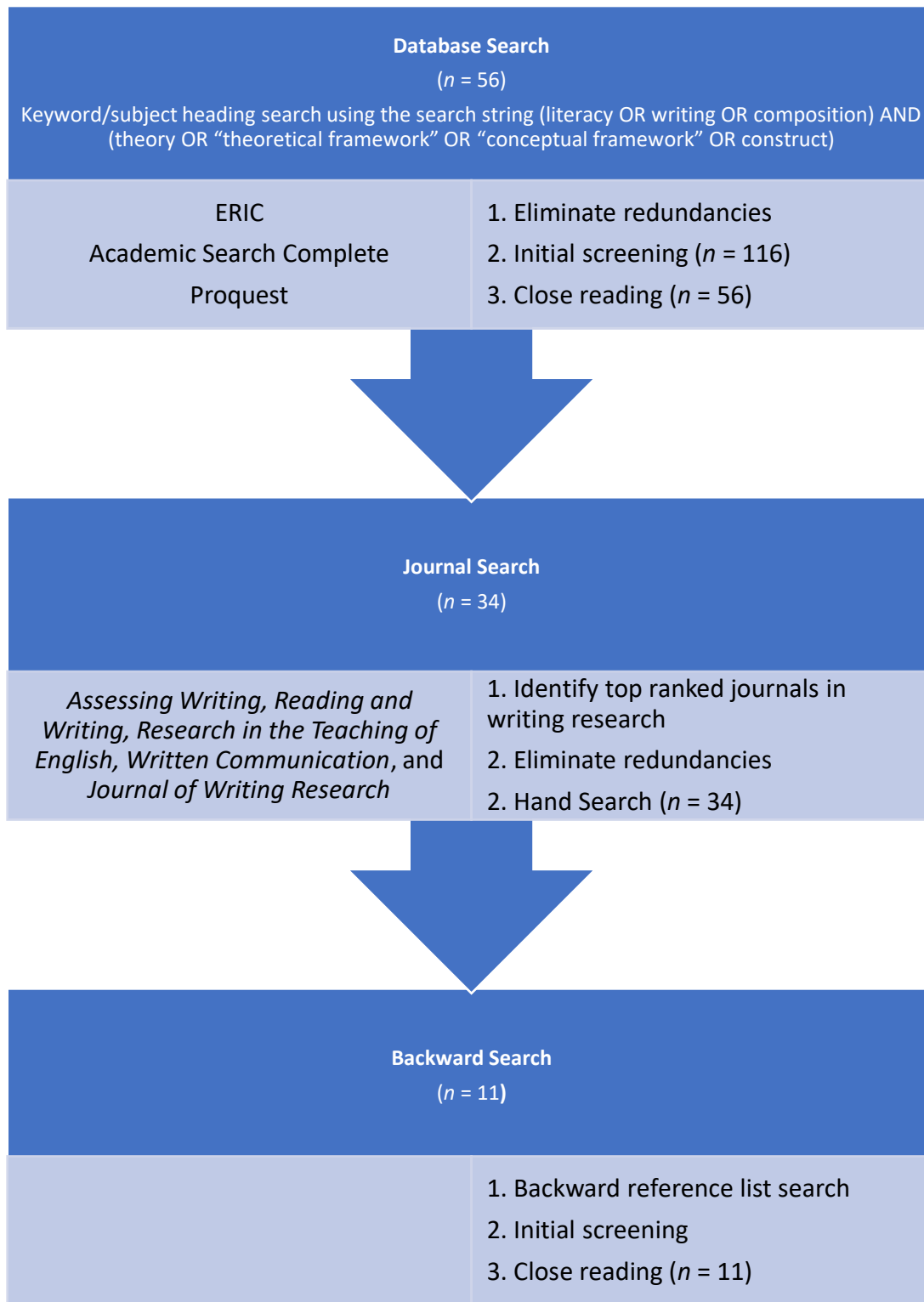
1. What writing knowledge domains are described in the literature over the last 50 years?
2. How do these domains coalesce to describe the construct of writing?
3. How has the emphasis on these domains changed over the decades? Specifically, is there a significant difference among the prevalence (i.e., coding densities, defined shortly) of domains from decade to decade (from the 1970s to the 2010s)?

Sources of Data

The data for this study was a corpus of studies on writing theory and constructs assembled through a search of two major sources: a database search and a journal search. A flowchart illustrating our process is included in Figure 3-2. We began with our database search delimited by the following inclusion criteria: the articles were to be peer reviewed, written in English, and had to be written between 1971 and the present (summer 2017). We chose 1971 as our starting date because that was the year during which Emig (1971) wrote her seminal work in which, for the first time ever, writing processes (as opposed to products) were examined (Voss, 1983). The following is a list of the search engines we consulted as well as the number of hits that met our inclusion criteria following an initial screening (i.e., reading of titles and abstracts) from these respective sources. Beginning with ERIC (i.e., Educational Resource Information

Center), we located $n = 81$ texts. From there, we searched Academic Search Complete and Proquest, further identifying $n = 12$ and $n = 9$ texts respectively, after eliminating redundancies. For our database search, we performed a keyword/subject heading search using the search string (literacy OR writing OR composition) AND (theory OR “theoretical framework” OR “conceptual framework” OR construct). In total, $n = 116$ texts were included after our initial screening, which was narrowed to $n = 56$ texts following a close reading of the texts. A number of the texts did not have an abstract, so a close reading was needed to determine if the texts were indeed relevant. Texts were eliminated following a close reading if they were not explicitly focussed on construct/theory development (e.g., if they were focussed on a classroom writing intervention, but made little contribution to construct development). Also, some were eliminated if they did not contribute anything new to theory/construct development due to saturation, operationalized as the point during which the collection of more data appears to have no additional interpretive worth (Sandelowski, 2008). For example, an empirical study using Hayes’ theoretical framework would not be included unless it added something new to this framework.

Figure 3-2. Flowchart of literature review process



Next, we conducted a second round of searching via a hand search of the top five ranked journals in writing research. This process included three phases. The first phase involved

identifying the top ranked journals in the field of writing research. To do this, we began with a search for all electronically available journals titles in the library database at the first author's institution. In the library's E-Journal search engine, we stipulated that words 'writing' or 'composition' or 'communication' needed to be in the journal's title. From this search, we included all journals which met the inclusion criteria of being peer reviewed, written in English, and focused on writing research or theory (thus, for example, practitioner journals or creative writing journals were excluded). Nineteen journals meeting these inclusion criteria were included at this juncture, and we added the journal *Research in the Teaching of English* because it is a journal known to focus on writing research that also carries a relatively high impact factor (Parsons & Gallagher, 2016). Next, we ranked those journals by CiteScore (Scopus®, 2016) and Impact Factor (*Journal Citation Reports*®, 2015) and selected the top five ranked journals for our hand search. Both rating procedures arrived at the same list of the top five journals, although their order differed (Table 3-2). Those journals included *Assessing Writing, Reading and Writing, Research in the Teaching of English, Written Communication, and Journal of Writing Research*. Phase two involved reading the abstracts of every article published by those journals between 1971 and the summer 2017. From there, we used the aforementioned inclusion criteria to narrow the results. This phase resulted in the discovery of $n = 34$ new articles after eliminating redundancies from the first round. In phase 3, we read the entire text of the articles and coded them. Finally, in addition to our database and journal searches, we conducted a backward (reference list) search wherein we added texts ($n = 11$) that that we deemed as making substantial contribution to the development of the construct of writing. After our database, journal, and backward search, we identified $n = 101$ (Table 3-3) texts for analyses. These

included mainly journal articles ($n = 92$), but also conference proceedings ($n = 2$), handbook chapters ($n = 3$), and one thesis.

Limitations

While we believe that our review of the literature was comprehensive, it was certainly not exhaustive. Particularly, we acknowledge the following biases that we employed in order to keep our review of the literature manageable: an emphasis on English, North American texts; an emphasis on more recent texts because we only reviewed texts from within the last 50 years; and, finally, we limited ourselves mainly to peer-reviewed journal articles thus ignoring other potential media formats (e.g., grey literature, websites, books).

Mixed Methods Analyses

Qualitative analyses. Once all relevant texts were identified, we used Nvivo and Microsoft Excel to complete three coding cycles. During the first coding cycle, two coders (the authors) read through all $n = 101$ texts and applied provisional codes, which refers to the process of beginning with researcher-generated codes based on preliminary research and revising, modifying, deleting, and expanding the codes as necessary (Miles, Huberman, & Saldana, 2014; Saldana, 2013). We began with a list of codes based on Beaufort's (2007) conceptual model derived from a longitudinal, ethnographic study because we believe this model to be one of the most robust and comprehensive models to date. Her model is comprised of five knowledge domains—discourse community, subject matter, rhetorical, genre, and writing process knowledge—all of which overlap and interact to describe the bodies of knowledge inherent to expert writers. In the model, the domain of discourse community knowledge encompasses the other four domains; this is one of the great strengths of the model, for it places emphasis on the situatedness of writing, accentuating the important role discourses play in shaping the domains

subsumed within. Each of the other four domains—subject matter, rhetorical, genre, and writing process knowledge—is not a static form of knowledge; rather, a successful writer will transform these knowledges in ways that are appropriate to a given discursive context. As Beaufort (2007) points out, the literature shows that expert writers not only have deep awareness of their discursive contexts, but they also possess mental schema, or heuristics, that they can draw upon in new contexts. Thus, Beaufort’s model does not set out to identify a model that is context specific, but rather mental schema that are transferable across writing contexts. While Beaufort’s model is certainly robust, our analysis of the literature suggested ways in which the model could be expanded. Thus, in first cycle coding, we used Beaufort’s five knowledge domains as our initial codes and expanded this list as new codes emerged. Our analysis revealed three additional codes (i.e., knowledge domains): metacognitive, critical discourse, and digital literacies knowledge. Provisional coding works well with exploratory studies such as ours (Saldaña, 2013) and aligns with our first research question (RQ 1) by aiding us in identifying the knowledge domains described in the literature.

Second cycle coding involved pattern coding, defined as developing a construct (or meta-code) that represents the corpus of codes generated in first cycle coding; furthermore, pattern coding serves to “attribute meaning to that organization” (Saldana, 2013, p. 209). This coding cycle aligns with our second research question (RQ 2) by describing how the domains coalesce to describe the construct of writing. Our analysis of the literature confirmed Beaufort’s (2007) model in which the domain of discourse community knowledge subsumes the domains of rhetorical, genre, writing process, and subject matter knowledge. We confirmed this by noting how discourse knowledge acted as a parent code (i.e., had a hierarchical relationship) to the previously mentioned codes. Pattern coding enabled us to extend Beaufort’s model by revealing

that two new domains—critical discourse knowledge and metacognitive knowledge—subsume Beaufort’s original model. We noted that critical discourse knowledge often acted as a grandparent code and metacognitive knowledge a great grandparent code (i.e., it subsumed all other codes). Meanwhile, digital literacies knowledge was frequently double coded with all other knowledge domains; thus, in our construct, digital literacies are inherent to every knowledge domain and not a domain of its own. In our results, we will describe the nature of the relationship among these codes. In lieu of calculating inter-rater reliability statistics, we opted for coder consensus wherein any coding discrepancies were debated and clarified until we mutually agreed on the appropriate use of the code.

Finally, in third cycle coding, we used operational diagramming to visualize and display the voluminous data generated from coding into a coherent and intelligible image (Miles, Huberman, & Saldana, 2014; Saldana, 2013). Using Nvivo, we analyzed the relationships and hierarchies that existed among the codes. We did this in order to create a figure that would illustrate the construct of writing.

Quantitative analyses. In addition to the qualitative analyses performed, we conducted a series of quantitative analyses that serve to answer research question three (RQ 3) regarding how the prevalence of the domains has evolved over the decades. To begin, we created an inter-respondent matrix (i.e., name of article x theme; Onwuegbuzie & Frels, 2016) wherein for every document, the coding density of each theme was recorded (Table 3-3). Nvivo automatically generates coding density but, since there is no native function in Nvivo to create an inter-respondent matrix, these data were imported to Microsoft Excel for further analysis. In Nvivo, coding density refers to the percentage of characters (i.e., each letter, number, or symbol represent one character) coded as text selections (i.e., highlighted for coding) in a document at

each node (i.e., theme, or in our case, knowledge domain). We used the data from the inter-respondent matrix to perform four analyses. First, we calculated the cumulative coding density scores of each text. Here, we refer to texts as journal articles mainly, but also conference proceedings, book chapters, and theses. To calculate the cumulative coding density (similar to what Onwuegbuzie refers to as qualitative effect sizes; Onwuegbuzie, 2003) we calculated the sum of the coding densities from each theme for every text. Essentially, the more densely a text was coded, the higher its cumulative coding density. Cumulative coding densities reveal information such as which articles were coded most densely and thus contributed more to our conceptualization of the construct. Second, to determine which domains were most heavily coded, we calculated the frequency of codes across all texts; whether the domain was coded only once or numerous times within a text, a score of one was allotted if the theme was mentioned in that text. Third, we calculated the percentage of texts coded by knowledge domain per decade. This helped us to see temporal patterns in the way the construct of writing was written and researched about over time. Finally, we ran multiple, one-way between groups analysis of variance (ANOVA) with post-hoc tests in order to see if there was a statistically significant difference among the coding densities of domains from decade to decade (from the 1970s to the 2010s).

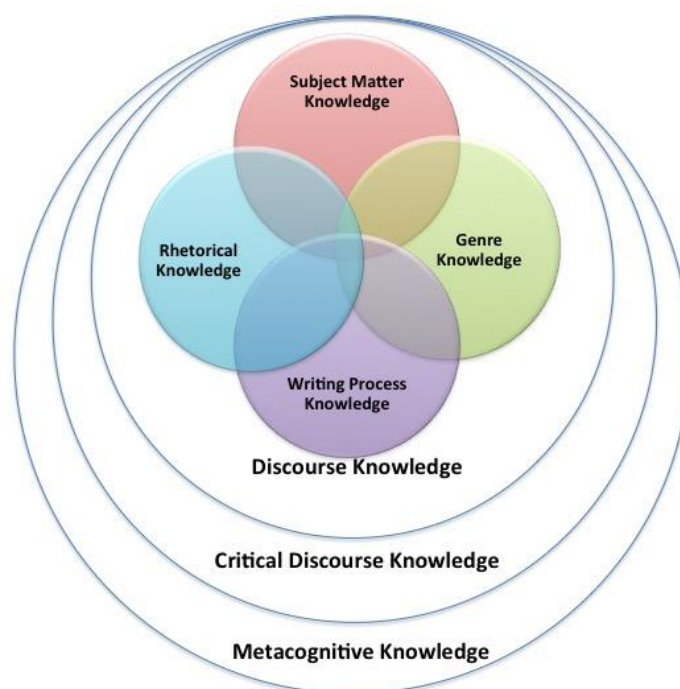
Results

Qualitative Results

Here we examine the results of the first two research questions wherein we identify the knowledge domains described in the literature (RQ 1) and describe how these domains coalesce to describe the construct of writing (RQ 2). The knowledge domains described in this section constitute an important set of resources necessary for the development of expertise in writing,

which we referred to in our bioecological model (Figure 3-1). We begin by synthesizing the data that correspond to Beaufort's (2007) knowledge domain of discourse community knowledge. Then, we show how our review of the literature revealed three new domains—critical discourse knowledge, metacognitive knowledge, and digital literacies knowledge—that subsume and extend Beaufort's original model. Given that digital literacies knowledge was interconnected with the other knowledge domains, we have chosen to embed this knowledge throughout each of the domains rather than making it a domain of its own. Next, we discuss how Beaufort's domains of rhetorical, genre, writing process, and subject matter knowledge fit within the construct. Taken together, these knowledge domains describe the integrated writing construct (Figure 3-3).

Figure 3-3. Integrated writing construct



Discourse knowledge. We define the domain of discourse knowledge as knowing the socially accepted association among ways of thinking, feeling, believing, valuing, and using

language (Beaufort, 2007; Gee, 1990) or signs (Kress, 2003; Kress & van Leeuwen, 1996; Witte, 1992) that can be used to identify oneself as a member of a social network. All humans have the biological endowment to acquire language, which begins in early childhood (Gee, 2014). This socially accepted language use among family and the immediate community constitutes one's vernacular, and the acquisition of all future non-vernacular social languages (e.g., the language of a nuclear physics, bird watching, or even that of a street gang) builds upon this foundational vernacular grammar. According to Gee (2014), learning a non-vernacular social language is a sociocultural process (we will argue later that it is also a socio-cognitive process) that involves recognizing new ways to pattern together one's lexical and grammatical resources to match the norms and values upheld by that discourse community. Thus, learning language is much more than learning English or any specific language; rather, it is learning to speak and write in ways that are appropriate for the discourse community within which one is communicating. Discourse communities can similarly be classified as primary versus secondary. Primary discourses are vernacular social languages (i.e., those learned through family and community) while secondary discourses are those learned "in order to interact with groups outside of our immediate community," which are typically institutional discourses, or "ones that [are learned] at school and other institutions" (Wood, 2002, p. 4).

Expanding on the first part of our definition, we take Gee's (1989) position that knowing the grammar of a language is not the same as knowing how to use that language. In other words, according to Gee, "[i]t is not just what you say, but how you say it" (Gee, 1989, p. 5). For example, while it would be appropriate to use proper grammar and esoteric jargon in a job interview for a position as an investment banker, it would be inappropriate to do so while meeting an old friend at the local pub. While the language is correct in the latter, it is

inappropriate for the situation. Moreover, as Gee (1989) has argued, Discourse⁵ is about more than language use; rather, it is the saying-writing-doing-believing-valuing combinations that matter. For example, an appropriate saying-doing combination would be to wear a designer suit to the above-mentioned interview while this same saying-doing combination would make a person look like an outsider at the local pub. Similarly, Beaufort (2007) describes college students who struggle to adapt to the ever-changing discourse communities presented to them as they transition among disciplines, professors, and from high school to disciplinary to professional discourses as they graduate and enter into the world of work. The writing style that got a student an 'A' in freshman writing simply does not work while writing a report in an engineering class on thermodynamics nor after graduation in an engineering firm. While we blend Gee's and Beaufort's conceptualizations of Discourse and discourse community into our own definition of discourse knowledge, we broaden this concept to include not only the appropriate use of language and its configurations (of writing-saying-doing-believing-valuing), but also the use of signs (Kress, 2003; Kress & van Leeuwen, 1996). In an age where writing has predominantly moved from the page to screen, the affordances of text go beyond linguistic design to include multimodal sign making via visual, audio, gestural, and spatial design features (New London Group, 1996).

⁵ Gee (1990) differentiates between upper-case Discourse, and lower-case discourse. The former refers to the saying-writing-doing-believing-valuing combinations mentioned above. The latter Gee describes as language-in-use, providing examples such as conversations, stories, reports, arguments, and essays; the latter we describe as genre knowledge, which will be discussed in the ensuing pages.

The second part of our definition describes identifying oneself as a member of a socially meaningful group. These groups take many forms such as disciplinary (Lea & Street, 2006), professional (Beaufort, 2007), academic (Ivanic, 2004), linguistic and/or cultural (New London Group, 1996), local (New London Group, 1996), online versus offline (Lankshear & Knobel, 2008), and the hybridity or heteroglossia of discourses ensuing from the combinations thereof (Ivanic, 2004; New London Group, 1996). It is not uncommon for a writer/speaker to simultaneously represent multiple discourse communities, which can either meet with resistance or in transformed practice or both; for example, Jesse Jackson mixed the discourse of politics with the discourse of African American religion in order to transform political discourse (Brady, 1993; New London Group, 1996).

Each discourse community possesses multiple layers, and gaining access to one layer does not guarantee access to the next. Those who have access to the innermost layer of a discourse community have insider status while those on the outer layers are colonized or subjugated and serve to reinforce the dominance of those on the inside (Berlin, 1992; Van Heertum & Share, 2006; Wood, 2002). The colonized are the language users who have just enough access to a discourse community to signal that they are not full members (Wood, 2002). Discourse knowledge is important because by signalling group affiliation, people gain the social capital necessary to be active participants in their personal, professional, and civic lives.

Critical discourse knowledge. We define the domain of critical discourse knowledge as knowing how writing is socially and historically constructed (Berlin, 1992; Ivanic, 2004) within specific power relations (Luke, 2012; Street, 2003), and knowledge of how writing reinforces discourses of both dominance and marginalization (Berlin, 1992). Because discourses are both social and historical constructions, what constitutes their correct usage and conventions is

completely arbitrary. This is not to say that convention is bad; for example, *Review of Educational Research* uses APA style, which signals to writers the norms and values considered appropriate by the community of educational researchers (e.g., how to cite sources or use grammar). As such, the *Publication Manual of the American Psychological Association* is a professionalization vehicle that assists new scholars in writing themselves into the field. However, as Canagarajah (1996) argued, (predominantly Western) scholars need to be cognizant of how the norms of academic publishing can resemble “ideological imposition and cultural hegemony” (p. 436) for periphery scholars.

With critical discourse knowledge, the writer recognizes the fluidity of language and that no one particular group should be able to discern and prescribe correct usage over another (New London Group, 1996). The writer also recognizes that when hegemonic views regarding discourse use become entrenched, “asymmetries of power, access, and opportunity along the lines of gender, race, class, and sexuality” are maintained (Van Heertum & Share, 2006, p. 263).

With regards to writing being historically constructed, Bakhtin (1981) and his peers emphasized that “a text is shaped by discourses or voices that exist prior to it” (Brady, 1993). This can be seen in the way certain discourse communities have been systematically silenced and marginalized throughout history, at least until something happens on a revolutionary scale (e.g., the Civil Rights Movement) that challenges the hegemonic thinking that marginalized those voices in the first place. To be able to situate a discourse chronologically (i.e., understand its history) is consequential because understanding the root causes behind a discourse’s marginalization is an essential step in producing counter-hegemonic texts (Van Heertum & Share, 2006), for it is important to understand the past in order to challenge the present.

Implicit in the notion that writing reinforces discourses of both dominance and marginalization (Berlin, 1992) is knowing that certain discourses tend to be privileged over others—such as monomodal versus multimodal (Hull & Nelson, 2005; Jewitt, 2008; Mills, 2013); the lingua franca (e.g., English) over other (e.g., indigenous) languages (Battiste, 1998; Lyons, 2000); and standard registers over local dialects (New London Group, 1996)—particularly in the institutionalized settings of secondary discourses (e.g., legal, corporate, and higher education). With critical discourse knowledge, the writer recognizes the wealth and diversity created by a heteroglossia of voices, languages, and modalities, particularly in an increasingly globalized and digitized society. Mills' (2010) review of the digital turn in *New Literacy Studies*, for example, highlights the changing emphasis from “print-based reading and writing practices to include new textual practices that are mediated by digital technologies” (p. 247). With the shift from page to pixel, the multimodal affordances of texts have increasingly expanded. Still, many institutional discourses rely on monomodal texts (e.g., legal documents, equity policies, fiscal documents, accountability processes) where the image is conspicuously absent thereby establishing “positions and ideologies that constrain readers not to view such discourse as either trivial or open to contestation” (Mills, 2013, p. 9). Additionally, privileging print-centric over multimodal representations limits the availability of sign systems for English language learners and/or low-literacy groups to translate. Concerning the privileging of one language over another, Battiste (1998) maintains that languages “provide a direct and powerful means of understanding the legacy of tribal knowledge” (p. 18). In addition to the privileging of a dominant language, Indigenous writers also face a privileging of modalities, with print-centric curricula serving as another colonial instrument to deprive Indigenous Peoples of their oral traditions (Battiste, 1998). Lyons (2000) cautions against oral-print binaries (as though there

exist no prolific Indigenous prose and verse), instead encouraging rhetorical sovereignty, the “inherent right and ability of peoples to determine their own communicative needs [be it oral or print] and desires in the pursuit of self-determination” (p. 462). Thus, the domain of critical discourse knowledge involves the ability of writers to discern the most appropriate means to convey their text—in the modality, language, and register of their choosing.

The major difference between discourse community knowledge and critical discourse community knowledge is being able to make visible and respond to the inequities among discourse communities that have been rendered invisible: “Whether the literacies user is aware or not, every literate act is either reinforcing or dismantling the hegemony present within the society in which it is produced and interpreted” (Wood, 2002). In other words, critical discourse knowledge is the process of seeing beyond the “hegemonic sheen” (Van Heertum & Share, 2006, p. 258), recognizing and understanding one’s own experiences of being Othered and also those of other groups. This is no easy task for the writer because accepted standards of the hegemonic group’s language are monolithic, ingrained, and rarely challenged. Unlike discourse knowledge, critical discourse knowledge goes a step further challenging the writer to understand the “appropriate and effective uses of literacy as more complex, dynamic, nuanced, situated, and involving both epistemological issues and social processes, including power relations among people, institutions, and social identities” (Lea & Street, 2006, p. 369).

Metacognitive knowledge. As originally coined by Flavell (1979), metacognition refers to the knowledge about and regulation of one’s cognitive activities in learning processes (Veenman, Hout-Wolters, & Afflerbach, 2006). From a Vygotskian perspective, metacognition is understood as inner speech—the internal monologue that shapes intelligence and guides performance. While cognition and metacognition are closely linked, much of the construct

model we have thus far developed describes the cognitive domains that contribute to excellence in writing. Metacognition is understood to be a higher-order function, and our analysis has revealed that metacognition in writing encompasses analysis of the rhetorical situation (Flower & Hayes, 1980; Huot, 1990), problem posing (Bereiter, Burtis, & Scardamalia, 1988; Rice, 2015), planning (De La Paz, S., & Graham, 1997; Flower & Hayes, 1981; Stotsky, 1990), execution of the plan, and an evaluation of the plan and its execution (Hayes, 2012; Hayes & Flower, 1980), which we will expound on later while discussing our bioecological model. Not all of these processes are evident in every writing situation nor do they always look the same: for example, revision may not be used when sending a text message to a friend, nor would there necessarily be much planning, at least not in a formal sense (Hayes, 2012). Research comparing novice and expert writers has demonstrated that metacognition is foundational to the development of expertise. Expert writers analyze the rhetorical situation they are writing in, they define the writing problem that needs to be solved, and they establish long-term and short-term goals related to solving this problem. They are more cognitively aware of what they are writing, more active in planning and monitoring their thinking and progress as they write, and more likely to evaluate their writing than are novice writers (Stolarek, 1994). Similarly, another study found that weaker writers demonstrated relative weakness with writing self-regulation and metacognition (Wakely, Hooper, de Kruif, & Swartz, 2006).

Rhetorical knowledge. We define rhetorical knowledge as knowing the rhetorical situation (i.e., the needs of a specific audience and the specific purpose[s] for a text; Beaufort, 2007; Flower & Hayes, 1981; Ronald & Volkmer, 2015), and knowing how to address those needs in the appropriate discursive context (Brady, 1993; Mailloux, 1989). Traditionally, rhetoric has been defined as the art of persuasion (Rapp, 2010), though as Moore (1997) has

argued, not all texts serve this purpose. Technical writing, for example, generally serves the purpose of describing how to perform a procedure or task (Moore, 1997). Thus, our definition goes beyond persuasion to include wide-ranging purposes for a given text (e.g., to describe, entertain, motivate, sell, narrate, argue, sympathize, advocate, proselytize), also acknowledging that any given text may serve a multiplicity of purposes.

In addition to knowing the purpose(s) of a text, knowledge of the rhetorical situation importantly includes knowing the needs of a specific audience. This explains why we follow Beaufort's (2007) precedent of embedding rhetorical knowledge within the domain of discourse knowledge. If writers wish a text to serve a particular purpose, they had better be aware of the needs of their audience (i.e., discourse community). As Mitchell and Taylor (1979) aptly remarked, the importance of rhetorical awareness changes the dictum from “know your grammar” to “know your audience,” for “[c]orrectness has its place only when the audience you are addressing values it” (p. 264). While English teachers rightly value grammar, other discourse communities value a range of ideas from the scientific process to the lyrical quality of a song to the balanced reporting of a current event.

Finally, our definition includes addressing the rhetorical situation by selecting the appropriate context. Even within a discourse community, there exist a variety of contexts. Take for example academia where an academic might engage in rhetorical contexts such as monomodal (e.g., writing a journal article) vs. multimodal (e.g., communicating research via a blog); formal (e.g., writing a white paper) vs. informal (e.g., discussing research with a colleague); or disciplinary (e.g., presenting a symposium at an academic conference) vs. general audience (e.g., communicating with stakeholders).

Traditionally, rhetorical knowledge has been taught and assessed in relation to monomodal texts, as an element of linguistic design (New London Group, 1996), which is complex in its own right. One contested aspect of rhetorical knowledge is that of traditional grammar because several robust meta-analyses have found no evidence to support the role of grammar instruction (Hudson, 2016; Jones, Myhill, Bailey, & Bailey, 2013). However, as Jones et al. (2013) have argued, previous studies into the efficacy of grammar instruction have investigated grammar taught in a decontextualized context (e.g., the ubiquitous grammar work sheets). When taught with a view that grammar is a tool to achieving a writer's rhetorical intention, by contrast, grammar instruction was shown to be effective (Jones et al., 2013). In other words, research has shown descriptive grammar instruction (i.e., teaching how language can be used to achieve a rhetorical goal) to be more effective than prescriptive approaches (i.e., teaching rules and correctness). At the same time as grammar and its role in the curriculum are being reconsidered, so too are other aspects of linguistic competence such as spelling and vocabulary (Schlagal, 2013). Perhaps, like grammatical instruction, teaching spelling and vocabulary has been ineffective in the past because it has neither considered the needs of the individual learner (Schlagal, 2013) nor the connection to rhetorical knowledge.

We see grammar as something beyond the correct usage of language and instead as a series of rhetorical choices to identify and select from the representational resources that are available “the various meanings these resources will have if drawn upon in a particular context, and the innovative potential for reshaping these resources in relation to social intentions or aims” (New London Group, 1996, p. 79). The New London Group (1996) uses two sentences to illustrate this point: “Lung cancer death rates are clearly associated with increased smoking” and “Smoking causes cancer” (p. 79). Depending on the discourse community reading these

sentences, what they connote and denote could vary considerably. Connotatively, the general public would read the latter sentence as making a stronger statement, while the former sentence seems more tentative. Denotatively, while the media and the general public would likely read these sentences as having the same meaning, statisticians understand that correlation does not equal causation. By contrast, statisticians would view the former sentence as an investigation of the size and relationship between two or more variables; the latter sentence, on the other hand, would point to a controlled study wherein a sample is split in two, with both groups being comparable in as much as possible—with the exception of the independent variable (i.e., smoker vs. non-smoker status). Thus, these sentences illustrate rhetorical knowledge by means of how grammar has been recruited to design different purposes for different discourse communities.

As if rhetorical knowledge were not complex enough when considering monomodal texts, another layer of complexity is added when considering the design of multimodal texts (Kress, 2003; New London Group, 1996). Expanding on the conceptualization of grammar hitherto mentioned, we draw on Kress' (2003) use of the term grammar as an “overarching term that can describe the regularities of a particular mode which culture has produced, be it writing, image, gesture, music or others” (p. 66). Knowledge of these new grammars is becoming increasingly important in a digital world. Where once only alphabetic text could be (somewhat easily) reproduced, now text, sound, image, and moving image are translated into 1s and 0s and disseminated around the world by anyone with a smart phone with one click. As Kress (2003) argued, each mode has distinct affordances and their potential for representing are only partial; thus, rhetorical knowledge includes the selection of the appropriate mode(s) through which to most appropriately and completely represent the writer’s intentions. If McLuhan's (1964) dictum

that “the medium is the message” (p. 7) is correct, then the act of writing and the technologies inherent to that mode are inseparable.

Error! Reference source not found. shows how rhetorical knowledge is embedded not only within discourse knowledge, but within critical discourse knowledge and finally metacognitive knowledge as well. Building pedagogy around critical rhetorical knowledge would follow a model whereby students would be presented with a gradual yet deliberate unfolding of texts revealing multiple and sometimes incongruous ideas so as to become conscientious of the “subtle and often dramatic rhetorical choices made by writers, which in turn affect them as readers” (Salibrici, 1999, p. 632). Once writers are able to deconstruct a text in order to understand the subtle ways it works on its readers, they may then appropriate those rhetorical strategies into constructing texts of their own to produce the same effect. The empirical evidence demonstrates that when students improve writing, reading follows; the reverse is also true (Graham & Harris, 2013). There is a reciprocity in using rhetorical knowledge critically, where students learn to view language both in terms of production and consumption: “On the one hand, rhetoric allows us to acknowledge how we as writers manage to develop a text, and on the other hand, rhetoric allows us to acknowledge how a text can then be understood and believed or not by its readers (Salibrici, 1999, p. 629). As Salibrici (1999) argues, increasing rhetorical awareness leads to greater critical thinking.

With critical rhetorical knowledge, writers go beyond radically simplifying the rhetorical situation (e.g., write another theme for English class; Flower & Hayes, 1981) to looking in, around, over, under, and through the problem in order produce critical insights for themselves and their audience. It is not enough to understand the rhetorical choices as a matter of the writer’s individual style or intention, but more so, as “inherently connected to different

discourses with their wider interests and relationships of power” (New London Group, 1996, p. 79).

Finally, rhetorical knowledge is subsumed under metacognitive knowledge (see **Error! Reference source not found.**). This metacognitive aspect of rhetorical knowledge is, according to research, partly what separates expert from novice writers (Flower & Hayes, 1981). While expert writers have developed heuristics over time that enable them to juggle many cognitive demands (Beaufort, 2007), including solving the rhetorical problem (Flower & Hayes, 1981), novice writers frequently reduce the complexity of the rhetorical situation by oversimplifying it. According to Flower and Hayes (1981), writers “only solve the problems they define for themselves. If a writer’s representation of her rhetorical problem is inaccurate or simply underdeveloped, then she is unlikely to ‘solve’ or attend to the missing aspects of the problem” (p. 369). Therefore, while novice writers oversimplify the rhetorical problem addressing it only partly or perhaps inaccurately, expert writers not only apply prior knowledge to solving the problem, they develop and refine the rhetorical problem to meet their own motivations for writing (Flower & Hayes, 1981) and those of their audience, while learning how to transfer their problem-solving strategies to contexts near and far (Brent, 2011).

Genre knowledge. Simply put, genre knowledge relates to one’s understanding of genres and how they function. As such, genre knowledge can be defined as understanding textual forms (e.g., how a website differs from a blog), organizational structures of texts (e.g., that websites have navigational bars), and both prescriptive and descriptive textual conventions (e.g., that websites use hyperlinks to link to additional information). More importantly, genre knowledge also involves an understanding of the purposes and social contexts that gave rise to these conventions, organizational structures, and textual forms: “As researchers from Miller

(1984) to Spinuzzi (2004) have demonstrated, genre mediates intention, exigency, and context, thus providing a way to understand organizational structure and professional behaviour” (Coppola & Elliot, 2013, p. 276). In this way, like the other domains of knowledge described in this construct model, genre knowledge is integrated with, and dependent upon, other domains of knowledge.

Understanding that genres reflect purposes of writing ties this domain of knowledge to that of rhetorical knowledge: “Genre can be said to represent typified rhetorical action (Miller, 1984, p. 151). As writers define their intentions for a text they are composing—whether that be narrating, describing, expounding, arguing, or evoking—they draw on rhetorical moves or structures that enable them to deliver on these intentions. Historically, these rhetorical structures often became linked to the genres themselves. This fossilizing notion of genre is problematic because it tends not to recognize that in addition to serving specific purposes, genres are designed to function within specific social contexts. Modes and structures of arguments that are effective in one social context, for example, may be highly problematic in others.

Success in writing depends on understanding that “genres are socially real and that to participate effectively in a discourse community one must adapt to (or around) readers’ generic expectations” (Salibrici, 1999, p. 165). Because they are “socially real”, genres are also fluid, evolving over time and changing across social contexts. Genre knowledge requires that writers understand this fluid contextual nature of textual structures, organizational features, and conventions. In this way, genre knowledge is tied to and dependent upon discourse community knowledge because it involves writers’ ability to understand the values and expectations of the discourse community, how those values and expectations are expressed in textual structures, features, and conventions, and of how to shape their texts in light of those understandings. We

emphasize here the transactional nature of our construct: meaning is constructed in the interaction between the writer's intentions and the context within which those intentions are enacted. In this respect writers are not simply free to choose to enact genre features as if they were neutral entities; rather each of these features themselves are socially situated and structured, their utility for the writer are also ecologically determined. This does not negate a writer's agency or intentionality, but rather positions it within an ecological context that both enables and constrains possibilities. It is through this process of transaction that genres evolve—we see this with respect to hybridity, intertextuality, and multimodality of text—because all are born of the transaction between a writer and the ecological contexts from within which and to which he or she is writing.

Writing process knowledge. Writing process knowledge involves knowing how to get discourse specific writing tasks accomplished, including meta-knowledge of the cognitive processes in composing and the practical phases of writing projects (Beaufort, 2007; Flower & Hayes, 1981). In other words, writing process knowledge can be understood as having both cognitive and practical dimensions (Stotsky, 1990).

In terms of cognitive writing process knowledge, there is an expanding body of researchers whose empirical work has revealed the complex mental processes involved in writing. This research began with the seminal work of Emig (1971) who documented the composing processes of twelfth graders, which marked a dramatic shift in writing research as it moved from product to process. This continued with the work of Flower and Hayes (1981) who developed a cognitive process theory of writing, based on a five-year study of protocol analysis. Hayes' work developing a cognitive model of writing continued over the decades (Chenoweth &

Hayes, 2001; Hayes, 1996, 2012; Leijten et al., 2014) expanding upon his earlier work with Flower.

Hayes' and colleagues' current writing model (Leijten et al., 2014) is divided into three levels: resource level, process level, and the control level. The resource level includes attention, motivation management, reading (i.e., the text produced so far or secondary sources), long-term memory, and working memory. This level describes the resources that the writer can draw upon in the other levels, including the process level and control level. For example, a writer can call upon long-term memory to recall grammatical rules needed by the translator in the process level. The process level consists of internal processes (i.e., searcher, proposer, translator, transcriber, and evaluator) and the external environment of those processes (i.e., collaborators and critics, production technology, task-related-sources/written plans, and text-and-graphics-created-so-far). The searcher looks for information in external sources, such as finding the right word in a dictionary or researching information from secondary sources (Leijten et al., 2014). Chenoweth and Hayes (2001) describe the proposer as the prelinguistic production of ideas. The translator then converts the prelinguistic ideas “into strings of language with appropriate word order and grammar” (Chenoweth & Hayes, 2001, p. 84). From here, the reviser evaluates what has been and is being written, while the transcriber turns these expressions of the mind into written language. At the control level, Hayes and his colleagues (Leijten et al., 2014) place motivation management, goal setting (for planning, composing, and revising), and the current plan/writing schemas/design schemas. Hayes (2012) noted that while schemas are represented at the control level, they are presumably stored in long-term memory. These writing schemas, which help writers accomplish specific writing or production tasks, are “are assumed to be modifiable by experience and instruction and to constitute an important part of writing skill” (p. 375).

Hayes and Flower (1980) identified several underlying cognitive activities in writing, which were subsequently analyzed by Breetvelt, van den Bergh, and Rjlaarsdam (1994) who concluded that the following activities explained 76% of the variance in the quality of papers produced by adolescent writers: reading the external sources (assignment, documentation), planning text (goal setting, generating, structuring), text production/translating, text reprocessing (rereading already written text, evaluating already written text, revising already written text), and monitoring (self-instructions, meta-comments, pausing). Another important finding of the Breetvelt et al. (1994) study was that the relationship between the cognitive activity and text quality was dependent upon the stage in the writing process where it occurred. For example, a writer who revises frequently at the beginning of the writing process might be experiencing start-up difficulties (i.e., beginning over, and over again) while another writer who revises just as frequently, but towards the end of the process, is revising for a very different purpose (van den Bergh et al., 2016).

While there has been considerable agreement regarding the underlying cognitive processes involved in traditional writing, more recent studies on writing in digital environments and those in professional contexts have revealed that the writing process needs to be expanded to include a greater variety of subprocesses. Two examples of these subprocesses include locating online information and collaborative writing, subprocesses that have been greatly complicated, yet also facilitated, by the Internet. A case study of professional writing revealed that locating online information is a crucial component of the writing process (Leijten & Waes, 2000). Those who can leverage Boolean operators (and, or, not), proximity searching strategies (within, near, before, after), and wildcard symbols (*, ?) are going to have greater proficiency using search engines and academic databases, particularly in a digital age where information is seemingly

infinite. Having proficient search strategies helps writers identify multiple sources of information for both content and ideas (Leijten & Waes, 2000) but goes beyond that to enable writers to redesign or remix content from multiple sources (Lankshear & Knobel, 2007; Leijten et al., 2014; New London Group, 1996). Another example of the way in which the construct of writing process could be expanded is to include collaborative writing (Marttunen & Laurinen, 2012). While collaborative writing has been around for some time, new communication technologies (e.g., Google Docs, workstreaming, online project management apps) have greatly facilitated this process. Though there is a paucity of research regarding how cognitive writing process knowledge has been affected by digital technologies, one can speculate that the cognitive load required of locating, critically evaluating, and managing the massive amount of information available in a digital age would certainly exert an influence over the writing process. Further, writers require new mental schemas in order to adapt to the new grammars demanded of ever-evolving communication technologies.

Writing process knowledge has been generally conceived as linear—despite insistence from writing researchers that this is not the case (Ronald & Volkmer, 1989). In fact, writers—particularly expert ones—continually establish different configurations of subprocesses with respect to the writing context, notably the topic, genre, and task complexity (van Weijen, 2009). Just as writing skill varies widely across contexts, so too does writing process knowledge; thus, van den Bergh and colleagues (2016) argued that we cannot rely on a single measurement of writing process in only one context should we wish to obtain a valid measurement of a student's writing process knowledge.

Included in writing process knowledge are the practical, iterative processes of “planning (setting goals, generating ideas, organizing ideas), translating (putting a writing plan into action),

and reviewing (evaluating, editing, revising)” Graham & Sandmel, 2011, p. 396). Early research into the effect of practical writing process knowledge on writing quality was discouraging. For example, Emig's (1971) study of twelfth graders showed no correlation between the subprocess of creating an outline and the grade a student receives with respect to an essay's organization. This may be explained by the context within which the students were writing: a timed, impromptu essay of approximately five paragraphs can easily be mentally organized without the need for a written outline, thus allocating more time for the actual writing of the essay itself. More recently, however, practical writing process knowledge has been shown to be positively correlated with writing quality. A meta-analysis of 29 experimental and quasi-experimental studies of the process approach to writing instruction revealed that this form of intervention was correlated with positive, albeit modest, gains to writing quality (Graham & Sandmel, 2011).

Subject matter knowledge. Knowledge of subject matter is an essential domain of the writing construct. After all, whenever we write, we write about something. Halliday (2003) described this domain as “the interpretation and expression in language of the different types of process of the external world, including material, mental, and abstract processes of every kind” (p. 314). This definition suggests that subject matter knowledge is not simply content knowledge, but more significantly an understanding both of how the world works and of the relationships between bodies of knowledge. It also involves the capacity to connect pieces of information for rhetorical effect.

In this way, subject matter knowledge does not function independently from other knowledge domains; it is linked to rhetorical knowledge and to many other facets of this construct. For example, ideational content, what we know about a subject, may structure the kind of argument we make, or the types of syntactic structures required to build that argument.

Thus, subject matter knowledge is linked to both rhetorical knowledge and genre knowledge. Of course, we often are expected to write about topics that we know very little about. In this respect metacognitive knowledge and writing process knowledge are closely linked to subject matter knowledge because they determine how one acquires, organizes, and utilizes content.

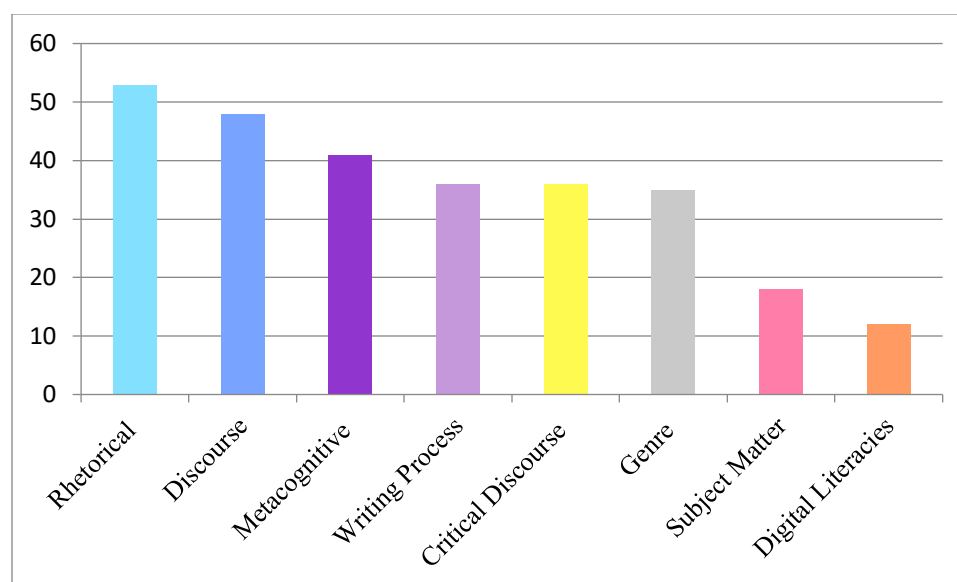
An important dimension of subject matter knowledge is its link to discourse community knowledge. While having some understanding of the topic one is writing about is important, it is just as important to know how the discourse community for whom one is writing views that subject matter. A paper on global warming for an environmental magazine would need to approach the topic of climate science differently than would a paper on the same topic for an oil industry magazine. The content may remain the same, but how one introduces or positions that content would very much be influenced by audience and rhetorical intent. The importance of both subject matter and discourse knowledge is empirically borne out in the literature: for example, for informative, narrative, and persuasive texts, between 30% and 43% of the variance in holistic scores was explained by topic knowledge in combination with discourse knowledge in fifth-grade students (Olinghouse, Graham, & Gillespie, 2015). Other studies showed that subject matter knowledge was correlated with longer, more organized, more rhetorically aware, and better argued writing (Proske & Kapp, 2013).

Quantitative Results

We end by revealing the results of RQ 3, which help us to understand how the texts and domains variously informed our construct development, and how these domains have prevailed over the last five decades. Our first quantitative analysis entailed the calculation of cumulative coding density scores revealing which articles overall and which knowledge domains more specifically were coded most heavily (Table 3-3).

Next, we calculated the number of texts coded per knowledge domain. The domain that was coded most frequently was rhetorical knowledge ($n = 53$ texts), followed by discourse ($n = 48$) and metacognitive knowledge ($n = 41$). The knowledge domain that was least coded was digital literacies knowledge ($n = 12$ texts). Figure 3-4 presents a frequency histogram of these results.

Figure 3-4. Number of articles coded per writing knowledge domain

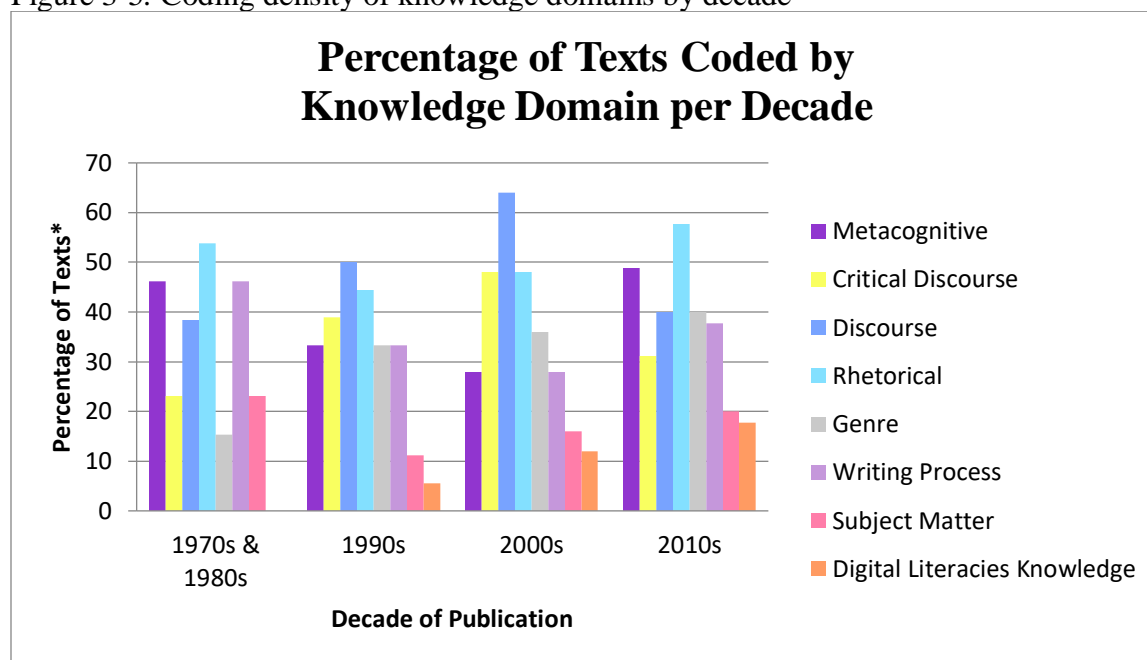


Third, we calculated the percentage of texts coded by knowledge domain per decade. In the 1970s and 1980s category (combined as there were only two texts in the 1970s category) 46% of texts were coded for metacognitive knowledge; this percentage diminished each decade until it rose to 49% in the current decade. Another domain that diminished (46 – 28%) and subsequently rose (38%) in the current decade was writing process knowledge. The domains of critical discourse (23 – 48%) and discourse knowledge (39 – 64%) had been rising until they

dropped in the current decade (31% and 40% respectively). The domains of genre (15 – 40%) and digital literacies knowledge (0 – 18%) have been increasing with each decade.

Figure 3-5 contains a side-by-side bar graph illustrating these trends.

Figure 3-5. Coding density of knowledge domains by decade



* Texts were often coded by more than one knowledge domain, thus total percentages per decade exceed 100%.

In order to determine if the coding densities of the domains varied in a statistically significant way, we conducted univariate Analysis of Variance (ANOVAs) tests using texts as our unit of analysis. The independent variable for each ANOVA was the decade in which the text was published: 1970s-1980s ($n = 13$), 1990s ($n = 18$), 2000s ($n = 25$), and 2010s ($n = 45$). The dependent variables (DVs) were the coding density scores derived from each of the domains of the integrated writing construct identified in our qualitative analysis (i.e., metacognitive, critical discourse, discourse, rhetorical, genre, writing process, subject matter, and digital literacies knowledge).

Next, we report the results of the tests for assumptions. A Shapiro-Wilk test ($p > .05$) revealed severe positive skewness across all DVs. To reduce skewness and to apply a constant so that the smallest score was $n = 1$, a $1/(x + 1)$ transformation was applied. Despite the transformation, the DVs subject matter and digital literacies knowledge still violated normality and were dropped from further analyses (note: relatively few texts were coded in these domains). The transformation also had the effect of eliminating most univariate outliers as determined by inspecting boxplots for values greater than 1.5 box-lengths from the edge of the box. Levene's test of homogeneity of variances was used to determine that metacognitive knowledge ($p = .37$) and discourse knowledge ($p = .27$) had homogeneity of variances based on their means. Critical discourse knowledge ($p = .00$), rhetorical knowledge ($p = .02$), genre knowledge ($p = .00$), and writing process knowledge ($p = .00$) had unequal variances. Descriptive statistics for the knowledge domains that we kept for analysis are reported in Table 3-1.

Table 3-1 Means and Standard Deviations for the Transformed Values of the Coding Density Scores of Texts by Decade

Domain	Coding Density Scores by Decade							
	1970s-1980s ($n = 13$)		1990s ($n = 18$)		2000s ($n = 25$)		2010s ($n = 45$)	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Metacognitive	.68	.38	.77	.34	.81	.31	.69	.33
Critical Discourse	.86	.30	.73	.36	.68	.35	.86	.23
Discourse	.78	.31	.69	.38	.61	.33	.77	.30
Rhetorical	.66	.35	.68	.39	.75	.29	.68	.31
Genre	.95	.13	.82	.29	.77	.32	.80	.28

Writing Process	.66	.39	.86	.22	.82	.31	.77	.32
-----------------	-----	-----	-----	-----	-----	-----	-----	-----

One-way Welch ANOVAs were conducted for critical discourse, rhetorical, genre, and writing process knowledge. Of these, only genre knowledge had a statistically significant difference in coding density scores among the decades, Welch's $F(3, 43.59) = 3.37, p = .03$. There was a decrease in coding density for genre from the 1970s-1980s ($M = .95, SD = .13$) to the 2010s ($M = .80, SD = .28$), a mean decrease of $-.15, 95\% CI [-.23, .00]$, which is a statistically significant mean difference ($p = .05$). The following were not statistically significant: critical discourse knowledge, Welch's $F(3, 34.35) = 2.11, p = .12$; rhetorical knowledge, Welch's $F(3, 36.03) = .35, p = .79$; and, writing process knowledge, Welch's $F(3, 37.78) = 1.16, p = .34$.

One-way ANOVAs were conducted for metacognitive and discourse knowledge. The coding density differences for metacognitive knowledge was not statistically significant, $F(3, 97) = .82, p = .49$. Also, the difference for discourse knowledge was not significant, $F(3, 96) = 1.52, p = .22$.

Discussion

This systematic review of the literature investigated what writing knowledge domains have been described in the literature over the last 50 years (RQ 1), how the domains might coalesce to describe the construct of writing (RQ 2), and how the emphasis on these domains has changed over the decades (RQ 3). Our qualitative analysis resulting from RQ 1 revealed the following knowledge domains that were described in the literature: metacognitive, critical discourse, discourse, rhetorical, genre, writing process, subject matter, and digital literacies knowledge. To answer RQ 2, we used operational diagramming via Nvivo's modeling function to notice the hierarchies and relationships that existed among these domains. Notably,

metacognitive knowledge was frequently used to describe all other domains, and so it had a hierarchical relationship to them. What this means is that you have metacognitive knowledge about something, for example, about genre. But, you can have genre knowledge without having metacognitive knowledge of it. In a similar manner, our model showed how critical discourse knowledge and discourse knowledge had a hierarchical relationship to other domains. While we had initially planned to incorporate digital literacies knowledge as a domain of its own, we noticed how it was intertwined with each knowledge domain. Thus, instead of making it a domain of its own, we decided to embed it into each domain. The results of this analysis led us to the creation of our integrated writing construct as illustrated in Figure 3-3. In RQ 3, we quantitatively analyzed the trends of the domains over the last five decades. We thought it best to understand where the construct has been in order to position it for where it might go. Through these analyses, we learned that the domains that were the most prevalent over the last 50 years have been rhetorical knowledge, followed by discourse and metacognitive knowledge. Meanwhile, digital literacies knowledge was the least prevalent, although its prevalence has been steadily growing in recent decades. More specifically, we wanted to know whether any of these trends of increasing or decreasing prevalence among the domains were statistically significant. Of the knowledge domains, only genre knowledge was statistically significant, seeing a decrease in coding density between the 1970s-1980s and the 2010s.

Implications for Curriculum, Instruction, and Assessment

The goal for both curriculum and assessment design should be for maximum construct representation. Thus, our goal for RQs 1 and 2 was to articulate a contemporary, integrated construct that might serve as a foundation of curriculum and assessment design. Ideally the construct sample captured in the curriculum, and the construct sample captured in any classroom

and large-scale assessment linked to that curriculum, should be aligned with one another. Misalignment of these construct samples tends to create negative consequences for teachers and students (Slomp, 2008). Hillocks demonstrates in *The Testing Trap* how more robust assessments (i.e., ones that capture more complete construct samples) better support teaching and learning than do more limited assessments (which capture a much smaller sample of the writing construct). Because writing portfolios more fully aligned with curriculum outcomes than did timed writing assessments, for example, they had more positive outcomes for student learning. Therefore, robust construct models are an integral part of designing effective curriculum, assessment, and instruction (Hillocks, 2002).

Situated within a bioecological model, the integrated writing construct (**Error! Reference source not found.**) serves to address ongoing limitations from both sociocultural and cognitive science models of writing, which individually, fail to provide a holistic model of writing development. Sociocultural models of writing are limited in that they generally ignore a writer's intrapersonal context. Failing to address a student's intrapersonal context means a failure to teach effectively. Take the disposition of motivation as an example. An abundance of research has shown that the importance of motivation to learning cannot be overstated. Motivation generates, sustains, and directs student learning (Ambrose, Bridges, DiPietro, Lovett, & Norman, 2010). Or, take the example of ethnicity, an aspect of a student's demand characteristics. Gee (2014) warns that students cannot learn the discourse community values of one community (e.g., school) if doing so conflicts with the socially-situated identity (e.g., ethnicity) that they take on as a member of another community:

If the identity required to be (and speak, write, and think) like a science student in this classroom here and now requires me, however tacitly, to disown, dishonor, or feel poorly

about my other identities and social languages (including languages other than English), then all bets are off—we have a perfect recipe for failure. This means that the community of practice created in the classroom must honor and allow for bridging across multiple identities. (Gee, 2014, p. 22)

Similarly, situating an integrated writing construct within a bioecological model addresses a key limitation of cognitive science models of writing that do not account adequately for the ecological context of writing. Flower and Hayes (1981) observed that people only solve the problems they define for themselves, and that weak writers often posed the wrong problems. In other words, novice writers are unaware of their rhetorical situation: they have limited knowledge of the discourse community for whom they are writing, they do not fully conceive the purpose or intention of their writing, and they are unfamiliar with the ways in which their ecological context constrains and enables their writing (Figure 3-1). Applying a bioecological model signals to those who teach and assess writing that, in order to understand a writing problem, it is crucial for students to have a metacognitive awareness of the boundaries of the writing problem thereby allowing for more robust and successful solutions to that problem.

Research supports the premise that metacognition plays an immense role in the development of writing expertise. A recent meta-analysis evaluating 25 years of empirical studies revealed that the teaching of metacognitive strategies has a strong effect on writing performance across all developmental levels (Kent & Wanzek, 2016a). Despite its importance, metacognition is rarely assessed on large-scale assessments (Jeffery, 2009; Slomp et al., 2014), which drive what is taught in the classroom (Hamp-Lyons, 1997; Hillocks, 2002; Messick, 1996).

Our modeling showed that metacognition frequently had a hierarchical relationship to other domains. For this reason, we positioned metacognition as superordinate to all other domains of writing knowledge in our integrated writing construct (**Error! Reference source not found.**). Additionally, our results show that metacognitive writing research is once again on the rise. This result suggests that we need to teach and assess all subsumed knowledge domains—discourse, critical discourse, genre, rhetorical, writing process, and subject matter knowledge—at a metacognitive level. As was discussed, a writer’s discourse community importantly shapes the transaction between the writer’s rhetorical situation and the ecological context—but if, and only if, a writer has the metacognitive awareness of this. Despite its importance, discourse community knowledge is more often than not neglected in the K-12 classroom. From our experience, a scan of most English textbooks will demonstrate how essays are written devoid of any discursive situation—as though a literary essay demands the same style or forms of evidence as an essay in political science or history. One popular English textbook, Oxford’s *The Canadian Writer’s Handbook* (Messenger, de Bruyn, Brown, & Montagnes, 2004), for example, devotes one page of 670 to the topic of audience and the word discourse is never mentioned.

Aside from metacognitive knowledge, discourse knowledge also frequently had a hierarchical relationship with other domains. Additionally, it was the second most frequently coded domain overall (Figure 3-4). Placing discourse community knowledge in the third most superordinate position in our integrated writing construct (**Error! Reference source not found.**) signals the important role it ought to play in teaching and assessing writing, and also in shaping the knowledge domains—namely, genre, rhetorical, writing process, and subject matter knowledge—contained within it. As our qualitative results showed, schools and other institutions serve a powerful role in regulating discourse, deciding which discourses will be

privileged and which will be Othered. Traditionally, educational systems have operated on a deficit model focusing on how students fail to appropriate a given discourse—namely, a dominant, monolithic one. Challenging the deficit model is a movement seeking to harness the diversity and complexity of a variety of discursive practices—multilingual, multicultural, multimodal, multidisciplinary, and multifarious more generally—to create a heteroglossia of learning (Gutiérrez, Baquedano-López, & Tejada, 1999; Van Heertum & Share, 2006). In our global village characterized by “increasing local diversity and global connectedness” (New London Group, 1996, p. 64), writers’ ability to respect, appreciate, and foster a multiplicity of discourses is vital to the success of their private, public, and civic lives. In increasingly globalized societies and globalized markets in particular—where it takes seconds for a company to conduct a transaction with another company in a different hemisphere—these multifarious discourses are not only accepted, they are encouraged; why should educational systems be any different?

Failing to integrate discourse knowledge into writing curriculum and instruction has resulted in students who are failing to transfer knowledge into new contexts (Beaufort, 2007). Taking up discourse knowledge in the curriculum will necessitate, for one, a shift in the construction of large-scale writing assessments away from decontextualized writing prompts toward more authentic writing assessments (Beck & Jeffery, 2007) that give due consideration to discourse knowledge, for it is generally acknowledged that assessment drives teaching and learning (Hillocks, 2002). Further, greater emphasis on discourse knowledge in the curriculum may create greater awareness of the varied discursive practices students bring to the classroom. Not all students arrive in class with equal social capital—students may be disadvantaged by class, gender, linguistic, or racial inequities—thus making it more challenging for certain

minoritized groups to attain insider status within academia (McKoski, 1995). Teachers' knowledge of this fact may help them recognize the strengths students bring from their varied discursive practices and build upon, rather than pathologize, those practices (Rodriguez, 2012).

Critical discourse knowledge also emerged as an important theme in our research. While it was not one of the most prevalent themes, when it did arise, it was frequently double coded with other domains. This suggests that critical discourse knowledge is an important resource in terms of writing domain knowledge and that it ought to play an important role in informing the domains subsumed within in. Recognizing critical discourse knowledge as a part of an integrated writing construct has important implications for curriculum, instruction, and assessment. Specifically, we believe that careful attention is needed to ensure that curriculum, instruction, and assessment do not further entrench inequalities along the lines of race, gender, sexuality, or language (Inoue, 2009; Slomp et al., 2014). Further, as schools are powerful agents of socialization, they are an ideal venue for developing students' awareness of social justice issues, challenging them to confront, rather than be complicit with, re/producing hegemonic discourses.

Another significant theme that emerged was that of rhetorical knowledge, and thus we chose to make this domain a vital part of the integrated writing construct and bioecological model. In our bioecological model, the rhetorical situation brings together both the social and physical considerations of writing. The National Writing Project (2010) logically advocates a double helix approach where the social and technological practices of writing are intertwined in educational standards and assessment. This will pose many a challenge as teachers who have been trained in Shakespeare and literary essay writing adapt to a world of websites and semiotics. However, teaching from an integrated writing perspective is "not about skills and

competence; it is aimed at creating a kind of person, an active designer of meaning, with a sensibility open to differences, change and innovation” (Cope & Kalantzis, 2009, p. 175). Thus, teacher education preparation ought not to focus on the skills of teaching technology, but rather on supporting students’ agency in their active and transformative creation of innovative, relevant, productive, and perhaps even emancipatory texts.

With respect to changing pedagogy around rhetorical knowledge, perhaps one of the fields most resistant to change will be that of educational assessment. While new literacies standards become more commonplace around the world (Leu et al., 2016),

forces of assessment and standardization exert a counter-pressure, asking us to prepare students to produce conventional, formulaic print texts in scripted ways [...] So it is that technology seems to be leading us forward to new forms of writing, but, as used by standardized testing programs, backward to the five-paragraph theme. (Herrington, Hodgson, & Moran, 2009, p. 2).

Until large-scale assessment catches up, there will inevitably be a certain amount of washback (Messick, 1996). In a similar manner to the ways in which rhetorical practices of multimodality are ignored on large-scale assessments, so too are critical thinking and metacognition (Slomp et al., 2014). This begs the question: why are important traits of the writing construct such as multimodality, critical thinking, and metacognition frequently precluded from large-scale tests? Hillocks has argued (2002) that traits such these, being more subjective, pose challenges to inter-rater reliability and thus are often ignored on large-scale assessments, a problem that some characterize as sacrificing validity for the sake of reliability (Moss, 1994; Wiggins, 1994).

Other themes that emerged from our review of the literature were the domains of genre, writing process, and subject matter knowledge. For these domains, metacognition is particularly

important because writing is, by its very nature, deictic. In our fast-paced global village, students need to be able to transfer their writing practices as they encounter new genres, writing processes, and subject matter. Additionally, students will face a multitude of new communications technologies, new modalities, and new socio-linguistic contexts. No teacher could ever prepare their students for this vast array of writing contexts (Downs & Wardle, 2007; Graves & Hyland, 2017). However, we would argue that teachers could help students develop the metacognitive practices to adapt to the vast array of contexts students will face upon graduation, many of which do not yet exist. And of course, “knowing how to create a digital text is not the same as knowing why” (National Writing Project, 2010) texts work or do not work for a given rhetorical situation. It is this deeper level of understanding that separates novice writers from expert writers.

Conclusion

But how do writers know what works or does not for a given rhetorical situation? How do writers arrive at meaningful transactions (purple section, Figure 3-1) that enable the analysis of the rhetorical situation, problem-posing, planning, the execution of a plan, and critically, the evaluation of the plan and its execution? In other words, how do writers develop metacognitive awareness? While these questions are beyond the scope of this paper, allow us to highlight the work of our colleagues who have sought to answer similar questions.

A first and obvious move, should we accept the premise advanced in our integrated writing construct and bioecological model, is that metacognition needs to be at the forefront of curriculum, assessment, and learning. A significant body of evidence suggests that assessment as learning (a means of promoting metacognition) is one of the most effective means of supporting student learning, particularly across new contexts: it supports student motivation,

lessens behavioural disruptions, leads to greater disciplinary expertise, leads to students taking greater responsibility for their own learning, and it has been documented to increase student achievement (Andrade & Boulay, 2003; Black, Harrison, Lee, Marshall, & William, 2004; Black & Wiliam, 1998; Chappuis & Stiggins, 2002; McDonald & Boud, 2003; Ross, 2006). According to Earl (2003)

[a]ssessment as learning is a metacognitive process in which students take ownership for improving their own learning. It involves students setting learning goals as well as monitoring, reflecting upon, and adjusting their own learning, often in response to feedback from the teacher and their peers. (p. 2)

But what types of assessment as learning practices are the most effective for improving writing? Do these practices consider both a writer's intrapersonal and ecological contexts? Do certain interventions work better for different subgroups? How and to what degree can technology be leveraged to provide students with a means of tracking learning goals and progress on writing, particularly through the use of e-portfolios? How and to what degree do students and teachers perceive the effectiveness of these assessment as learning strategies? These and other questions will importantly guide researchers who wish to break down the social versus cognitive research barrier. This research has the potential to assist students on their path to becoming not only independent writers, but also independent thinkers about writing—and that will go a long way in our fast-paced global village.

A second important point is that students cannot develop a metacognitive awareness of a variety of contexts (discursive, rhetorical, genre, technological, etc.) without exposure to and practice in them:

So when anyone is trying to speak/write or listen/read within a given social language within a given domain of practice, the crucial question becomes, what sorts of experiences (if any) has this person had that can anchor the situated meanings of words and phrases of this social language? (Gee, 2014, p. 21)

It will not suffice to artificially scaffold student learning by providing templates for generic, five-paragraph essays and assume that, as if by magic, students will know that five is an arbitrary number of paragraphs and that the rhetorical situation differs from essay to essay. Rather, we need to scaffold student learning with the focus of developing metacognitive awareness that begins with being able to analyze the rhetorical situation (purple section, Figure 1). Students will require much practice and embodied experience within a discourse community in order to “trigger the pattern” (Gee, 2014, p. 20), in other words, develop the appropriate writing and design schemas (Leijten et al., 2014). Cognitive apprenticeships have shown promise in helping students acquire these valuable metacognitive practices (A. Collins, Brown, & Holum, 1991), though their use has not been studied with regards to digital writing.

Finally, a major implication of our construct and bioecological model for writing pertains to large-scale assessment, for these vastly complicate the construct of interest being measured. The integrated writing construct and bioecological model make clear that to assess writing ability, we require measures that examine more than the written products students produce. These are imperfect artifacts of applied metacognition. Instead, what is required are measures that capture more directly the procedural and declarative knowledge about writing possessed by students and that situates this information within the bioecological contexts in which students are working (Leu et al., 2016; Slomp, 2012; Wardle & Roozen, 2012).

Appendix

Table 3-2 Writing Journals Identified for Hand Search and their Ratings

Journal	Impact Factor	
	CiteScore (Scopus, 2016)	<i>(Journal Citation Reports®)</i> , 2015)
Assessing Writing	2.23	1.095
Reading and Writing	1.96	1.308
Research in the Teaching of English	1.77	1.297
Written Communication	1.75	2.2
Journal of Writing Research	1.09	n/a

Table 3-3 Texts Identified in Literature Review and their Associated Coding Cumulative and Domain Densities

Text	CD								
	MK	K	DK	RK	GK	WPK	SMK	DLK	CCD
Alvermann, D., & Hagood, M. C. (2000). Critical media literacy: Research, theory, and practice in "New Times." <i>The Journal of Educational Research</i> , 93, 193–205. doi:10.1080/00220670009598707	0	0.87	0.88	0	0	0	0	0	1.75
Anderson, K. T., Stewart, O. G., & Kachorsky, D. (2017). Seeing Academically marginalized students' multimodal designs from a position of strength. <i>Written Communication</i> , 34, 104–134. doi:10.1177/0741088317699897	0	1.56	0	1.53	1.25	0	1.25	1.53	7.12
Aull, L. (2015). Connecting writing and language in assessment: Examining style, tone, and argument in the U.S. Common Core standards and in exemplary student writing. <i>Assessing Writing</i> , 24, 59–73. doi:10.1016/j.asw.2015.03.002	0	0	1.26	1.26	0	0	0	0	2.52
Barbeiro, L. F. (2011). What happens when I write? Pupils' writing about writing. <i>Reading and Writing</i> , 24, 813–834. doi:10.1007/s11145-010-9226-2	2.2	0	0	2.2	2.2	2.2	0	0	8.8
Bazerman, C. (2016). What do sociocultural studies of writing tell us about learning to write? In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), <i>Handbook of writing research</i> (2nd ed., pp. 11–23). New York: Guilford Press.	2.8	2.52	3.3	3.68	3.18	0.86	0	0.89	17.23
Beaufort, A. (2000). Learning the trade: A social apprenticeship model for gaining writing expertise. <i>Written Communication</i> , 17, 185–223. doi:10.1177/0741088300017002002	0	1.8	9.41	0.47	1.78	0	1.94	0	15.4
Beaufort, A. (1997). Operationalizing the concept of discourse community: A case study of one institutional site of composing. <i>Research in the Teaching of English</i> , 31, 486–530.	0	0.99	7.27	0.82	2.96	0	0	0	12.04
Beck, S. W., & Jeffery, J. V. (2007). Genres of high-stakes writing assessments and the construct of writing competence. <i>Assessing Writing</i> , 12, 60–79. doi:10.1016/j.asw.2007.05.001	1.68	0	1.46	0	1.68	0	1.68	0	6.5

Beers, S. F., Quinlan, T., & Harbaugh, A. G. (2010). Adolescent students' reading during writing behaviors and relationships with text quality: An eyetracking study. <i>Reading and Writing, 23</i> , 743–775. doi:10.1007/s11145-009-9193-7	0	0	0	0	0	0	0	0	0
Berlin, J. A. (1992). Poststructuralism, cultural studies, and the composition classroom: Postmodern theory in practice. <i>Rhetoric Review, 11</i> , 16–33. doi:10.1080/07350199209388984	0	3.11	0	0	0	0	0	0	3.11
Best, L. (1995). <i>A critique of cognitive research on writing from three critical perspectives: Theoretical, methodological, and practical</i> (Report). Retrieved from Kean College of New Jersey: https://archive.org/details/ERIC_ED377516 .	1.73	0	0.79	0.65	0	0.94	0.79	0	4.9
Brady, L. A. (1993). A contextual theory for business writing. <i>Journal of Business and Technical Communication, 7</i> , 452–471. article. doi:0803973233	1.07	1.68	3.95	4.24	0	0	0	0	10.94
Brandt, D. (1986). Toward an understanding of context in composition. <i>Written Communication, 3</i> , 139–157. doi:10.1177/0741088386003002001	2.01	0	1.9	2.89	0.47	0	0	0	7.27
Bulger, M. E., Mayer, R. E., & Metzger, M. J. (2014). Knowledge and processes that predict proficiency in digital literacy. <i>Reading and Writing, 27</i> , 1567–1583. doi:10.1007/s11145-014-9507-2	0.98	0	0	1.59	3.09	3.21	3.21	3.21	15.29
Burdick, H., Swartz, C. W., Stenner, J. A., Fitzgerald, J., Burdick, D., & Hanlon, S. T. (2013). Measuring students' writing ability on a computer-analytic developmental scale: An exploratory validity study. <i>Literacy Research and Instruction, 52</i> , 255–280. doi:10.1080/19388071.2013.812165	1.24	0	0	1.39	0	2.32	0	0	4.95
Burnett, C., Merchant, G., Pahl, K., & Rowsell, J. (2014). The (im)materiality of literacy: The significance of subjectivity to new literacies research. <i>Discourse: Studies in the Cultural Politics of Education, 35</i> , 90–103. doi:10.1080/01596306.2012.739469	0	0	1.01	0.45	0	0	0	0.45	1.91

Camp, H. (2012). The psychology of writing development--And its implications for assessment. <i>Assessing Writing</i> , 17, 92–105. doi:10.1016/j.asw.2012.01.002	1.83	0.78	1.61	0.13	0.13	0	0	0	4.48
Camp, H. (2012). The psychology of writing development-And its implications for assessment. <i>Assessing Writing</i> , 17, 92–105. doi:10.1016/j.asw.2012.01.002	0	0	0	0	0	0	0	0	0
Canagarajah, A. S. (1996). “Nondiscursive” requirement in academic publishing, material resources of periphery scholars, and the politics of knowledg production. <i>Written Communication</i> , 13. doi:10.1177/0741088396013004001	0	2.59	0	0	0	0	0	0	2.59
Cheng, X., & Steffensen, M. S. (1996). Metadiscourse: A Technique for Improving Student Writing. <i>Research in the Teaching of English</i> , 30, 149–181.	1.77	0	0	4.84	0.54	0	0	0	7.15
Cleary, L. M. (1991). Affect and cognition in the writing processes of eleventh graders. <i>Written Communication</i> , 8, 473–508. doi:10.1177/0741088391008004003	2.89	0	0	0	0	0.39	0	0	3.28
Condon, W., & Kelly-Riley, D. (2004). Assessing and teaching what we value: The relationship between college-level writing and critical thinking abilities. <i>Assessing Writing</i> , 9, 56–75. doi:10.1016/j.asw.2004.01.003	0	0.91	2.14	0	0	0	0	0	3.05
Crossley, S. A., Roscoe, R., & McNamara, D. S. (2014). What is successful writing? An investigation into the multiple ways writers can write successful essays. <i>Written Communication</i> , 31, 184–214. doi:10.1177/0741088314526354	0	0	0	0	0	0	0	0	0
Cumming, A. (2002). Assessing L2 writing: Alternative constructs and ethical dilemmas. <i>Assessing Writing</i> , 8, 73–83. doi:10.1016/S1075-2935(02)00047-8	0	3.67	0	0	0	0	0	0	3.67
de Milliano, I., van Gerlderen, A., & Slegers, P. (2012). Patterns of cognitive self-regulation of adolescent struggling writers. <i>Written Communication</i> , 29, 303–325. doi:10.1177/0741088312450275	2.04	0	0	2.88	2.88	2.88	3.29	0	13.97

Deane, P. (2013). On the relation between automated essay scoring and modern views of the writing construct. <i>Assessing Writing</i> , 18, 7–24. doi:10.1016/j.asw.2012.10.002	0	0	0.57	1.62	2.19	0	0.22	0	4.6
DePalma, M. J., & Ringer, J. M. (2011). Toward a theory of adaptive transfer: Expanding disciplinary discussions of “transfer” in second-language writing and composition studies. <i>Journal of Second Language Writing</i> , 20(2), 134–147. doi:10.1016/j.jslw.2011.02.003	1.84	0	0	0.3	0	0	0	0	2.14
Drijbooms, E., Groen, M. A., & Verhoeven, L. (2015). The contribution of executive functions to narrative writing in fourth grade children. <i>Reading and Writing</i> , 28, 989–1011. doi:10.1007/s11145-015-9558-z	1.49	0	0	0	0	2.34	0	0	3.83
Dryer, D. B. (2013). Scaling writing ability: A corpus-driven inquiry. <i>Written Communication</i> , 30, 3–35. doi:10.1177/0741088312466992	0	0.45	0	1.1	0.82	0.06	0.38	0	2.81
Dunmire, P. L. (2000). Genre as temporally situated social action: A study of temporality and genre activity. <i>Written Communication</i> , 17, 93–138. doi:10.1177/0741088300017001004	0	0	0.93	1.6	6.35	0	0	0	8.88
Durst, R. K. (1989). Monitoring processes in analytic and summary writing. <i>Written Communication</i> , 6, 340–363. doi:10.1177/0741088389006003005	6.55	0	0	1.91	0	1.42	2.01	0	11.89
Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. <i>College Composition and Communication</i> , 32, 365–387. doi:10.2307/356600	11.04	0	0	3.91	0	8.07	1.05	0	24.07
Gere, A. R. (1980). Written composition: Toward a theory of evaluation. <i>College English</i> , 42, 44–58. Retrieved from http://www.jstor.org/stable/10.2307/376032	0	0	0	1.01	0	0	0	0	1.01
Graham, S., Gillespie, A., & McKeown, D. (2013). Writing: Importance, development, and instruction. <i>Reading and Writing</i> , 26, 1–15. doi:10.1007/s11145-012-9395-2	2.01	1.19	1.73	1.73	0	0	0	1.18	7.84

Gregg, N., Sigalas, S. A., Hoy, C., Wisenbaker, J., & McKinley, C. (1996). Sense of audience and the adult writer: A study across competence levels. <i>Reading and Writing, 8</i> , 121–137. doi:10.1007/BF00423929	0	0	0	0	0	0	0	0	0
Hamp-Lyons, L. (2005). What is writing? What is “scholastic aptitude”? What are the consequences? SAT I Writing - A trip down memory lane. <i>Assessing Writing, 10</i> , 151–156. doi:10.1016/j.asw.2005.09.002	0	0	2.16	2.16	2.16	4.95	0	0	11.43
Hassett, D. (2006). Technological difficulties: a theoretical frame for understanding the non-relativistic permanence of traditional print literacy in elementary education. <i>Journal of Curriculum Studies, 38</i> , 135–159. doi:10.1080/00220270500363703	0	0	0	0	0	0	0	0.33	0.33
Hawkey, R., & Barker, F. (2004). Developing a common scale for the assessment of writing. <i>Assessing Writing, 9</i> , 122–159. doi:10.1016/j.asw.2004.06.001	0	0	1.04	1	1.04	0	0	0	3.08
Hayes, J. R. (2012). Modeling and remodeling writing. <i>Written Communication, 29</i> , 369–388. doi:10.1177/0741088312451260	0	0	0	0	0	4	0	0	4
Hull, G. A., & Nelson, M. E. (2005). Locating the semiotic power of multimodality. <i>Written Communication, 22</i> , 224–261. doi:10.1177/0741088304274170	0	0.85	0.34	0.96	0	0	0	0.96	3.11
Humes, A. (1983). Putting Writing Research into Practice. <i>The Elementary School Journal, 84</i> , 3–17. doi:10.1086/461339	0	0	0	0	0	1.36	0	0	1.36
Ivanic, R. (2004). Discourses of writing and learning to write. <i>Language and Education, 18</i> , 220–245. doi:10.1080/09500780408666877	0	6.08	1.76	0.64	3.15	1.5	0	0	13.13
Jeffery, J. V. (2009). Constructs of writing proficiency in US state and national writing assessments: Exploring variability. <i>Assessing Writing, 14</i> , 3–24. doi:10.1016/j.asw.2008.12.002	0	0.93	1.26	0	1.26	0	0	0	3.45
Jones, S., Myhill, D., Bailey, T., & Bailey, T. (2013). Grammar for writing? An investigation of the effects of contextualised grammar teaching on students’ writing. <i>Reading and Writing, 26</i> , 1241–1263. doi:10.1007/s11145-012-9416-1	0	1.44	0.68	1.62	1.19	0.87	0	0	5.8

Karlen, Y. (2017). The development of a new instrument to assess metacognitive strategy knowledge about academic writing and its relation to self-regulated writing and writing performance. <i>Journal of Writing Research</i> , 9, 61–86. doi:10.17239/jowr-2017.09.01.03	0	0	0	0	0	0	0	0	0
Kelly-Riley, D. (2015). Toward a validation framework using student course papers from common undergraduate curricular requirements as viable outcomes evidence. <i>Assessing Writing</i> , 23, 60–74. doi:10.1016/j.asw.2014.10.001	0	0.26	0	2.12	0	0	0	0	2.38
Kennedy, M. L. (1985). The composing process of college students writing from sources. <i>Written Communication</i> , 2, 434–456. doi:10.1177/0741088385002004006	0.96	0	0	0	0	7.76	0	0	8.72
Kent, T. (1991). On the very idea of a discourse community. <i>College Composition and Communication</i> , 42, 425–445. doi:10.2307/357995	0	0	4.48	0	0	0	0	0	4.48
Kent, S. C., & Wanzek, J. (2016). The relationship between component skills and writing quality and production across developmental levels: A meta-analysis of the last 25 years. <i>Review of Educational Research</i> , 86, 570–601. doi:10.3102/0034654315619491	0	0	0	0	0	0	0	0	0
Kirkpatrick, L. C., & Klein, P. D. (2016). High-achieving high school students' strategies for writing from internet-based sources of information. <i>Journal of Writing Research</i> , 8, 1–46. doi:10.17239/jowr-2016.08.01.01	0.55	0	0	1.49	0	0	0	0	2.04
Knoch, U., & Sitajalabhorn, W. (2013). A closer look at integrated writing tasks: Towards a more focused definition for assessment purposes. <i>Assessing Writing</i> , 18, 300–308. doi:10.1016/j.asw.2013.09.003	0	0	0	0	0	0	0	0	0
Kopple, W. J. Vande. (1985). Some Exploratory Discourse on Metadiscourse. <i>College Composition and Communication</i> , 36, 82. doi:10.2307/357609	0	0	0.57	0	0	0	0	0	0.57

Proske, A., & Kapp, F. (2013). Fostering topic knowledge: Essential for academic writing. <i>Reading and Writing</i> , 26, 1337–1352. doi:10.1007/s11145-012-9421-4	2.74	0	1.58	1.27	0	2.55	2.55	0.35	11.04
Reiff, M. J., & Bawarshi, A. (2011). Tracing discursive resources: How students use prior genre knowledge to negotiate new writing contexts in first-year composition. <i>Written Communication</i> , 28, 312–337. doi:10.1177/0741088311410183	1.19	0	1.59	0	1.19	0	0	0	3.97
Ritchey, K. D. (2008). The building blocks of writing: Learning to write letters and spell words. <i>Reading and Writing</i> , 21, 27–47. doi:10.1007/s11145-007-9063-0	0	0	0	0	0	0	0	0	0
Ronald, K., & Volkmer, J. (1989). Another competing theory of process: The student's. <i>Journal of Advanced Composition</i> , 9, 83–96. Retrieved from http://www.jstor.org/stable/20865673	0	0.5	0	0.78	0	3.52	0	0	4.8
Rowan, K. E. (1990). Cognitive correlates of explanatory writing skill. <i>Written Communication</i> , 7, 316–341. doi:10.1177/0741088390007003002	0	0	5.58	7.77	2.06	0	5.75	0	21.16
Ryan, M. (2014). Reflexive writers: Re-thinking writing development and assessment in schools. <i>Assessing Writing</i> , 22, 60–74. doi:10.1016/j.asw.2014.08.002	2.32	1.28	0.74	0.35	0.35	0	0	0	5.04
Salibrici, M. M. (1999). Dissonance and rhetorical inquiry: A Burkean model for critical reading and writing. <i>Journal of Adolescent & Adult Literacy</i> , 42, 622–626.	0	7.41	0.78	2.7	1.59	0	0	0	12.48
Slomp, D. H. (2012). Challenges in assessing the development of writing ability: Theories, constructs and methods. <i>Assessing Writing</i> , 17, 81–91. http://doi.org/10.1016/j.asw.2012.02.001	4.93	0	3.31	1.26	0.98	2.04	2.02	0	14.54
Spellmeyer, K. (1989). Foucault and the freshman writer: Considering the self in discourse. <i>College English</i> , 51, 715–729.	0	15.4 9	6.12	0	0	0	0	0	21.61

Stolarek, E. A. (1994). Prose modeling and metacognition: The effect of modeling on developing a metacognitive stance toward writing. <i>Research in the Teaching of English</i> , 28, 154–174.	6.38	0	0	0	0	0	0	0	6.38
Sydow, D. L. (1994). Composition theory: Taking it apart and putting it back together again. Paper presented at the Conference on College Composition and Communication, Nashville, TN. Paper retrieved from https://eric.ed.gov/?id=ED376471	2.96	0	0	0	0	2.16	0	0	5.12
Troia, G. A., Harbaugh, A. G., Shankland, R. K., Wolbers, K. A., & Lawrence, A. M. (2013). Relationships between writing motivation, writing activity, and writing performance: Effects of grade, sex, and ability. <i>Reading and Writing</i> , 26, 17–44. doi:10.1007/s11145-012-9379-2	1.41	0	0	0	0	0	0	0	1.41
van Arragon, K. B. (2003). <i>Balancing the roles of explicit instruction of text form language and schema theory in student non-fiction writing: Problems and possibilities</i> (Report). Retrieved from O.I.S.E. University of Toronto: https://archive.org/details/ERIC_ED539725 .	2.59	0	2.47	0	0	0	0	0	5.06
Van Heertum, J., & Share, J. (2006). A new direction for multiple literacy education. <i>McGill Journal of Education</i> , 41(3), 249–265.	0	9.41	1.04	0	0	0	0	0	10.45
Vanderberg, R., & Lee Swanson, H. (2007). Which components of working memory are important in the writing process? <i>Reading and Writing</i> , 20, 721–752. doi:10.1007/s11145-006-9046-6	2.7	0	0	0.85	0	0.85	0	0	4.4
von Koss Torkildsen, J., Morken, F., Helland, W. A., & Helland, T. (2016). The dynamics of narrative writing in primary grade children: Writing process factors predict story quality. <i>Reading and Writing</i> , 29, 529–554. doi:10.1007/s11145-015-9618-4	1.68	0	0	0	0	1.51	0	0	3.19

Wagner, R. K., Puranik, C. S., Foorman, B., Foster, E., Wilson, L. G., Tschinkel, E., & Kantor, P. T. (2011). Modeling the development of written language. <i>Reading and Writing, 24</i> , 203–220. doi:10.1007/s11145-010-9266-7	1.34	0	3.54	2.41	0	1.34	0	0	8.63
Weigle, S. C. (2013). English language learners and automated scoring of essays: Critical considerations. <i>Assessing Writing, 18</i> , 85–99. doi:10.1016/j.asw.2012.10.006	0	0	0	0	0	0	0	0	0
Witte, S. P. (1992). Context, text, intertext: Toward a constructivist semiotic theory of writing. <i>Written Communication, 9</i> , 237–308. doi: 10.1177/0741088392009002003	0	0	0.18	0	0.2	1.1	0	0	1.48
Wood, J. W. (2002). Defining literacies. Paper presented at the Annual Meeting of the <i>American Educational Research Association</i> , New Orleans, LA. Retrieved from https://eric.ed.gov/?id=ED465980	0	2.14	4.29	0.52	0	7.98	0	0	14.93
MK = Metacognitive Knowledge; CDK = Critical Discourse Knowledge; DK = Discourse Knowledge; GK = Genre Knowledge; SMK = Subject Matter Knowledge; RK = Rhetorical Knowledge; WPK = Writing Process Knowledge; DLK = Digital Literacies Knowledge; CCK = Cumulative Coding Density									

Table 3-4 *Texts with the Top 10 Cumulative Coding Densities*

Bibliographic Information	MK	CDK	DK	RK	GK	WPK	SMK	DLK	CCD
1. Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. <i>College Composition and Communication</i> , 32, 365–387. doi:10.2307/356600	11.04	0	0	3.91	0	8.07	1.05	0	24.07
2. Leu, D. J., Slomp, D. H., Zawilinski, L., & Corrigan, J. A. (2016). Writing research through a New Literacies lens. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), <i>Handbook of writing research</i> (2nd ed., pp. 41–53). New York: Guilford Press.	2.34	0.05	3.37	4.43	0.81	0	0	12.01	23.01
3. Spellmeyer, K. (1989). Foucault and the freshman writer: Considering the self in discourse. <i>College English</i> , 51(7), 715–729.	0	15.49	6.12	0	0	0	0	0	21.61
4. Rowan, K. E. (1990). Cognitive correlates of explanatory writing skill. <i>Written Communication</i> , 7, 316–341. doi:10.1177/0741088390007003002	0	0	5.58	7.77	2.06	0	5.75	0	21.16
5. MacArthur, C. A., & Graham, S. (2016). Writing research from a cognitive perspective. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), <i>Handbook of writing research</i> (2nd ed., pp. 24–40). New York: Guilford Press.	6.51	0.94	0.75	2.3	0.31	5.93	0.72	0	17.46
6. Bazerman, C. (2016). What do sociocultural studies of writing tell us about learning to write? In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), <i>Handbook of writing research</i> (2nd ed., pp. 11–23). New York: Guilford Press.	2.8	2.52	3.3	3.68	3.18	0.86	0	0.89	17.23
7. New London Group. (1996). A pedagogy of multiliteracies: Designing social futures. <i>Harvard Educational Review</i> , 66, 60–92. doi:10.17763/haer.66.1.17370n67v22j160u	0	3.18	5.51	5.01	2.05	0.18	0	0.49	16.42
8. Beaufort, A. (2000). Learning the trade: A social apprenticeship model for gaining writing expertise. <i>Written Communication</i> , 17(2), 185–223. doi:10.1177/0741088300017002002	0	1.8	9.41	0.47	1.78	0	1.94	0	15.4
9. Bulger, M. E., Mayer, R. E., & Metzger, M. J. (2014). Knowledge and processes that predict proficiency in digital literacy. <i>Reading and Writing</i> , 27, 1567–1583. doi:10.1007/s11145-014-9507-2	0.98	0	0	1.59	3.09	3.21	3.21	3.21	15.29

10. Wood, J. W. (2002). Defining literacies. In *Annual Meeting of the American Educational Research Association*. New Orleans, LA.

0	2.14	4.29	0.52	0	7.98	0	0	14.93
---	------	------	------	---	------	---	---	-------

MK = Metacognitive Knowledge; CDK = Critical Discourse Knowledge; DK = Discourse Knowledge; GK = Genre Knowledge; SMK = Subject Matter Knowledge; RK = Rhetorical Knowledge; WPK = Writing Process Knowledge; DLK = Digital Literacies Knowledge; CCD = Cumulative Coding Density

Table 3-5 Means (and Standard Deviations) for Transformed Values of Coding Densities by Decade

	Total <i>N</i> = 101	1970s & 1980s <i>n</i> = 13	1990s <i>n</i> = 18	2000s <i>n</i> = 25	2010s <i>n</i> = 45
Metacognitive Knowledge	.73 (.33)	.68 (.38)	.77 (.34)	.81 (.31)	.69 (.33)
Critical Discourse Knowledge	.79 (.30)	.86 (.30)	.73 (.36)	.68 (.35)	.86 (.23)
Discourse Knowledge	.72 (.33)	.78 (.31)	.69 (.38)	.61 (.33)	.77 (.30)
Rhetorical Knowledge	.69 (.32)	.66 (.35)	.68 (.39)	.75 (.29)	.68 (.31)
Genre Knowledge	.82 (.28)	.95 (.13)	.82 (.29)	.77 (.32)	.80 (.28)
Writing Process Knowledge	.78 (.32)	.66 (.39)	.86 (.22)	.82 (.31)	.77 (.32)
Subject Matter Knowledge	.90 (.23)	.88 (.24)	.93 (.22)	.90 (.24)	.89 (.24)

Table 3-6 *Intercorrelations among Knowledge Domains*

Knowledge Domain	1.	2.	3.	4.	5.	6.
1. Metacognitive Knowledge	--	-.27**	-.04	.22*	.00	.41**
2. Critical Discourse Knowledge		--	.33**	.04	.18	-.14
3. Discourse Knowledge			--	.26**	.44**	-.08
4. Rhetorical Knowledge				--	.45**	.26**
5. Genre Knowledge					--	.09
6. Writing Process Knowledge						--

Note: $N = 101$

** $p > .01$ (2-tailed)

* $p > .05$ (2-tailed)

**CHAPTER 4 - EXPLORING THE DIFFERENCES BETWEEN STUDENT
WRITERS AND KNOWLEDGE WORKERS IN ONLINE RESEARCH
WRITING**

**Exploring the Differences between Student Writers and Knowledge Workers in Online
Research Writing**

Julie A. Corrigan

University of Ottawa

Author Note

Julie A. Corrigan, Faculty of Education, University of Ottawa

This work was generously supported by a Social Sciences and Humanities Research Council of Canada Bombardier Canada Graduate Scholarship, Ontario Graduate Scholarships, and internal funding from the University of Ottawa.

Abstract

Though the Internet and digital technologies have revolutionized the way we write, there is a relative paucity of research demonstrating what new cognitive and metacognitive processes are required beyond those involved in traditional writing. Further, little is known about how these processes are enacted differently between novices and experts. Therefore, the purpose of this study was to explore the cognitive and metacognitive processes involved—and their spectrum across the novice-expert continuum—as people research and write in online environments. In this exploratory, novice-expert study, I used a mixed methods approach to investigate the differences between student writers (ranging from sixth to ninth grade) and knowledge workers (including a technical writer, lawyer, professor, registered dietician, financial management analyst, and nuclear safety analyst, among others) who were invited to complete a large-scale assessment of online researching and writing. Data sources for this study included test scores, but also cued retrospective reporting and interview transcripts, questionnaires, and writing artefacts from school and work. The results suggest that there are complex and sophisticated cognitive and metacognitive processes underlying online research writing and that some of these processes are unique to online contexts. Secondly, both quantitative and qualitative results suggest significant differences between novice and expert groups. Implications for research, theory, and practice are discussed.

Exploring the Differences between Student Writers and Knowledge Workers in Online Research Writing

As writing moves from page to pixel, there is a relative paucity of research demonstrating what new cognitive and metacognitive processes are required beyond those involved in traditional writing (Corrigan & Slomp, 2017; Leijten et al., 2014). Further, while there is a limited tradition of researching the differences between novices and experts in writing (Schriver, 2012), there is an even greater paucity of this research in online contexts. Understanding the cognitive and metacognitive processes involved in online writing, and their developmental range, plays an important role in the development of constructs upon which test designers and curriculum developers can base their assessments and curricula. Without an empirically based construct of online writing, test designers and curriculum developers are left to base their work on out-dated constructs. This then carries over to the classroom because curriculum, assessment, and instruction are ideally integrated and aligned with one another (Gareis & Grant, 2015). The end result is students getting left behind for not having the digital writing practices that they need to be effective in their social, professional, and democratic lives.

In this exploratory, novice-expert study, I investigated the cognitive and metacognitive processes underlying one of the most common practices of online writing, which I describe as online research writing (ORW). ORW involves locating, critically evaluating, and synthesizing online information in order to communicate it across online platforms such as email or wiki (Leu et al., 2012). I used a mixed methods approach to investigate the differences between student writers (ranging from sixth to ninth grade) and knowledge workers (including a technical writer, lawyer, professor, registered dietician, financial management analyst, and nuclear safety analyst, among others) who were asked to complete an assessment of ORW known as the Online

Research Comprehension Assessment (ORCA). The ORCA was created to be a reliable, valid, and practical assessment of online research and communication ability among seventh-grade students, particularly in the subject of science (Leu et al., 2012). Further description of this assessment tool will be provided in the Methods section. Data sources for this study included ORCA test scores, but also cued retrospective reporting, interviews, surveys, and writing artefacts from school and work. The purpose of the study is twofold: to identify what cognitive and metacognitive processes underlie the construct of ORW; and, to identify the differences between novices and experts with regards to these processes.

Online Research Writing

In this article, ORW refers to the practice of locating, critically evaluating, and synthesizing online information for the purpose of communicating it via an online platform (such as an email, website, wiki, blog, microblog [e.g., Twitter], or multimodal report). Leijten and colleagues (2014) called this process constructing documents using multiple digital sources or composing in distributed environments. The closest analog terms would be research writing, reading to write (Spivey & King, 1989), or writing from multiple sources (O'Hara, Taylor, Newman, & Sellen, 2002). What these analog terms have in common with ORW is that all acknowledge the hybridity of processes—such as reading, writing, synthesizing, and integrating—required to research and communicate information (O'Hara et al., 2002; Spivey & King, 1989). According to the research in this study and that of my colleagues (Leijten et al., 2014; Leu et al., 2007), ORW diverges from its analog counterpart in marked ways. Specifically, additional, distinct processes are involved in ORW that include, for example, locating online information, critically evaluating online sources, managing online information, and using multimodal and hypertextual means to represent information.

Several indicators have suggested that the importance of ORW is on the rise. Surveys have suggested that professionals spend an average of 24% of their work week writing, while careers in technical communication are expected to rise by 18% from 2008 to 2018, which is “faster than the average of all occupations” (Schriver, 2012, p. 275). According to Madden and Jones (2008), 62% of all employed adults in the United States could be considered networked workers—those who use the Internet or email in their workplace. A subset of these workers is classified as knowledge workers, a term which refers to someone who primarily “thinks for a living” (Davenport, 2005, p. 3). More specifically, a knowledge worker is someone who works primarily with information, developing it and using it in the workplace. While many occupations require thinking to some extent, “it is the perennial processing of non-routine problems that require non-linear and creative thinking that characterizes knowledge work” (Reinhardt, Schmidt, Sloep, & Drachsler, 2011, p. 150). Estimates put the number of knowledge workers in developed nations at a quarter to a half of the entire workforce (Davenport, 2005), a number that will continue to rise as the labour market moves toward knowledge-based industries (Yan, 2005). Regardless of the term used, a large proportion of the workforce in developed countries is working in tertiary industries where the prime commodity is knowledge. Because knowledge is encoded as writing—broadly conceived as any form of encoding of monomodal or multimodal text—the act of writing figures prominently in the jobs of these workers (Brandt, 2005). Additionally, to discover knowledge, one conducts research. In a digital age, it is fair to assume that this primarily takes place online—particularly in the case of secondary research, which this study examined. Thus, it is a fair deduction that ORW is an important component of the jobs of knowledge workers. Also, because the knowledge economy is growing (Powell & Snellman, 2004), another fair deduction is that the number of workers required to do ORW will grow, too.

Past Research on ORW

Relatively little is known about ORW. Although it seems reasonable that some of what we know about research writing from print-based sources (Kennedy, 1985; O’Hara et al., 2002; Spivey & King, 1989) also applies to ORW, there is a need to expand our understanding to consider the ways in which the Internet has revolutionized this practice. With regards to online reading, Coiro and Dobler (2012) were among the first to support the claim that, while online and offline reading have much in common, new cognitive processes are required in online contexts. For example, the processes of locating and evaluating information are vastly different in online versus traditional contexts (Coiro & Dobler, 2007; Coiro, Knobel, Lankshear, & Leu, 2008). Locating online information has been likened to drinking from a fire hose as the process can be truly overwhelming, particularly for novices. Further, critically evaluating information requires unique processes in online contexts, particularly because a feature of online writing is that there is very little gate-keeping—anyone with a smart phone and an Internet connection can publish online in an instant. We are only beginning to expand our understanding of online writing, especially with regards to hybrid reading-writing practices such as ORW. For the most part, these hybrid activities have only been examined separately (Mateos, Martín, Villalón, & Luna, 2008).

Preliminary research into hybrid reading-writing activities has revealed that better writing was generated by students with “high recursivity” which was described as “reading of the source text(s), writing while rereading the sources and revising of the text produced (reading the text produced and editing and reformulating it)” (Mateos et al., 2008, p. 690). This is an important finding, one which would benefit from being expanded to include the complexities of online contexts, as Kirkpatrick and Klein's (2016) research does. Their research found that students

resorted to one of two global writing strategies: using mediated planning documents (used to record their research findings and organize their thinking as they completed their writing tasks); or, writing directly from source texts. Kirkpatrick and Klein's (2016) research also revealed sub-strategies particular to online contexts, those “not seen in either writing without the Internet or using the Internet without having to write” (p. 34). For example, students engaged in rhetorically-driven Internet searches in order to retrieve content that would align with their rhetorical stance (e.g., “animal testing cons”). Taking somewhat of a departure that ORW requires processes unique to online contexts, Bulger, Mayer, and Metzger's (2014) research showed that proficiency in ORW could be predicted better by academic experience and domain knowledge than by digital literacy proficiency. In other words, academic literacies are more important than digital literacies. I would argue that it is a misconception to consider traditional/academic and digital literacies as separate; rather, traditional literacies work in tandem with new literacies as writers require new forms of strategic knowledge in online contexts (Leu et al., 2004). Or as Stroupe (2000) would describe it, digitality and literacy are intertwined. Another important advancement regarding ORW was the development of a model of writing that considered, for the first time, the impact of the Internet on writing (Leijten et al., 2014), which will be described in more detail forthwith. As the above studies illustrate, online writing and ORW in particular are gaining research momentum, though there is still much to be learned. The current study will expand this knowledge by considering differences between novices and experts during ORW, with a particular emphasis on the cognitive and metacognitive processes and their situatedness in intrapersonal and ecological contexts.

Models of Writing and Writing Expertise

The diversity of approaches taken by writing researchers—ranging from cognitive to sociocultural to postprocess—has produced varied and complex models to explain the practice of writing. Beginning with Flower and Hayes' (1980) seminal work, writing became recognized as a problem-solving, cognitive and metacognitive process. Flower and Hayes (1980) were also the first to recognize that writers only solve the problems they represent for themselves. Stated another way, even if a group of writers were given the same assignment, it is they—not the assignment—who define the problem.

Although novice-expert differences in writing have not been studied extensively, findings from these studies have been enlightening. Sommers' (1980) seminal novice-expert writing study used a case study approach to examine the differences between student writers and experienced writers in their revision processes. According to Sommers, while the expert writers completed global revisions, with a deep consideration for audience, novice writers exhibited more of a focus on local revisions. But what might explain these very differing foci between novices and experts? Bereiter and Scardamalia's (1987) research seven years later would shed some light on this question.

Bereiter and Scardamalia's (1987) landmark work *The psychology of written composition* was informed by 120 empirical (mostly novice-expert and experimental) studies completed by the pair and their colleagues over eight years. This work illuminated the apparent paradox of why novice writers seem to write with comparatively little effort when compared to experts who seem to have a tremendous investment of mental effort and are plagued by false starts and much longer start-up times. Bereiter and Scardamalia (1987) explain this paradox by means of two models of composition: knowledge telling and knowledge transforming. What the pair

discovered was that novice writers tend to think of what to say and how to say it (knowledge telling) while expert writers engage in a more complex process considering not only the content, but also their rhetorical goals (knowledge transforming). As a strategy carried over from oral language production, novices reduce the complexity of writing to the content problem space (i.e., the knowledge telling model). Experts, conversely, engage with the complexity of a dual-problem space including both the content problem space and its complex interplay with the rhetorical problem space (i.e., the knowledge transforming model). It is important to note here that these models refer to the process used to construct the texts, and not the texts themselves. In the case of an expert conducting a very routine writing task, for example, it may appear that knowledge transforming processes were used when in fact that may not have been the case.

Kellogg (2008) later speculated that a third model might exist called knowledge crafting, wherein expert writers shape “what to say and how to say it with the potential reader fully in mind” (p. 7). Sommers' (1980) earlier research appears to support the validity of Kellogg's knowledge crafting model:

The experienced writers imagine a reader (reading their product) whose existence and whose expectations influence their revision process [...] The anticipation of a reader's judgement causes a feeling of dissonance when the writer recognizes incongruities between the intention and execution, and requires these writers to make revisions on all levels. (p. 385).

The knowledge telling versus knowledge transforming/crafting models are a powerful way of explaining the different problem-solving processes engaged with by writers. Novices pose problems around what to write; meanwhile, experts consider what to write in relation to how to

write it, and also give due consideration to their audience. But what cognitive developmental processes might account for the differing use of these models?

According to Alamargot and Fayol's (2009) research, writing expertise develops as transcription and orthography become automatized thereby freeing up working memory to represent rhetorical problems, and as metacognitive knowledge is expanded through instruction and practice. "Knowledge-transforming, and especially knowledge-crafting," argued Kellogg (2008) in a similar fashion, "occur only when sufficient executive attention is available to provide a high degree of cognitive control over the maintenance of multiple representations of the text as well as planning conceptual content, generating text, and reviewing content and text." This may also to explain why knowledge telling heuristics do not necessarily coincide with poor writing while, likewise, knowledge transforming heuristics do not necessarily coincide with strong writing: "Such deliberate control of normally unmonitored activity [i.e., the rhetorical problem space] exacts a price in mental effort and it opens up possibilities of error" (Bereiter & Scardamalia, 1987, Chapter 1, Writing as both natural and problematic, para. 6). So, why would any writer pursue a more costly knowledge transforming process? According to Bereiter and Scardamalia, doing so can lead to expertise far beyond the knowledge telling model. Specifically, using a knowledge transforming approach allows a writer to shape a text to achieve its intended effect, that is, its rhetorical goal.

Yet, despite such novice-expert research, the verdict is still out on whether or not such a thing as writing expertise exists. Bazerman (1994), for example, argued that writing is so contextualized that competence in one domain (e.g., journalism) often has no bearing on competence in another (e.g., poetry or novel writing). He even goes so far as to argue that within a domain, that competence oftentimes does not transfer (e.g., going from writing dramatic poetry

to lyric poetry), and he illustrates his point using compelling examples (e.g., “T.S. Eliot is expert at writing Eliot poems but not nearly so good at writing Yeats poems, and certainly not Wordsworth poems” [p. 132]). Conversely, Beaufort (2007) has argued that “while expertise does not transfer wholesale from one context to another, it is possible to identify the common knowledge domains within which writers must develop context-specific knowledge” (p. 17). This begs the question, is there such a thing as expertise in writing, or is the field so diffuse that writing experts lack a common protocol for solving rhetorical problems?

It may be fruitful to this discussion of whether or not writing expertise exists to consider more nuanced categories of expertise. Studies in fields outside of writing propose that there are two types of expertise: generic (a.k.a. adaptive or prospective) and specific. Warren (2011) defines generic expertise as “the ability to represent field-specific problems accurately and efficiently, regardless of whether they are working in their area of specialization” (p. 350). An expert then, according to Patel and Groen (1991), is someone who possesses a high level of both generic and specific expertise. Thus, a writer who has specific expertise would be an expert on the content area and have the ability to apply writing knowledge to meet the needs of the discourse community within that content area. Expert writers are those who possess “very rich, deep, context-specific knowledge, but they also have mental schema, or heuristics, with which to organize knowledge and aid problem-solving and gaining new knowledge in new situations” (Beaufort, 2007, p. 17). Patel and Groen (1991) referred to doctors solving medical problems outside of their domain as subexperts. Applied to writing, subexperts are those who write outside of their discourse community; these writers represent rhetorical problems accurately and efficiently, but lack the specific subject matter and discourse community knowledge to write as an insider. Lay people who write manuals for high tech companies would be a good example of

subexperts. These classifications call to mind Bereiter and Scardamalia's (1987) conceptualization of writing expertise in terms of dual problem spaces: content problem space and rhetorical problem space. Expertise, according to Bereiter and Scardamalia, is demonstrated through the interplay between these two problem spaces. This study will suggest, however, that this dual problem space model is too simplified, and that writing is in fact the interplay among at least seven problem spaces.

This interplay between subject matter knowledge and rhetorical knowledge is indirectly represented in Leijten and colleagues' (2014) model of skilled professional communicators, which is an extension of Hayes' earlier models (Chenoweth & Hayes, 2001; Flower & Hayes, 1981; Hayes, 1996, 2012). As in Flower and Hayes' (1981) earlier model, the recent model represents subject matter knowledge at the resource level under long-term memory while rhetorical memory would also be represented there, as well as at the control level under writing and design schemas. Hayes' and colleagues' most recent model (Leijten et al., 2014) updates earlier models of writing by taking into consideration the multimodal nature of text in our digital age. For example, the term transcription technology becomes production technology, text-produced-so-far becomes text-and-graphics-created so far, and writing schemas becomes writing and design schemas.

What is less evident in the models of Hayes and other cognitive writing theorists is the socially situated nature of writing. The act of writing and the development of writers is a complex, psychosocio-(meta)cognitive phenomenon (Corrigan & Slomp, 2017) influenced not only by biopsychological factors (e.g., working memory, long-term memory, motivation, fine motor control, and intelligence), but also by writers' situatedness in time and place. Ecological factors at the individual, community, and societal levels influence writers and their development

making causal inferences challenging due to the vast number of variables individually or in combination that may be exerting influence (Schultz & Fecho, 2000). Flower (1989) conceded that early cognitive models, including her own, largely downplayed the social function of writing. Social factors such as task environment and awareness of audience are taken in account, but presented in a static, limited manner only as they relate to composing processes.

The social turn in writing and literacy studies (Gee, 1999) brought about an increased awareness to the “view that reading and writing only make sense when studied in the context of social and cultural [and we can add historical, political, and economic] practices of which they are but a part” (p. 3). Consequently, sociocultural models of writing help us understand why and when people write, how writing gets done, how writing is learned, and what consequences amass from the practice of writing (Bazerman, 2016). But do they paint the full picture? Yes, writers are products of their environment—highly influenced by a range of social and cultural, local and global, recent and historical factors. But writers are also individuals with autonomy and agency who become the writers that they are sometimes in spite of their backgrounds, their sociocultural influences. Writers are also individuals with unique DNA leading to unique biopsychological combinations of intelligence, fine-motor control, affect, motivation, personality, and memory (Bronfenbrenner & Morris, 2006).

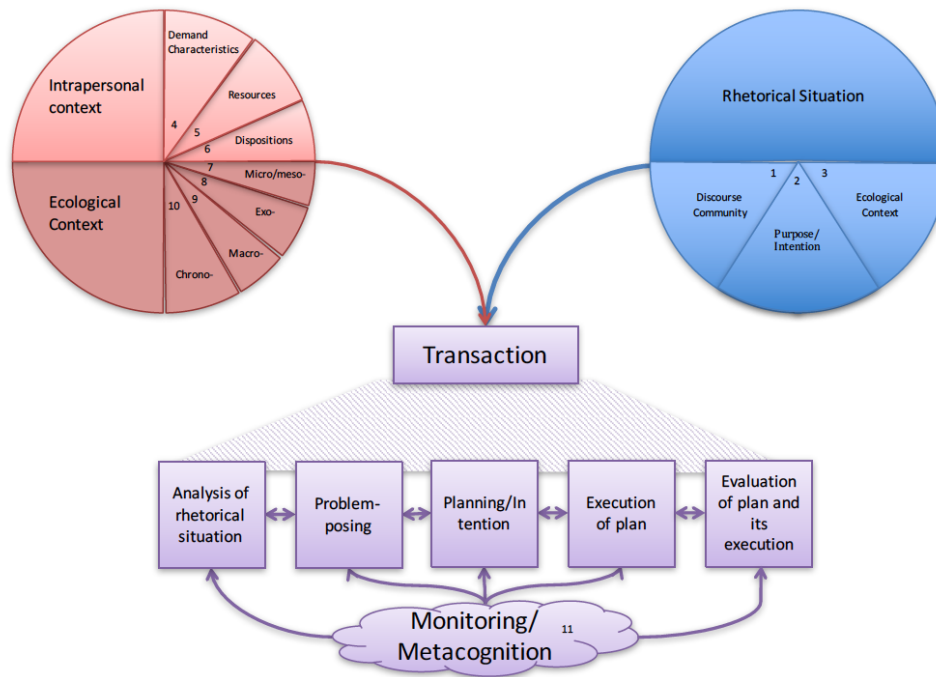
A Bioecological Model of Writing

Owing to the fact that humans are influenced by nature and nurture—and to ignore either would be doing a disservice to researchers and those whom their research benefits—many researchers within the field of writing and in social sciences more generally have advocated for theory that integrates sociocultural, cognitive, metacognitive, and a range of disparate

perspectives (Roger Beard, Myhill, Riley, & Nystrand, 2009; Bronfenbrenner & Morris, 2006; Flower, 1989; Graham & Harris, 2013).

Such was the intent with the formation of a bioecological model of writing (Figure 4-1; Corrigan & Slomp, 2017). Informed by a synthesis of 50 years worth of writing theory and construct development, my colleague and I developed a construct that not only recognized the influence the Internet has had on the practice of writing, but also that the practice and development of writing is influenced by both intrapersonal (nature) and ecological (nurture) contexts. Writers' intrapersonal contexts include their demand characteristics (e.g., personality, physical appearance, and demographic factors), resources (i.e., biopsychological liabilities [e.g., concussive injury and poor fine motor control] and assets [e.g., long-term memory, knowledge, abilities, skills, and experience]), and dispositions (e.g., motivation and self-efficacy). Writers' ecological contexts include those systems—from micro to macro to chronosystems—that situate them in time and space. This ranges from a writer's immediate context (e.g., family), to community (e.g., school and work), to broader societal norms and values that support or limit writing and its development (e.g., broadband Internet access, funding for education and the arts, free speech, and trends in standardized testing).

Figure 4-1. A Bioecological Model of Writing



Source: Corrigan & Slomp, 2017

In the model, transaction—the cognitive process of writing wherein writers construct meaning—is driven both by the writer’s context (intrapersonal and ecological) and by the rhetorical situation. The more expert the writer, the more clearly defined the rhetorical situation (Flower & Hayes, 1980, 1981; Kellogg, 2008). The rhetorical situation is defined by the discourse community, purpose/intent of the author, and the author’s ecological context (including the social [e.g., collaborator, critics, and audience] and physical context [e.g., modality of communication, transcription technology, text-and-graphics-produced-so-far, location of production and audience] in which the writing occurs).

Hypothetically, this transaction succeeds—to a greater or lesser extent—when a writer possesses the metacognitive awareness to drive the transaction, and not be driven by it. Research supports the vital role that metacognition plays in the ability of a writer to effectively and efficiently define the rhetorical situation (Kent & Wanzek, 2016a; Stolarek, 1994; Wakely et al.,

2006). Thus, the bioecological model suggests that the success of the transaction is presupposed by the writer's analysis of the rhetorical situation (Flower & Hayes, 1980; Huot, 1990), problem posing (Bereiter et al., 1988; Rice, 2015), planning (De La Paz, S., & Graham, 1997; Flower & Hayes, 1981; Stotsky, 1990), execution of the plan, and an evaluation of the plan and its execution (Hayes, 2012; Hayes & Flower, 1980). These processes are not linear, but rather occur recursively throughout the writing process.

Elsewhere, my colleague and I (Corrigan & Slomp, 2017) defined analyzing the rhetorical situation as knowing the needs of a specific audience and the specific purpose(s) for a text (Beaufort, 2007; Flower & Hayes, 1981; Ronald & Volkmer, 2015), and knowing how to address those needs in the appropriate context (Brady, 1993; Mailloux, 1989). Writers need to identify an appropriate context for their writing, making decisions about whether to construct their texts, for example, as monomodal or multimodal, formal or informal, or, disciplinary or for a general audience. In addition to analyzing the rhetorical situation, writers must pose problems. The ability to pose to pose correct and complex problems (as opposed to inappropriate and simplified ones) is a key attribute separating novices from experts (Flower & Hayes, 1981; Stolarek, 1994). Furthermore, writers are involved in planning. I borrow Flower and Hayes' (1981) broad view of planning, that is, taking abstract internal representations and turning them into concrete words on a page—or in our digital age, a variety multimodal representations (e.g., alphabetic text, graphics, video, audio, hypertext, or combinations thereof). Thus, planning involves translating (putting thoughts to words), generating ideas, organizing, and setting goals (Flower & Hayes, 1981). Next comes executing the plan. Execution involves transcription (turning mental representations into visual, spatial, and/or auditory text) via a transcription technology (e.g., pen and paper, computer, mobile device; Olive & Passerault, 2012).

Transcribed text (i.e., text and graphics produced so far) becomes a resource influencing future text production (Hayes, 2012; Leijten et al., 2014). Finally comes evaluating the plan and its execution, which it should be noted, happens throughout the writing process and not merely at the end.

Collectively, these metacognitive processes (i.e., analyzing the rhetorical situation, posing the problem, planning, executing the plan, and evaluating the plan and its execution) are known as high-level executive functions (EF; Drijbooms, Groen, & Verhoeven, 2015). By contrast, research has demonstrated the importance of low-level EF in writing including inhibition (including selective attention, sustained attention, and response inhibition); updating (storing and updating information in working memory); and, shifting (switching between tasks and mental sets; Drijbooms et al., 2015). Low-level EF underlies and supports high EF, particularly in more complex composition tasks (Drijbooms et al., 2015). Thus, in the bioecological model, low-level EF is represented by the term monitoring (purple cloud; Figure 4-1). These metacognitive processes present uniquely—or not at all—depending on the rhetorical situation and context. When sending a text message, for example, there might be little evidence of evaluating the plan and its execution. In summary, I hypothesize that it is these metacognitive processes—both low- and high-level EF—that crucially support cognitive ORW processes.

The Current Study

The purpose of this exploratory study was to compare and contrast novice and expert writers engaging in ORW on the Online Research Comprehension Assessment (ORCA), and at work and school more generally. More particularly, the study compared the cognitive and

metacognitive practices used by writers ranging from novice to expert developmental levels.

These aims prompted the following research questions:

1. What cognitive and metacognitive processes do writers engage with during ORW?
2. How do these processes differ—both quantitatively and qualitatively—between novices and experts?

Method

Overview

I invited novice and expert participants to complete an ORCA after which I guided them through a think-aloud protocol known as cued retrospective reporting (CRR, also known as stimulated recall). While there are two major types of think aloud methods—concurrent and retrospective—I chose the latter so as not to tax the participants' cognitive load, which is already challenged by the completion of the ORCA. CRR is a type of retrospective think aloud wherein “participants are instructed to report retrospectively on the basis of a record of observations” (van Gog, Paas, van Merriënboer, & Witte, 2005, p. 238). Participants' memories were augmented by playing for them video screen capture of their performance on the ORCA. Because retrospective think alouds are prone to fabrication due to their being based on memory, CRR is thought to help mitigate this shortcoming (van Gog et al., 2005). During CRR, participants watched a screencast of their performance while being prompted to describe the processes they used to complete the online research writing tasks found in the ORCA. Following this, I interviewed them about their online research writing backgrounds and I asked them to share and discuss an artefact that demonstrated their online research writing ability.

Participants

The study included $n = 10$ experts and $n = 10$ novices. Owing to the fact that this is a novice-expert study, maximum variation sampling (Cohen, Manion, & Morrison, 2013) was used in order to investigate the broad range of cognitive and metacognitive processes that writers use during ORW. Table 4-1 and Table 4-2 give an overview of the demographic characteristics describing these participants. For the purposes of this study, participants in Grades 6 to 9 were considered novices while adults who met the recruitment criteria were considered experts. Junior high school students were recruited for the novice sample of this study because the ORCA was originally designed for Grade 7 students and thus that age range was considered an appropriate level of difficulty. Further, well-established developmental research demonstrates that students at this age range are typically transitioning from the knowledge-telling to the knowledge-transforming macro-stage of cognitive development in writing skills (Kellogg, 2008). It is highly unlikely that students at this age range would fall into the knowledge-crafting (i.e., expert) stage because it typically takes over two decades to reach this level, with many adults never reaching this level at all (Kellogg, 2008). To be eligible for the study, expert participants were required to be employed full-time in the knowledge economy (i.e., service/information producing industries), be fluent in English, not have a major visual impairment, and do online research writing as part of their employment (as determined by a pre-administered survey). All adults in this study were classified as subexperts in the context of writing for the ORCA (i.e., they were writing outside of their domain of expertise) and experts when it came to the writing of their professional writing artefacts. I recruited both adult and youth participants by posting a call for participants on local Facebook groups. Some of my participants I knew professionally or personally while others were strangers. I ended up recruiting more participants than needed and

I invited additional participants to be put on a waiting list should another participant drop out. Adult participants with a diversity of writing experiences, educational backgrounds, and employment classifications were selected to promote maximum variation (Creswell, 2007). I should note that all participants who responded to my recruitment call were Caucasian, middle class, and mostly female (75%). All participants had a strong command of English and a number of them were bilingual (i.e., English and French), which is not unusual in this region of Canada.

In this exploratory study, I used the logic of replication, not sampling logic, to select an appropriate number of participants. Replication in case study research is analogous to that which is used when conducting multiple experiments in order to replicate findings. In general, 4-5 cases per study are recommended (Creswell, 2007); or, to make a theoretical replication, 6-10 cases are recommended (Yin, 2009). The latter approach was used in order to make analytic generalizations (Yin, 2009). Therefore, the study met the required number of participants to make analytic generalizations for each group of novice and experts.

Table 4-1 Demographic and ORW Characteristics of Expert Group

Pseudonym	Occupation	Age	Gender	Educational Attainment / Grade Level	Frequency of ORW	Preferred Communication Media	Preferred Research Media
Trina	Technical writer / Business analyst	51	Female	Bachelor's degree	Daily or more	Social networking sites; email	Twitter; Google alerts; news articles; research papers
Catherine	Community Outreach Lead, education programs and partnerships	30	Female	Doctorate	Weekly or more	Social networking sites; email; blogs/wikis	Twitter; professional / peer-reviewed journal articles
Robert	Nuclear safety analyst	41	Male	Master's degree	Weekly or more	Email; reports; websites	Web sites; professional publications and newsletters; reports
Lisa	Registered Massage Therapist	35	Female	College diploma	Monthly or more	Social networking sites; social bookmarking sites; blogs/wikis	News and magazine articles; meetings and/or presentations; video (e.g., television news, vodcasts) and/or Audio (e.g., radio news, podcasts)
Margaret	English Teacher, Curriculum Coach	55	Female	Bachelor's degree	Daily or more	Social networking sites; microblogs; blogs/wikis	Social networking sites; professional publications/newsletters; meetings and/or presentations
Robin	Financial Management Analyst	42	Female	Advanced degree (i.e., Chartered Professional Account)	Daily or more	Email; reports; meetings/presentations	Reports; websites; meetings and/or presentations
Laura	Director and Senior Counsel (lawyer)	52	Female	Advanced degree (i.e., LL.B)	Daily or more	Email; meetings/presentation; websites	Professional publications/newsletters; websites; dictionaries and language guides
Monica	Professor of Educational Technologies in a Faculty of Education	45	Female	Doctorate	Weekly or more	Blogs /Wikis; email; social networking sites	Professional / peer-reviewed journal articles; websites; social networking sites
Judy	Registered Dietician	34	Female	Bachelor's degree	Monthly or more	Meetings/presentations; email; pamphlets/brochures	Listservs; professional / peer-reviewed journal articles; electronic medical records
Colin	Program Engineer, Regulatory Administration	37	Male	Bachelor's degree	Weekly or more	Email	Email; reports; websites

Table 4-2 Demographic and ORW Characteristics of Novice Group

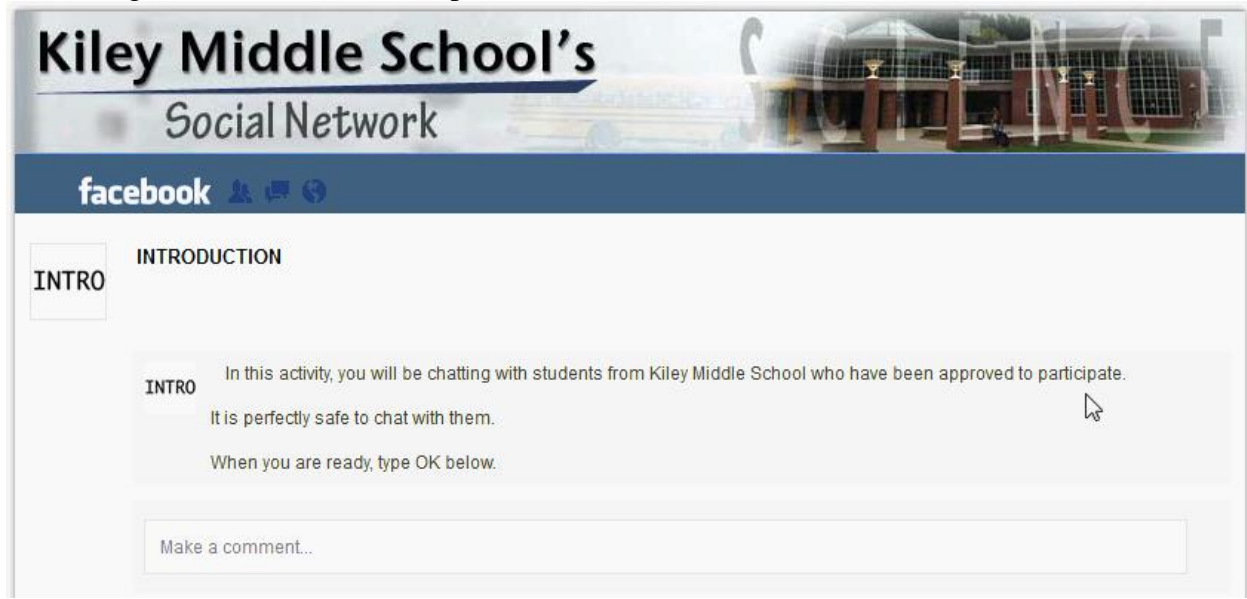
Pseudonym	Grade	Age	Gender	Frequency of Internet Use	Preferred Online Activities	Other Details
Helen	7	12	Female	20 minutes/day at school; 20 minutes/day at home	Email; online research for school projects; blogging	Fraternal twin of Jake; loves to draw and write fictional stories; uses Google Classrooms and Chromebooks at school
Jake	7	12	Male	20 minutes/day at school; 20 minutes/day at home	Web searching to learn more about something; research for school projects	Fraternal twin of Helen; won a national Remembrance Day essay writing contest
Sue	6	11	Female	20 minutes/day at school 60 minutes/day at home;	Research for school projects; listening to music; watching TV shows; computer games involving coding; Googling funny things about cats; Finding recipes	Uses iPads and Chromebooks at school; mom is a teacher; goes to a French immersion school
Jill	8	13	Female	20 minutes/day at school 20 minutes/day at home;	Research for school projects; messaging friends; occasional use of Instagram	Goes to French school; swims competitively
Cecilia	8	13	Female	60 minutes/day	Research for school projects; email	Homeschooled; sibling of Darryl; dad is a computer programmer
Darryl	9	14	Male	90 minutes/day	Email; social media (Instagram); researching school projects; gaming (MineCraft ; Mario Kart)	Homeschooled; speed reader; gamer; sibling of Cecilia
Natasha	7	12	Female	60 minutes/day at school; 60 minutes/day at home;	Playing games on iPod; YouTube; doing research for homework; email	Each class has a set of Chromebooks; goes to French school;
Fiona	7	12	Female	30 minutes/day at school (twice a week); 10 minutes/day at home	Online language translators; research for school projects; watching movies; email	Identical twin of Anna; goes to French immersion school; mom is a teacher; Dad is VP product development at a high tech firm
Anna	7	12	Female	50 minutes/day at school (a couple of times a week); 30 minutes/day at home	Research for school projects; email	Identical twin of Fiona; goes to French immersion school; mom is a teacher; Dad is VP product development at a high tech firm; class uses Chromebooks
Cody	8	13	Male	“most of the day”	Online games (e.g., Papa’s Pizzeria and Fire Ball); research for school projects	Won a spoken word contest where he talked about his experience living with Asperger’s

Online Research Comprehension Assessment (ORCA)

For this study, I used the ORCA as a means of observing participants' ORW practices in a controlled environment. The ORCA is a large-scale assessment of online research and writing designed for seventh graders. The ORCA was developed at the University of Connecticut's New Literacies Lab as a part of the ORCA Project grant from the Institute of Education Sciences (Leu et al., 2009-2014). Currently, the ORCA is not a high-stakes test. Rather, data from these assessments are used to inform research, provide feedback to teachers and students, and to inform the development of professional development resources. An overview of the ORCA including video capture of some its main features can be found [here](#).

There are eight ORCA scenarios, with each scenario focussing on a different research problem from the domain of health and human body systems, an area common to many seventh grade science curricula. The ORCA scenario selected for this study posed the research problem "Do decorative contact lenses harm your eyes?" Cosmetic contact lenses, such as those used by [Lady Gaga](#) in her "[Bad Romance](#)" video, are lenses designed to alter the colour or appearance of a person's eyes. Each ORCA scenario is framed around one of two types of research (Learn More About or Investigate Conflicting Claims) and one of two types of digital writing platforms (email or wiki). For this study, I used an Investigate Conflicting Claims (ICC) version wherein the participants' final written product took the form of a wiki. The research problems appeared within a Facebook-like environment via avatars named Brianna and Jordan (Figure 4-2) who are introduced as students from another school. Students were guided to engage in ORW via requests and questions from Brianna and Jordan. A video of a high-performing participant taking the ORCA-Closed assessment, "Are Energy Drinks Heart Healthy?" may be viewed by linking to this URL: <http://neag.uconn.edu/orca-video-ira/>

Figure 4-2. Facebook-like platform used in ORCA.



The ORCA is designed to measure online research writing ability in four skill areas, three of which are research focussed (locating, evaluating, and synthesizing online information) and one of which is writing focussed (communicating online information). These skill areas were determined through a systematic review of the literature, and ORCA test items were designed and revised following numerous think-aloud sessions held over the course of the ORCA's five-year development (Leu, Kulikowich, Sedransk, & Coiro, 2014). Each version of the ORCA consisted of sixteen items, with four component items in each of the research and writing skill areas named Locate, Evaluate, Synthesize, and Communicate (LESC).

Assessment tasks did not appear in a linear sequence according to skill area (i.e., LESCE). Rather, a more natural and logical sequence was used according to the nature of the research task. The ORCA began with Brianna introducing herself and inviting participants to collaborate in solving an online research problem, in this case, "Do decorative contact lenses harm your

eyes?” Participants were also given a general overview of the assessment, in which they would be asked to read online information and present their findings on a classroom wiki for Mr. Henry’s class, a fictional class from a neighbouring school. Following this, participants were instructed to read an email from a school board president, which provided greater context to the task. From there, Brianna directed students to investigate conflicting claims about cosmetic contact lenses. At this point, participants are directed to a simulated search engine called “Gloogle” (Figure 4-3), where they were to locate and read four online articles conveying multiple viewpoints pertaining to the safety of cosmetic contact lenses (for example, an online article from the *New York Times* entitled “[What Big Eyes You Have, Dear, but Are Those Contacts Risky?](#)”). While reading the four articles, participants were instructed to summarize the main points from each article using a virtual notepad within the ORCA. Towards the end of the assessment, participants were asked to synthesize the information from all four articles. Then, they were asked to evaluate the credibility of one of their online sources by answering questions posed by an avatar named Jordan. The final task of the ORCA invited participants to take a position regarding the safety of cosmetic contact lenses and to communicate their findings via Mr. Henry’s classroom wiki (Figure 4-4).

Figure 4-3. "Gloogle" Search Engine used in ORCA

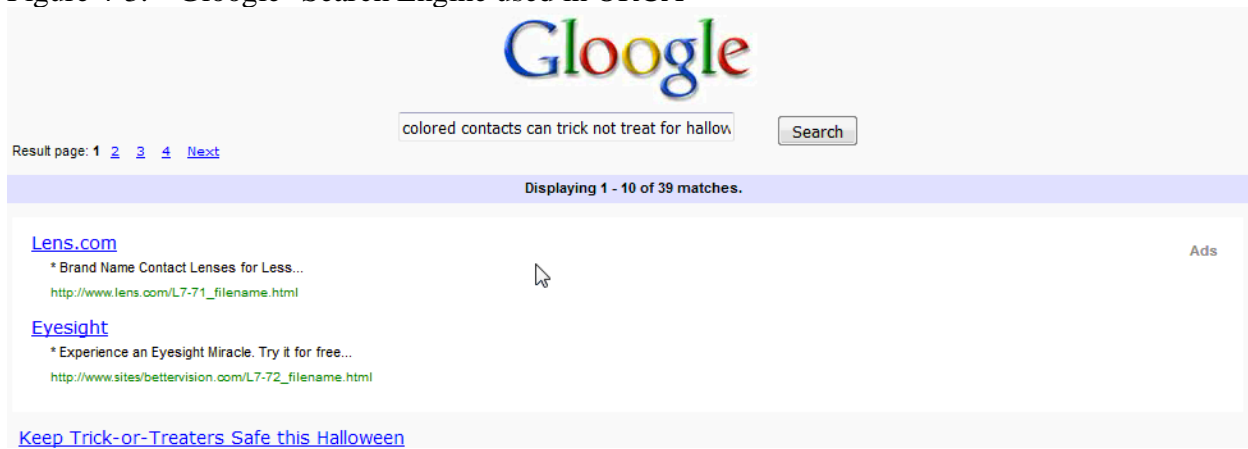
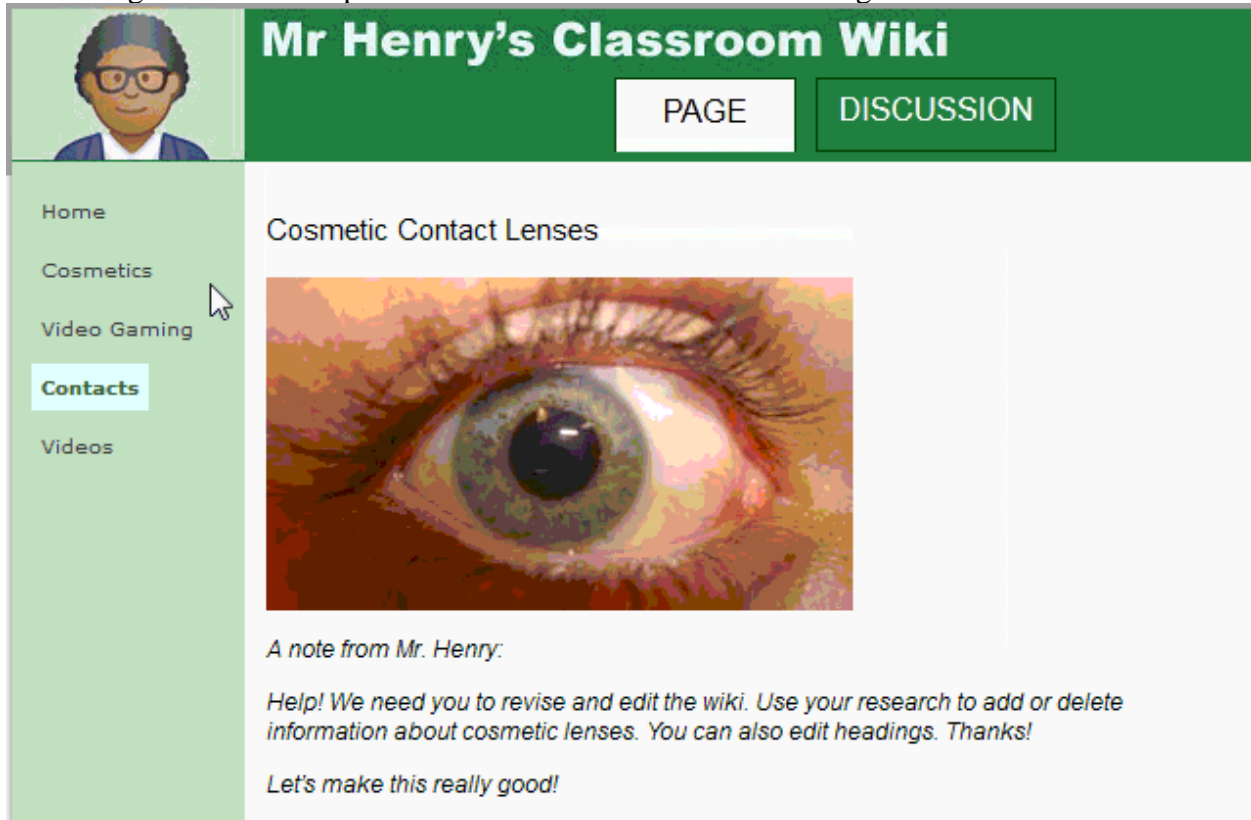


Figure 4-4. Participants communicate their research findings via wiki.



While the ORCA was developed using three different format types, this study used the ORCA Virtual, which enabled the analysis of data in a closed (i.e., controlled) Internet environment, and which is a direct writing assessment (as opposed to the multiple choice version; please see Leu et al., 2012, for further details regarding the versions of the ORCA). In the Virtual format, students conducted their research and selected information from pre-determined websites from the project's search engine named "Gloogle." As the tasks were in an open-ended format, the ORCA Virtual is a performance-based measure.

An auto-capture system recorded students' responses. Anytime a student typed a response to a task in the ORCA, it was automatically recorded in an online database. Video screen captures recorded students' performance as a backup for the auto-capture and to use during CRR.

Procedures

I used a protocol to ensure replication of procedures from participant to participant in order to establish transparency and replicability in my findings (Appendix A; Yin, 2009). I invited participants one at a time to a private computer lab at a large university in Ontario, Canada. I asked participants to read through the consent/assent form and I gave them the opportunity to ask any questions. For minors, I also obtained parental consent. Next, I read participants a brief set of instructions, as per the protocol, and then they completed the ORCA on a PC with a 17" monitor. While participants completed the ORCA, I monitored them from a remote station to ensure that data were being captured properly (i.e., video screen capture), and to answer any technical questions. I also captured eye-tracking data, but chose not to use this data as the stimulus proved too long to reliably capture data (i.e., participants moved their heads too much while taking the ORCA). I instructed participants that I could not answer any questions relating to the tasks on the ORCA. Built in prompts throughout the ORCA reminded the participants that they had a limited time to complete various segments, although I advised participants that they could ignore those prompts. Duration of the ORCA for participants in this study ranged from 25 minutes to an hour.

Cued Retrospective Reporting (CRR). Following their completion of the ORCA and a short break, I instructed participants in thinking aloud, a method in which participants verbalize the thoughts they had during problem solving. Research shows that thinking aloud is not always intuitive, and that having it modeled beforehand helps to improve the quality of the feedback from participants (Holmqvist et al., 2011). Prior to beginning the CRR process, I instructed participants as follows: "I'm going to ask you to constantly say out loud whatever was going on in your mind as you worked. Please don't just read aloud the questions and say your responses.

Instead, tell me everything that happened in your head as you worked on the problem. You can say anything that was on your mind. If you are silent, I will say something like ‘Please keep talking.’ As previously explained, during CRR participants viewed a video screen capture of their performance during the ORCA. I paused the video following each task and cued participants to verbalize their problem-solving processes. Participants’ responses to the CRR and subsequent interview were audio recorded and later transcribed for research purposes.

Semi-structured interviews. The final activity the participants were invited to complete was an interview. The goal of the interview was to learn more about how the participants engaged in ORW in their personal and school/work lives. Participants were asked to bring samples of their online research writing in order to elicit additional cognitive and metacognitive ORW practices that may not have been called upon during the ORCA. This helped ensure that I more fully captured the construct. In a manner similar to the CRR, I asked participants to describe and explain the problem solving approaches they took in writing their school or work artefact. The interview followed a semi-structured approach, as outlined in the protocol (Appendix A).

Data Sources and Analyses

I collected both quantitative and qualitative data throughout this study in order to “compensate for inherent method weaknesses, capitalize on inherent method strengths, and offset inevitable method biases” (Greene, 2007, p. 120). Specifically, I mixed methods for the purposes of triangulation and initiation. While the focus of triangulation tends to be on consonance (using different methods to confirm and corroborate results), mixing methods with the purpose of initiation evokes dissonance (using different methods to see where results diverge and contradict one another, all in the service of developing novel perspectives and

understandings; Greene, 2007). My data analysis process involved examining participants individually, then within groups (i.e., students and knowledge workers), and finally between groups (i.e., students vs. knowledge workers).

ORCA Results. ORCAs were scored using both an analytic rubric developed for the ORCA project and a holistic rubric developed for the current study. An auto-capture system recorded participants' responses for later scoring. Each ORCA was hand scored by the author as well as an additional scorer who had played a major role in training scorers for the ORCA Project. A scoring guide developed for the ORCA project was used to improve inter-rater reliability. In cases where the scorers did not reach an inter-rater reliability level of 90%, the scorers discussed the discrepancy and reached a consensus as to the most appropriate score. First, ORCAs were scored on an item-by-item basis using an analytic rubric (Appendix B, Table 1). The rubric was divided into four categories (Locate, Evaluate, Synthesize, and Communicate) and each of the four categories consisted of four items for a total of 16 items reflecting tasks found on the ORCA. The rubric consisted of dichotomous scoring with participants scoring either a 1 or a 0 for a total of 16 possible points. A total ORCA unidimensional score was calculated for each participant by summing the scores for each of the 16 items. Additionally, multidimensional scores were calculated for each of the locate, evaluate, synthesize, and communicate domains by totalling the sum of the scores from each of these scores. Second, each wiki response was hand scored using a holistic rubric modified from the TOEFL iBT Test (Appendix B, Table 2). The preceding analyses resulted in two scores: ORCA (from the analytic rubric) and Holistic (from the holistic rubric). The major difference between these measures is that the first focuses on the cognitive processes of ORW (including locating, evaluating, synthesizing, and communicating online information) while the latter focuses on the

product (an overall evaluation of the wiki, a product of ORW). Both the ORCA and holistic scores were rescaled to range between 0 and 1 to facilitate comparisons and for further statistical analyses.

CRR and interview data. Coding of CRR and interview data involved three coding cycles. In cycle one, I read the CRR and interview transcripts from all participants to get a sense of the data. Then, I applied provisional codes, which refers to the process of beginning with researcher-generated codes based on preliminary research and revising, modifying, deleting, and expanding the codes as necessary (Miles, Huberman, & Saldana, 2014; Saldana, 2013). The data were coded using two broad categories of codes based on previous studies: cognitive ORW practices (locating, evaluating, synthesizing, and communicating; (Leu et al., 2014) and metacognitive writing practices (analyzing the rhetorical situation, posing the problem, planning, executing the plan, and evaluating the plan and its execution; Corrigan & Slomp, 2017). Additionally, I allowed for new codes and child codes to emerge throughout the analysis. Each time a participant uttered a skill, strategy, disposition, or social practice relating to one of these cognitive or metacognitive practices, it was coded as one episode. A codebook was created and revised throughout the study detailing the code's name, its descriptor, inclusion and exclusion criteria, as well as typical exemplars. Because the interview immediately followed the CRR, and because participants referred back to the ORCA while discussing their personal and school/work writing, I kept the CRR and interview transcript as one data source.

Second cycle coding involved magnitude coding (Saldaña, 2013) of the five metacognitive themes previously mentioned. For each participant, I read all of the codes within each theme (e.g., problem posing) and gave that participant's metacognitive process a score on a scale of 1 (lowest) to 4 (highest) based on their expertise. For example, novice participant Cecilia told me,

“So I just typed zebras in the search and opened a whole bunch of links,” and this was given an expertise score of 1. Meanwhile, expert participant Julie described how she strategically searched for information with her audience in mind, conducting different searches using different databases and/or search engines based on the needs of her stakeholders. This was given an expertise score of 4. In mixed methods research, this procedure is known as quantizing, a process wherein qualitative, textual data are transformed into quantitative, numerical data for the purposes of data integration and analysis (Sandelowski, Volis, & Knafl, 2009). Additionally, an overall Metacognitive ORW score was calculated for each participant by summing the scores from each of the five metacognitive process sub-scores. Once again, these scores were rescaled to range between 0 and 1 to facilitate comparisons. [Dedoose](#), a computer assisted mixed methods data analysis software program, was used to code all transcripts and the ORCA descriptive data described above.

Writing artefact analysis. Each student and knowledge worker participant was asked to bring an artefact that typified his or her experience with online research writing. Expert artefacts included a diversity of texts such as legal decisions, financial forecasts, nuclear regulatory reports, blog posts, emails, and patient records. Meanwhile, novice artefacts were a variety of school assignments, from a spoken word project about living with Asperger’s syndrome, to an essay about why we need to banish rules at recess to research projects on topics such as polar bears and how to make chocolate. Participants were asked to redact any confidential or proprietary information contained within their artefacts prior to sharing them with me. I read the artefacts prior to the interviews in order to provide greater context for semi-structured interviews. I wrote descriptive notes for each artefact, and later coded both the artefact and my notes in

terms of ways in which these aligned with the cognitive and metacognitive themes previously noted.

Demographic data. I collected demographic data from two sources: pre-administered surveys (adult participants only) and during the interviews. For all participants, I collected data regarding age, gender, educational attainment/grade level, and frequency and type of research, writing, and online research writing. For adult participants, I also collected data regarding their occupation, number of years of employment, and ORW practices.

Statistical analysis. In order to understand the ways in which novices and experts differed among measures of ORW more generally, and cognitive and metacognitive measures more specifically, a variety of statistical analyses were undertaken. To begin, a one-way multivariate analysis of variance (MANOVA) was conducted to determine whether experts and novices performed differently on the three previously described measures of ORW, specifically, the ORCA, Metacognitive, and Holistic measures. Next, Mann Whitney U tests were conducted to investigate the differences between novices and experts with regard to the cognitive domains of the ORCA (Locate, Evaluate, Synthesize, and Communicate) and the domains of the Metacognitive score (Analyzing the Rhetorical Situation, Posing the Problem, Planning, Executing the Plan, and Evaluating the Plan and its Execution) derived from the CRR and interviews. Mann Whitney U tests were used in these instances because these data violated assumptions of normality, as will be shown.

Results

In this study, I investigated the following research questions: (1) What cognitive and metacognitive processes do writers engage with during ORW? (2) How do these processes differ—both quantitatively and qualitatively—between novices and experts? In brief, analyses

showed that participants engaged with a wide variety of cognitive and metacognitive processes during ORW, many of which appear to be unique to online contexts. Further, both quantitative and qualitative analyses demonstrated that there were differences in the ORW abilities of novices and experts, both generally, and with regards to their cognitive and metacognitive processes.

Quantitative Results

Overall ORW differences between novices and experts.

Prior to running a one-way MANOVA to determine whether experts and novices performed differently on measures of ORW, preliminary assumption checking was conducted and revealed minor violations of assumptions. Specifically, an inspection of boxplots to check for univariate outliers determined that there were three outliers (greater than 1.5 box-lengths from the edge of the box), but no extreme outliers (3 box lengths or more; Figure 4-5). In terms of normality, scores (specifically the expert ones) were moderately positively skewed for the ORCA, CRR, and holistic measures, as assessed by a Shapiro-Wilk's test ($p > .05$) and by skewness and kurtosis values. After comparing the results of a MANOVA with the original data to those of the transformed data (i.e., following a square root transformation), a decision was made to keep the outliers and original data because the transformation did not lead to any appreciable difference in the results. There was no multicollinearity, as assessed by Pearson correlation (Table 4-3). An inspection of scatterplots revealed a positive linear relationship between the ORCA and holistic writing scores and the CRR and holistic writing scores. However, the relationship was less obvious between the ORCA and Metacognitive scores. There were no multivariate outliers in the data, as assessed by Mahalanobis distance ($p > .001$). There was a violation of the homogeneity of variance-covariance matrices, as assessed by Box's test of

equality of covariance matrices ($p = .57$); however, the sample sizes were equal thus mitigating the issue.

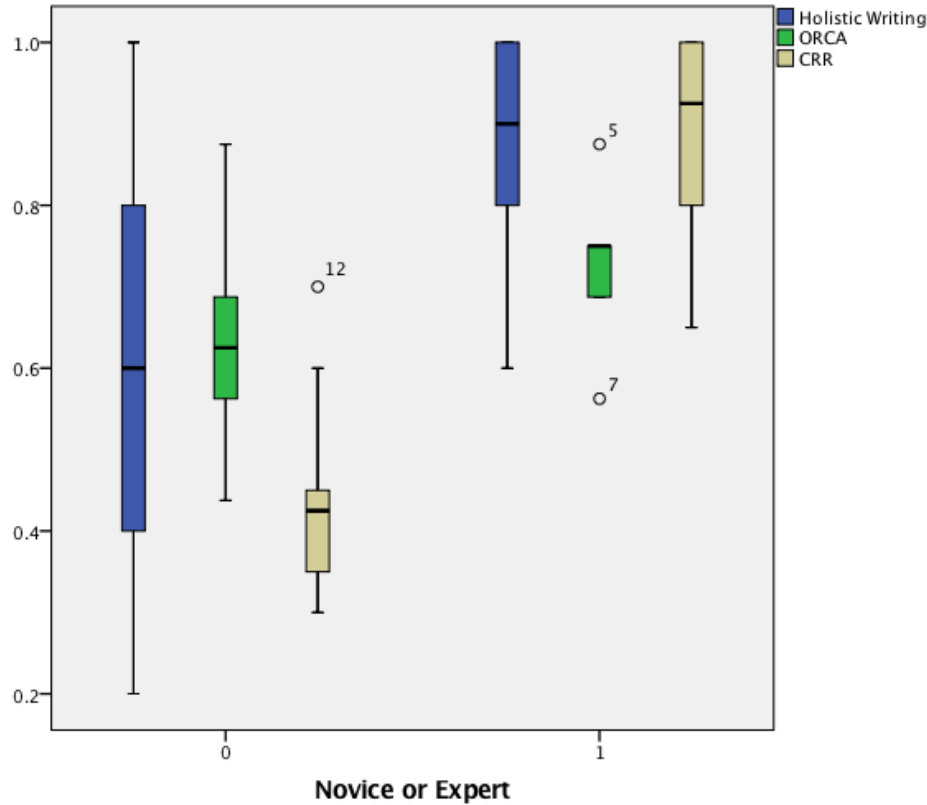
Table 4-3 Correlations among ORW Measures (N = 20)

	1.	2.	3.
1. Holistic ORW Measure	--		
2. ORCA	.488*	--	
3. Metacognitive ORW Measure	.701**	.359	--

* $p < .05$. ** $p < .01$.

Following assumption checking, a one-way MANOVA was performed to investigate differences between novices and experts on three ORW measures, namely the Holistic, ORCA, and Metacognitive measures. Experts scored higher than novices across all measures, including the holistic ($M = .86, SD = .06$; $M = .60, SD = .06$), ORCA ($M = .73, SD = .03$; $M = .64, SD = .03$), and metacognitive ($M = .90, SD = .04$; $M = .45, SD = .04$, respectively) measures. There was a statistically significant difference between novices and experts on the combined dependent variables, $F(3, 16) = 28.62, p < .0005$; Pillai's Trace = .84; partial $\eta^2 = .84$. When the results for the dependent variables were considered separately, two measures achieved statistical significance, using a Bonferroni adjusted alpha level of .017. These were the Holistic [$F(1, 18) = 8.40, p < .01$; partial $\eta^2 = .32$] and Metacognitive [$F(1, 18) = 72.43, p < .001$; partial $\eta^2 = .80$] measures. However, there was no statistically significant difference for the ORCA, $F(1, 18) = 4.63, p < .05$; partial $\eta^2 = .21$.

Figure 4-5. Differences between novices and experts on three ORW measures



Domain differences between novices and experts.

Metacognitive differences.

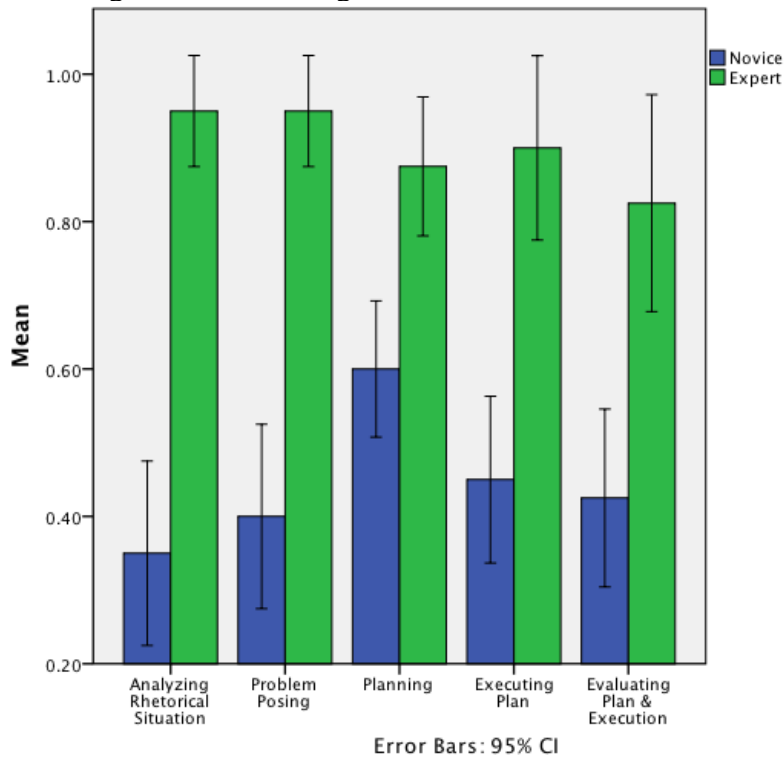
Mann-Whitney U tests were also run to determine if there were differences in metacognitive domain scores (Analyzing the Rhetorical Situation, Problem Posing, Planning, Executing the Plan, and Evaluating the Plan and Its Execution) between novices and experts. Distributions of the domain scores for novices and experts were not similar, as assessed by visual inspection. Experts scored statistically significantly higher than novices on all metacognitive domains, as per Table 4-4 and Figure 4-6.

Table 4-4 Differences in metacognitive domains between novices and experts
Differences in metacognitive domain between novices and experts

	Mean Rank				
	Novice	Expert	U	z	p
Analyzing the Rhetorical Situation	5.60	15.40	99.00	3.92	.001***
Problem Posing	5.60	15.40	99.00	3.88	.001***
Planning	6.50	14.50	90.00	3.25	.002**
Executing the Plan	6.00	15.00	95.00	3.57	.001***
Evaluating the Plan and its Execution	6.35	14.65	91.50	3.26	.001***

* $p < .05$; ** $p < .01$; *** $p < .001$

Figure 4-6. Metacognitive domain differences between novices and experts

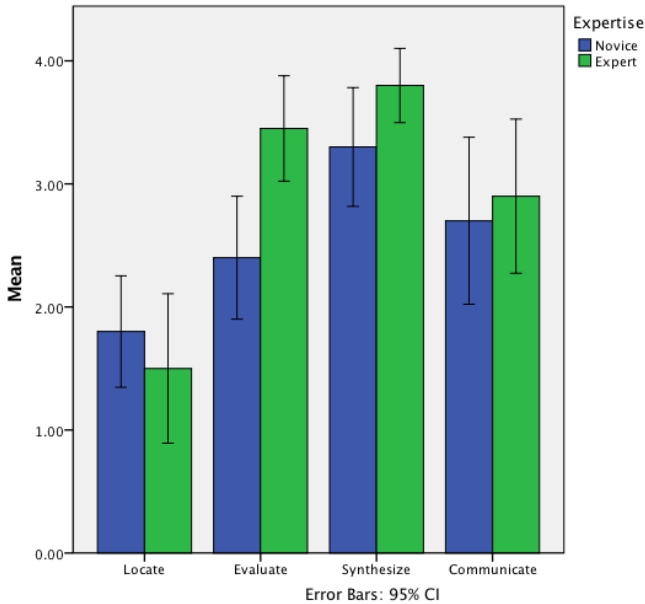


Cognitive differences.

Mann-Whitney U tests were also run to determine if there were differences in cognitive domain scores (Locate, Evaluate, Synthesize, and Communicate) between novices and experts.

Distributions of the domain scores for novices and experts were not similar, as assessed by visual inspection. Evaluate scores for experts (mean rank = 14.00) were statistically significantly higher than for novices (mean rank = 7.00), $U = 35.50$, $z = -1.20$, $p = .23$. Expert scores for Synthesize (mean rank = 12.60) were also higher than for novices (mean rank 8.40), though the difference was not statistically significant, $U = 71.00$, $z = 1.84$, $p = .12$. The same was true of Communicate scores between experts (mean rank = 11.00) and novices (mean rank = 10.00), $U = 55.00$, $z = .45$, $p = .66$. Locate scores, conversely, were lower for experts (mean rank = 9.05) compared to novices (mean rank = 11.95), though this difference was not statistically different, $U = 35.50$, $z = -1.20$, $p = .23$. I believe this was due to technical problems with the communicate task, as I will explain later. Figure 4-7 illustrates these differences. Both experts and novices found scored lowest on the locate domain ($M = .38$; $M = .45$) and highest on the synthesize domain ($M = .95$; $M = .83$). In order from the lowest to the highest scoring domains for novices were locate, evaluate, communicate, and synthesize. For experts, the order from lowest to highest was locate, communicate, evaluate, and synthesize.

Figure 4-7. Cognitive Domain Differences between Novices and Experts



Qualitative Results

As was seen in the quantitative results, differences between novices and experts existed on both metacognitive and cognitive measures, and their associated domain measures. These findings were corroborated by qualitative findings. In this next section, I present my qualitative findings regarding how experts and novices differ between their metacognitive and cognitive processes. As metacognitive processes support cognitive ones and are intertwined (Veenman et al., 2006), I present these results concurrently, embedding the cognitive processes (i.e., locate, evaluate, synthesize, and communicate) into the metacognitive processes (i.e., analyzing the rhetorical situation, problem posing, planning, executing the plan, and evaluating the plan and its execution) described in the bioecological model of writing. It should be noted that while these processes are described linearly, they were enacted recursively. In other words, during the CRR, participants described moving among numerous cycles of these processes in rapid succession.

Table 4-5 and Table 4-6 highlight the findings of the cross-case analyses of the metacognitive and cognitive differences between novices and experts with regards to ORW.

Additionally, Appendices C - G provide interview and CRR excerpts demonstrating these differences.

Table 4-5 Case summaries from cross-case analysis of processes in ORW

Metacognitive Domain	Experts	Novices
<i>Analyzing the Rhetorical Situation</i>		
<i>Knowing the purpose of a text</i>	Possessed a clear sense of purpose or even multiple purposes in relation to the needs of the audience	Possessed a superficial sense of purpose that ORW is for a school grade
<i>Anticipating the needs of the audience</i>	Considered the audience when determining diction, tone, genre, and rhetorical strategies	Described as finding out what the teacher wants
<i>Anticipating the objections of the audience</i>	Took on the persona of their audience in order to imagine how the text might be challenged	n/a
<i>Analyzing the historical context</i>	Considered the history of problem before presenting solution; considered intertextuality	n/a
<i>Analyzing the rhetorical situation multimodally</i>	Used pictures, hypertext, multimodal design features; considered differences between online and offline genres; considered digital footprint	Used pictures
<i>Analyzing the rhetorical situation in order to read online texts</i>	Considered the reciprocal nature of constructing and deconstructing online texts; Understood online design features and how to manipulate them (e.g., following hyperlinks to find out more about the author; using Boolean search operators; understanding the non-linear nature of online writing)	n/a
<i>Problem Posing</i>		
<i>Problem posing</i>	Posed correct problems, often complicating them.	Variously posed correct and incorrect problems; often simplified problems
<i>Planning</i>		

<i>Translating and generating ideas</i>	Experts addressed all novice processes, in addition to annotating source material, thinking over a period of time, having a daily writing routine, referring to texts that one previously wrote or read (texts often stored in reference managers), using technology, and following social media on subjects of interest.	Generated ideas by reading source material, free writing, using graphic organizers, using mediating documents, talking to others, writing about interests, and writing about experiences.
<i>Goal setting</i>	Considered short-, medium-, and long-term writing goals; breaking down tasks into sub-tasks; reflecting on previous experience with the task	Focused on short- and medium-term writing goals; Breaking down tasks into sub-tasks
<i>Triaging</i>	Determined how much time to allocate to ORW task based on importance to self and others	Determined how much time to allocate to ORW task based on importance to self
<i>Organizing</i>	Followed the genre conventions appropriate to their profession (e.g., international dietetic standards, standards for the nuclear industry, APA standards)	Followed prescribed format (e.g., five paragraph essay)

Executing the Plan

<i>Summarizing, synthesizing, and generative synthesis</i>	Summarized, synthesized, and synthesized generatively, putting an emphasis on generating new knowledge;	Summarized and synthesized, relying heavily on paraphrasing
<i>Assuming a stance</i>	Considered multiple, competing perspectives and weighed the evidence considering its credibility; took more nuanced stances	Considered evidence from one perspective while discounting evidence from competing perspectives; took binary stances (yes or no answers)
<i>Using rhetorical strategies</i>	Used diction, rhetorical devices (e.g., metaphor), pictures, video, figures, tables, hypertexts, elements of web design, and aesthetic choices to make arguments more compelling; considered rhetorical strategies	Used diction and pictures to make arguments more compelling

	appropriate to the discourse community and genre	
<i>Using a linear / recursive process</i>	Used recursive writing processes	Used linear writing processes (e.g., research, brainstorm, write, revise)
<i>Writing collaboratively</i>	Collaborative writing was necessitated by the complexity of the writing task; tasks (e.g., writing safety standards for a nuclear facility; patient records written by a family health care team) required distributed expertise; process tended to be asynchronous and necessitated project management; attention was given to maintain a unified voice, typically through the use of style guides	Collaborative writing was dictated by the teacher; process was typically synchronous
<i>Digital remixing</i>	Repurposed their own texts and those of others; applied previously written information to a new context; used reference management software to manage digitally stored texts	Used pictures from online sources; paraphrased from online sources

Evaluating the Plan and its Execution

<i>Revising</i>	Focussed on local, but especially global revision; Offset writing/researching weaknesses by seeking out peers with complementary expertise	Focussed on local revision
<i>Transfer</i>	Reflected so as to improve current or future ORW	n/a

Table 4-6 Case summaries from cross-case analysis of cognitive processes in ORW

Cognitive Domain	Expert	Novice
<i>Locating Online Information</i>		
<i>Primary Locating: Locating Sources of Information</i>	Used conventional search strategies including key word searches; title searches; narrowing search by date, location, and name of publication; and reading the resume in the search engine results	Also used conventional search strategies including key word searches; title searches; narrowing search by date, location, and name of publication; and reading the resume
	Used more sophisticated search strategies including using Boolean operators; using divergent sources of information (e.g., Twitter feed; reference managers; listserv; professional/expert blog posts); backward and forward reference searching; second-level backward reference searching; and, backward author searching	Used less sophisticated search strategies such as clicking on the top hit; using trial and error; and using very generalized key word searches
<i>Secondary Locating: Locating Information within the Source</i>	Skimmed texts in order to efficiently identify relevant information	Read the whole text
	Identified relevant information	Became distracted by irrelevant information and click bait
<i>Metacognitive Practices: Using Metacognitive Practices to Locate Information</i>	Engaged in a broad range of metacognitively-supported search strategies including triaging; using technology to manage information; remixing information for new purposes; locating primary source information (being sceptical of secondary sources); and locating information with an eye to the audience for whom one is writing	Engaged in emergent metacognitively -supported search strategies

Evaluating Online Information

<i>Primary Evaluation: Evaluating Argument Credibility</i>	Viewed arguments (including their claims, evidence, and warrants) with scepticism	Viewed arguments (including their claims, evidence, and warrants) with naivet�
	Questioned the degree of the source of the information (e.g., primary vs. secondary source)	Took information from secondary sources at face value
	Viewed information credibility as a matter of degree	Viewed information credibility as a binary (right or wrong)
<i>Secondary Evaluation: Evaluating Source Credibility</i>	Applied scepticism toward sources (author, organization, and/or publisher) such as questioning the author’s education, expertise, and purpose for writing; background checking the author and/or organization	Applied emerging scepticism towards sources (author, organization, and/or publisher) such as by examining the URL; looking for spelling and grammatical errors in online sources as a way to discredit them)
	Used sophisticated strategies to verify source credibility such as following hyperlinks both within the website (clicking on the author’s name to learn more) and without (background checking via sources not hyperlinked to the text being verified)	Used less sophisticated strategies to verify source credibility (only two out of 10 participants clicked on the hyperlink of the author’s name and none checked external sources)
<i>Metacognitive Practices: Using Metacognitive Practices to Evaluate Information</i>	Used sophisticated and diverse metacognitive strategies to evaluate online information (e.g., triangulating evidence; considering multiple, diverse perspectives; triaging; backward and second-level backward reference searching)	Used emergent metacognitive strategies to evaluate online information (e.g., background reading; backward reference searching; occasionally triangulating information)

Synthesizing Online Information

<i>Summarizing</i>	Paraphrased information in more original language; summarized key ideas	Paraphrased directly; summarized key ideas
--------------------	---	--

<i>Synthesizing</i>	Weighed credibility of evidence from contradictory sources when synthesizing information; considered multiple perspectives	Ignored sources that did not corroborate their own position; considered a singular perspective
<i>Generative Synthesizing</i>	Used extant information to generate new ideas	n/a

Communicating Online Information

<i>Communicating Critical Discourse Knowledge</i>	Communicated with a sense of promoting equity and social justice; wrote accessibly; considered intertextuality and historical perspectives	n/a
<i>Communicating Discourse Knowledge</i>	Communicated with a strong sense of disciplinary, professional, and/or community norms and values	Communicated with little to no sense of community norms and values
<i>Communicating Rhetorical Knowledge</i>	Chose diction to communicate rhetorical purpose, but also with a sense of discourse community values	Chose diction to communicate rhetorical purpose
	Used a range of multimodal rhetorical strategies including using video, using hyperlinks, elements of web design, and aesthetic choices in online writing	Relied on images as a multimodal rhetorical strategy
<i>Communicating Genre Knowledge</i>	Chose genre by considering the rhetorical situation (e.g., deciding to send client an email rather than a formal legal decision); used genre conventions as appropriate to the community/profession	Relied on school-based genres (e.g., five paragraph essay) with little consideration given to rhetorical situation
<i>Using Writing Process Knowledge</i>	Used a recursive process Used collaborative writing processes necessitated by distributed expertise; complex collaborative writing processes required project management	Used a linear process Engaged in teacher prescribed collaborative writing processes; little project management

	Collaborative writing was asynchronous	Collaborative writing was synchronous
<i>Communicating Subject Matter Knowledge</i>	Communicated subject matter knowledge with consideration given to the norms and values of the discourse community	n/a

Analyzing the rhetorical situation.

One of the first prompts participants encountered during the ORCA was to visit the wiki page of Mr. Henry's class, the site in which they would eventually share the results of their research findings regarding whether or not cosmetic contact lenses are harmful to eye health. During the subsequent CRR, all of the experts discussed how they took time at that moment to analyze the rhetorical situation (e.g., knowing their purpose for writing and knowing the needs of their audience) of the of the class wiki, while for most of the novices, visiting the class wiki was perfunctory. Further, during the subsequent interview, experts were more likely to discuss the importance of analyzing the rhetorical situation during the composition of their work writing artefacts, while fewer novices mentioned doing this for their school writing. Trina, a technical writer and one of the expert participants, says she always asks the same two questions when writing: "Why am I writing this? Who is it for?" Monica, a professor and another expert participant, echoed that sentiment: "I can't write without understanding a communicative purpose or an audience."

Knowing the purpose of a text.

One of the most common ways in which participants discussed analyzing the rhetorical situation was by knowing the purpose(s) of the texts they were writing. While nine of experts discussed having a purpose for their writing, novices solely discussed their purpose as being a required school assignment. Laura, a lawyer by profession, described how during the ORCA she kept reminding herself of the original task, "which was are you going to argue that decorative contacts are risky or not?" During his interview, Robert, a nuclear analyst, had a very clear idea of the purpose behind his work writing, which was to write company-wide processes, procedures, and standards to ensure greater safety and compliance at the nuclear plant.

Similarly, Catherine, who creates programs in partnership with Indigenous Peoples for an educational technology company, clearly articulated two purposes behind her writing artefact: “It’s the showcase piece for my first blog at my new job” and “I need to demonstrate the value of using video games to teach critical pedagogy.” Even when the communication of research results was oral, expert participants still had a clear idea of their purpose. Lisa, a registered massage therapist, said she spends time at work researching particular injuries, learning how to treat them, and learning how to prevent them. She said she knows she has a brief window while she is massaging her clients to educate them about how to avoid getting injured in the future.

Anticipating the needs of the audience.

A second way in which participants demonstrated knowledge of the rhetorical situation was by anticipating the needs of their audience. While nine of experts discussed anticipating the needs of their audience, only three of novices did so. For example, Monica discussed how while doing the ORCA she learned what she could about the community for whom she would be writing, Mr. Henry’s class, deferring to her knowledge as a former teacher. She decided that when she wrote her wiki entry, she wanted to keep her language accessible and engaging. Similarly, teacher and literacy coach Margaret spoke about how she considered her audience: “Because it’s a student wiki, and students are going to be drawn by celebrities and model their behaviour, I wanted to get students to think critically about the risk they put themselves at when they use these circle lenses that aren’t regulated by the FDA.” Here is Margaret’s wiki response:

Cosmetic Contact Lenses

According to several articles found online ("Lady Gaga's Contacts Are Risky For Eyes" and "Contact Lens Risks and News"), cosmetic contact lenses can be harmful to your eyes. However, just as regular contact lenses may harm your eyes, the problem lies in the behaviour of individuals who choose to purchase them

without getting them fitted by a professional, and without seeking the advice of a professional regarding their own eye needs. According to "What Big Eyes You Have, Dear, but Are Those Contacts Risky?", cosmetic lenses are available in an unregulated market, and this makes them a potential health hazard. They are not monitored by the Food and Drug Administration making them a possible risk to eye health. However, "Avoid the tricks that come with the Treats: Tips for a healthy, happy Halloween" points out some steps to minimize the risk of cosmetic contact lenses.

Colin demonstrated how he anticipates the needs of his clients as a program engineer working on emissions regulations for the federal government. Colin's writing artefact was an email he sent to a client regarding what he described as a regulatory submission that the client had sent him for review: "I'm going over the deficiencies of their submission to help them fix it. You need to fix these things because it isn't satisfactory and doesn't meet the regulatory requirements. I'm trying to be specific and pointed about what corrections need to be made and why there's an issue." When writing such emails, Colin researches the regulations specific to his client's context and provides hyperlinks to these regulations, as well as supplementary material.

Laura discussed how it was common practice at her legal practice to communicate to clients their legal options via email, instead of by a formal written opinion. Formal written opinions are more resource intensive, she said, but mainly she and other lawyers stay away from them because they are not what clients are looking for: "In a formal written opinion, you're going to set out the facts, you're going to provide the legal framework, you're going to do a legal analysis where you apply the facts to the framework, and then you're going to have a conclusion. Meanwhile, the client just wants the bottom line. They don't care about all of your analyses. Not really. Often, they just want the bottom line so they can get on with their business." Emails, she said, are an effective way to meet her clients' needs.

When novices discussed anticipating the needs of their audience, they described it as figuring out what their teachers want. When describing a class presentation she did on how to make chocolate, Susan spoke about what her teacher wanted: “I know [the teacher] wanted a lot of pictures.” Other students spoke about asking the teacher for clarification (Darryl), reading the assignment directions (Jake), and reading the rubric (Fiona) to figure out what their teachers wanted. In all cases, students viewed their audience solely or mainly as their teacher—even in Susan’s case where the audience was not only the teacher, but the whole class who listened to her presentation.

Anticipating the objections of the audience.

Beyond anticipating the needs of their audience, three of the expert participants also remarked on how they anticipate the objections of their audience. Catherine described her blog post (excerpt in Figure 4-8) aimed at convincing teachers to use video games like *Assassin’s Creed* and *Medal of Honor* as a part of their curriculum: “So I’m slowly building an argument for the audience discussing the merits and the challenges of using these games in the classroom. If it’s acceptable to use movies like *Saving Private Ryan* to teach World War II, then how come I can’t use a piece of digital media like a video game? They’re both fictional accounts.” Likewise, Laura described how she went to great lengths to write a legal opinion that was “bullet proof” because she “anticipated this being challenged and reviewed at the highest levels.”

Did you know that [ninety-seven percent of North American teens play some type of digital game on a regular basis](#)? How many of these entertainment games, do you think, are played in a classroom setting? Probably not many, because a conflict over the value of gaming in the classroom exists in our society. [James Paul Gee](#) addresses this as a value we as a society, educators and parents place on learning, and our need to demonstrate actual knowledge based on *content*.

At first glance, video games do not contain, for the most part, the kind of knowledge content offered in our math textbooks and history lectures. Gee suggests that the contents of games, when “they are played actively and critically, situate meaning in a multimodal space through embodied experiences to solve problems and reflect on the intricacies of the design imagined worlds and the design of both real and imagined social relationships and identities in the modern world” (pp. 40-41). What Gee forgets in his assessment of a game’s value depending on content is the social interaction the content *instigates*. Teaching young people to engage with games, then to examine the content critically, and assign value to the learning that they engage in is a step in the right direction.

Figure 4-8. Excerpt from Catherine’s blog post in which she anticipates the objections of her audience

Analysing the historical context / intertextuality.

Expert participants Tina and Laura spoke about the need to understand the historical context behind the issues they were writing about. During her interview, Trina described an interesting scenario where she was tasked with writing a specification to fix a piece of computer code that failed to operate as it should. After asking herself who wrote the code and for what purpose, she eventually discovered the answer: “I found a section of code that had been modified, and it turned out it had been modified in response to a union grievance. Someone felt that an automated button had replaced their job.” Similarly, Laura drew from a historical context when writing a legal opinion: “In this case, there were previous legal opinions that came up with different conclusions so I was pulling together a whole bunch of history in this file in order to

give the client an interpretation that was compelling, that they could live with, and that they could use to make an informed decision.”

Analysing the rhetorical situation multimodally.

While the ORCA contained multimodal aspects to read, there was no capacity within the assessment to compose multimodally. However, participants did discuss multimodal choices with regards to their writing artefacts. Darryl, a ninth grader, spoke about his use of pictures (Figure 4-9) on his class project about the endocrine system to appeal to the audience.

Pictures are really powerful for the human brain, especially colour. It will draw you in more than text. When people see text and pictures, most of the time they will look at the picture first and then read the text. This picture I chose has text in it so you can go look and little arrows lead you in and you get an idea, almost a teaser because you don't know what most of this stuff means, and so you'll want to look in for an explanation.

Another participant, a seventh grader named Helen, spoke about how she used pictures in her class research project on polar bears. She selected a picture of a polar bear on a piece of ice to convey how global warming is affecting polar bears: “One of my ideas was that the ice was melting and the polar bears live and fish on the ice. This shows the ice melting and the polar bear.”

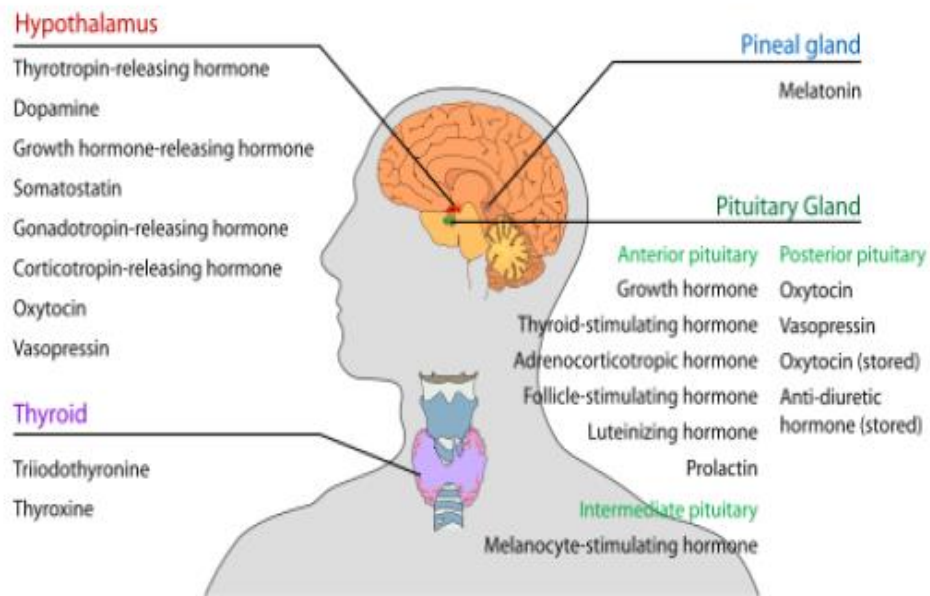


Figure 4-9. Image used by Darryl on the title page of his school project.

While novices only spoke about using images in terms of multimodal design choices, experts went beyond that. Monica discussed the design choices she makes when she blogs about educational issues, something she does regularly: “I think people are better able to deal with short, multimodal texts in an online environment than long, protracted missives like in a scholarly piece of work.” Likewise, Margaret also spoke about breaking text into shorter chunks and using graphics when she writes in online environments. Half of the expert participants (Colin, Margaret, Monica, Catherine, and Lisa) also commented on their use of hyperlinks to draw their reader to additional information online.

In addition to the multimodal design choices such as using images, writing in shorter chunks, and using hyperlinks, two experts (Catherine and Margaret) also discussed the importance managing your [digital footprint](#) to control your reputation in online environments, which they say has an effect on how your online audience perceives you. Catherine described how she once challenged a student to find out all he could about her, and all the student found

was what she *wanted* him to find: information about where she went to school, what she had published, awards she had won, and where she works.

Analyzing the rhetorical situation in order to read online texts.

Expert participants referred not only to analyzing the rhetorical situation in order to write, but also to analyzing the rhetorical situation behind the texts they were reading. Margaret phrased this eloquently:

And the act of reading I see as deconstruction, taking something apart and figuring out how each part is and how it communicates. And writing is an act of construction, and the way to construct something meaningfully is to understand how people read it and deconstruct it. It's this constant back and forth.

While completing the ORCA, participants—experts in particular—discussed analyzing the rhetorical situation across all of the tasks on the ORCA, which included locating, evaluating, synthesizing, and (as has been discussed) communicating online information.

One of the initial tasks on the ORCA challenged participants to locate online information from four online articles, two of which supported and two of which cautioned against the safety of cosmetic contact lenses. Participants used the ORCA's internal search engine, Google, to find these articles. In order to execute a successful search, participants had to be able read the rhetorical situation of the search engine. A search engine has a rhetorical situation in the sense that it has authors (e.g., computer programmers author the algorithm used by a search engine and graphic designers author the layout of a search engine) and those authors wrote it for a purpose: to help people find information—for a particular means (e.g., commercial, educational, informational) and by a particular means (e.g., using keywords, using Boolean operators). Robert discussed his knowledge of reading the rhetorical situation of the search engine when he

discussed his approach to finding one of the online articles he was tasked to find in the ORCA: “I was given the title to search for, so my preferred way would be to copy and paste that into the search bar or address bar. It's faster and more accurate. When I'm told to find a specific page, it should achieve that.” Laura also discussed how during the ORCA and at work she used Boolean operators. During the ORCA, she placed quotation marks around the titles of the articles she searched for. At work, she uses keywords with Boolean operators to search court decisions or other tribunal decisions in online databases.

Once participants located their online information on the ORCA, they were also tasked with evaluating its credibility—and this was one of the keys areas in which participants analyzed the rhetorical situations of the texts they were reading. In this CRR excerpt, Trina discusses how she analyzed the rhetorical situation by understanding the author's purpose for writing the text:

I went back to the top [of the article] and started looking at why she [the author] was writing the story. She's a health reporter, so she's got the health angle, right? She's trying to write to, I can assume, parents mostly, but not strictly, about how to keep your kids safe. She's taking the safety angle. She's not trying to sell contact lenses, but rather she has a broad interest in keeping us safe.

While synthesizing information on the ORCA, experts once again displayed how they were conscious of the rhetorical situation, including the aspects of discourse community, purpose, and ecological context (Figure 4-1). For example, participants spoke about synthesizing their information in a way that was relevant to seventh graders (discourse community knowledge), synthesizing information in a way to prove whether or not cosmetic

contact lenses are harmful to eye health (knowledge of purpose), and of writing their synthesis bearing in mind the affordances and limitations of the classroom wiki (ecological context).

Posing the problem.

When it came to problem posing, the trend was for experts to pose the right problem and to complicate the problem. Meanwhile, novices variously posed right and wrong problems, often simplifying the problem. Also, problem posing was evident across cognitive tasks, whether research oriented (locating and critically evaluating online information) or communication oriented (synthesizing and communicating online information).

Posing the right versus the wrong problem.

In this CRR excerpt, expert participant Trina is problem posing about how she will locate information to support her position regarding the safety of cosmetic contact lenses: “Okay, so there are lots of voices happening here. There's a narrative happening in the background and I really try to park the narrative. I'm asking myself, who's talking? I gravitated towards that association there because it's not just one person, it's a collective.” Similarly, expert participant Colin correctly problem poses in relation to finding evidence to support his position on the ORCA:

If I'm told to find something specific to whether the contact lenses do or do not harm your eyes, then I can scan each paragraph or group of paragraphs to determine if it's going to be relevant. I'm focussed on a very specific task. I'm not asked to read and understand the whole article. In the case of the Miss Dilday's article, it's covering stuff like glow sticks and other treats and things. I knew that I could skip over that. I'm not asked to analyze the whole article or to understand everything she is saying about Halloween

safety. I narrowed it down to a paragraph where there's a quote from an optometrist and I basically just reiterated that sentence in my summary. (Colin)

Many participants also spoke about the problems they pose when they are evaluating the credibility of online information. Here, Trina identifies the questions she asks herself when evaluating online information: “Are they experts? Is this credible? Can I trust this?” Novices tended to be less critical while problem posing when it came to evaluating the credibility of online information, such as is demonstrated in this interview excerpt:

We were asked, ‘Is Ms. Dilday an expert?’ Yes, since there were a lot of quotes from experts in eye care. She was making a lot of good points. Since she was a specialist. She knew a lot about that topic. (Anna)

Compare this to how an expert, Margaret, problem posed and arrived at this more nuanced position regarding the credibility of that same author:

I clicked on the author's name to get details about her. I read the description about her. At first glance, it sounded like she was a credible source, but I realized she only worked as an assistant at an optometric clinic, which is not an ophthalmologist. An ophthalmologist is a doctor. An optometrist gives glasses. Whereas the author, she's primarily a journalist. To me she's not totally credible. If you're talking about health and harm to health, she does not have the medical training to say whether or not those are harmful.

The above excerpts also demonstrate how novices tended to view the credibility of authors and publications as a binary (credible or not; “she *was* a specialist”), while experts viewed it on a spectrum (“she’s not *totally* credible”). Trina explains the way she weighs the credibility of evidence: “So the health reporter studied at Berkeley—that's a reasonable

institution, right? But depending on if it were some made up college name, I wouldn't trust it. I sort it [the evidence] that way, by the magnitude of how much I can trust it."

In addition to posing the right problems while locating and evaluating online information, participants also posed the right problems while synthesizing and communicating this information. Here, Cody, one of the novice participants, shows the problem posing he engaged in while creating his wiki entry: "So, I'm asking myself, are cosmetic contact lenses dangerous or not?" An expert participant, Robert, explained how he synthesizes research in his job by doing a gap assessment: "Then it's also looking at what we [the nuclear agency] already have and doing a gap assessment of what the differences are. Also [we bring together] other standards and use external standards from the U.S."

When it came to synthesizing online information, once again, experts tended to take a more nuanced perspective than novices. Here, novice participant Jake discusses how he synthesized information from one perspective: "The most important idea was that cosmetic contact lenses can hurt your eyes, so I looked for things that said how it could hurt your eyes and why." By contrast, expert participant Monica refuses to take a single perspective: "I wasn't actually willing to take a one side or the other given the information that I had. In each of the articles that I read, I found information that could help me develop a more complex understanding of the issue, not a yes it's harmful or no it's not."

Finally, all expert participants also posed the right problem when they were communicating online research. For example, Catherine spoke about the problems she poses for herself when writing: "I start by asking, 'Why is this important? Why should you care?' And I should already have an idea, if I'm going to be writing about this topic, about why they should care."

By contrast, novices participants seemed at times confused about or distracted from the central problem: “I didn’t really include my opinion [about the safety of the lenses] but I added the stuff that I read and found interesting [about celebrities wearing the lenses],” said Susan of her wiki entry, although eventually she notes, “I changed some of [the wiki] here and wrote why not to wear cosmetic contact lenses.” Below is Susan’s wiki entry in which it is evident that Susan is distracted from the central research problem, first writing information that is tangential, though eventually she writes an implied stance that wearing these lenses is harmful because “Your eye needs oxsigen [*sic*].”

A cosmetic contact lens is designed to change the appearance of the eye. These lenses sometimes also correct for vision problems [e.x.cross eye, reading, far away e.c.t...] However, most of the time they are just used to enhance appearance It can cause blindness and other problemes. They are also called decorative lenses.

who uses the cosmetic contact lenses ?

Some famous people use them for videos or it may look like that but it could only be edited by a cumputer. That is what happend to Laddy gaga fans exsepet fans thought they were real and started to follow her fake apperence witch is not good or ok for you and your eyes

why not to were cosmetic contact lenses?

Your eye needs oxsigen so wearing some thin over top is not a good idea

Similarly, Cecelia shows confusion about problem posing during the process of communicating her research findings:

JC: How did you decide what to include in your wiki entry?

Cecilia: I don't know... I kinda just put everything that came to mind.

However, while Cecilia was unable to articulate the central research problem during her interview, she did in fact edit the ORCA wiki to take a stance regarding the safety of the decorative lenses:

cosmetic lens can be a danger

when people wear cosmetic lens they are putting their eyes at risk. If the do the smart thing it is not a danger.

Eventually, all of the novices posed the correct problem when it came to the central research problem posed on the ORCA (i.e., arguing whether or not cosmetic contact lenses are harmful to eye health), though their responses were achieved with varying degrees of success.

Complicating versus simplifying the problem.

While a number of the novice participants commented that locating the websites was one of the easiest tasks on the ORCA, Judy explained that in her job as a registered dietician, this can in fact be quite challenging. Here, Judy discusses the problem posing she does when she searches for nutritional information:

Just trying to find the right source—that's probably the hardest part. Not necessarily in the ORCA because it gave you the specific website or article to find, but if you're giving a talk and you have to find out about lupus and headaches for example, where do I even start? That's probably the hardest thing to do is to say, okay, first I'm going to look in medical literature database and then I'm going to look at a nutrition database put together by other dieticians in the field that's evidence based. Then I might look at [Up to Date](#), which is a database for the physicians. Then I might look at blogs and websites of other health professionals and then I'll Google just to see what the average person is saying just to know what my patients are reading about too.

This shows how Judy is complicating the problem by not only finding information, but also finding information with consideration to her discourse community. Judy told me that in her job, she often gives presentations on nutrition to medical residents and physicians on the one hand, but also needs to be able to communicate the same ideas to her patients in lay terminology. Here, Judy is also showing a metacognitive awareness that her Internet searches are generated in a filter bubble (Pariser, 2011), with search engines such as Google using

algorithms (based on location and search history) to predict what she is searching for. She notes that she wants to read not only what dieticians are saying about this issue, but what a myriad of others are reading and saying too. Juxtapose that approach with novice participant Cecelia's approach to locating information for her school project: "So I just typed zebras in the search and opened a whole bunch of links."

Similarly, novices tended to simplify the problem when evaluating online information. In this excerpt, Natasha demonstrates an implicit trust of one of the online sources she evaluated during the ORCA: "When she [the reporter] said stuff, she had something to prove what she said. There were quotes from organizations and she said what they said, and she said what doctors said." By contrast, expert Laura is skeptical of that same online article in the ORCA: "I would look at a newspaper to gather information, but I would never quote from a newspaper as a reliable source. At work, I would go to the initial sources, for example in this article, what is the FDA [Food and Drug Administration] doing or saying? And I wouldn't make the assumption that it's illegal—that's someone's interpretation of the law. I would want to see the actual law."

Once again, while synthesizing and communicating online information, novices mainly engaged in simplifying the research problem while experts tended toward complicating the problem. In this example, novice Fiona simplifies the complexity of the information when she is tasked with taking a position: "I wrote that they [decorative lenses] do harm the eyes because all the websites said something about the bad things that could happen [if you wear them]." By contrast, expert Judy's approach is much more complex: "I'll weigh the evidence and use the most reputable evidence first. In this example, I'll use quotes from the optometrist or the FDA. But it's not always trying to balance the pros and cons necessarily because the

pros might carry 90% of the weight and the cons 10%. You're not necessarily giving a balanced view, but rather you're giving an accurate view.”

Planning.

Participants described planning in terms of translating and generating ideas, setting goals, triaging, and organizing.

Translating and generating ideas.

Both novices and experts used a variety of means to generate ideas. Novices spoke about generating ideas by reading source material, free writing, using graphic organizers, using mediating documents, talking to others, writing about interests, and writing about experiences. Experts addressed all of these processes, in addition to annotating source material, thinking over a period of time, having a daily writing routine, referring to texts that one previously wrote or read (texts often stored in reference managers), using technology, and following social media on subjects of interest.

In terms of reading to generate ideas, Trina how before “diving in,” she would think about her writing task and then she would go and read everything she could. When asked about how she came up with ideas, Cecilia stated, “I don't know. I just kinda put everything that came to mind.” For Cecilia, the very act of typing elicited ideas. Trina spoke about how she will often talk to others, sometimes someone in the field or completely outside the field, depending on her purpose. “I like to talk to people. Even in this other article, it triggered a fascinating conversation with my friend Owen. We went and had schwarmas and talked for like an hour because I needed to get more from the conversation. It was the tip of the iceberg for me,” explained Trina. Meanwhile, Anna described how her interest in war stories prompted her school research.

Thirteen-year-old Cody spoke about his choice to do his school spoken word project on Asperger's because he lives with this form of autism; in other words, Cody wrote about what he knows. "I wrote about my feelings, how I don't like sensations and stuff, how I get stuck on stuff, and how I get frustrated," explained Cody. Similarly, Jill wrote about what she knows as a 13 year old: "I had been complaining that there are too many rules at recess. No snow forts, no snowballs, no playing on the ice, no going on the snow. There's nothing to do in the winter and that's why I wrote this."

Other participants spoke about going for a walk or sleeping on it to generate ideas: "I said [to my daughter] you go back to what you're being asked to do and you read that before you go to sleep. Just put it in your head before you go to sleep and you'll wake up with the right question," said Trina. Catherine described how she wrote her blog in her head first, adding that "it was the process of thinking about things on a daily basis." Sometimes, said Catherine, an idea would come to her in the middle of the night or while she was out doing something else. So that she would remember these ideas, she would often send an email to herself.

Experts spoke about using the affordances of technology to generate ideas, particularly with regards to long-term writing goals. This involved practices such as following social media on subjects of interest. For example, several expert participants spoke about following blog posts, RSS feeds, Twitter feed of experts in their field or Twitter feed on subjects of interest. Experts used various apps and reference managers to keep track of the large amounts of data they gathered.

Finally, participants spoke about using mediated planning documents in order to record their research findings and organize their thinking as they completed their writing tasks. Participants spoke about doing this on the ORCA and in their school and work writing task. For example,

Jake explained his use of mediated planning documents on the ORCA: “Looking at my notes, I discovered that cosmetic contact lenses can hurt your eyes but only if you let them [...] Then I took all of the research I found together and it all got into my head. I looked at my notes and I wrote what I thought.” Similarly, Colin described how when responding to client emails, he’ll often search archived emails on a similar topic as use those as a starting point for a new email. Otherwise, participants described writing from the original documents: “It’s only recently that I got two computer screens set, which makes a big difference in how you work, being able to have reference documents on one screen and your writing on another screen. Otherwise you have your reference documents, your cases, and your law in paper on your desk, and you’re writing on your computer screen” (Laura).

Goal setting.

Several novice and expert participants described the notion of breaking up a larger task into several smaller tasks. “If it were something longer, after I got a project—either that night or the next day or as soon as I can—I divvy up what I have to research,” explained novice participant Darryl. “I figure out whatever the teacher wants me to put in this report, I write that all down again, and break it down into what I need to research and write.” Of course, at other times goals were not achieved as planned, according to one novice. “To get it done on time, we did one thing one day and another thing another day and repeated that. But, then we weren’t quite finished and we had to do it at the last minute. We were panicking and had to finish it all in one day” (Natasha). While novices set short- and medium-term goals (for example, Cecilia described getting her information one week, writing over the next couple of weeks, and revising in the final week), experts were often engaged in multiple writing projects simultaneously at various states of completion. These more protracted and often collaborative writing projects

required short, medium, and long-term goals, which were frequently tracked using project management applications.

Triaging.

Some participants engaged in triaging as a part of the planning process, which I define as determining how much time to allocate to ORW, depending on the availability of time, and the importance of the research question to one's self or others. Here, Margaret engages in triaging: "The other thing is, because I knew this was a child's wiki, I was less concerned about the credibility of the authors in those online research articles than I might have been if I were purposefully doing academic research. And even though this is academic, this is at a different level." Margaret also went on to explain how she allocates more time to ORW depending on how high the stakes are. Thus, for example, when writing a peer-reviewed article, she has a professional obligation to make sure she gets the information correct because it could impact educational policy decisions or inform future research. Darryl, the oldest novice participant and somewhat of an anomaly among his peers (in the sense that he engaged in more critical thinking and metacognitive practices), also described using triaging: "If it's not a school report, if it's something I do for fun, like air guns, I just look up where's the cheapest place to buy air guns. Things that are more low key, I just look it up online [not paying particular attention to the source]. But if I have to write an essay or do a book report or whatever, then I'll look up sites that are 99% reliable and find the sites that they [he gave [Encyclopaedia Britannica](#) as an example] recommend to branch off."

Organizing.

For novices, the organization of their writing was most often prescribed. Jill described the format her teacher had asked for: "We already had an outline for how we had to do it: we had

to have a beginning, first, second, and third argument, and a counter argument, and conclusion.” Expert participants, conversely, matched the organization of their writing to the genre conventions appropriate to their profession. Sometimes this too was prescribed (by professional/industry standards). Robert described that in his job writing or revising standards for the nuclear industry, he usually uses a template with information such as the purpose of the document, division of responsibilities, what those responsibilities involve, what training is required to perform those responsibilities, technical details, and finally, a section for approvals and revisions. Similarly, Judy explained how all dieticians follow a standard protocol for patient records: “We follow an international dietetic terminology standard using the SOAP format: Subjective, Objective, Assessment, Plan.”

When not given a prescribed format, participants described organizing their writing using outlines, or sometimes as they wrote, the ideas seemed to organize themselves. Here Laura describes the process she used to organize her writing when drafting legal opinions: “I start pencilling out a framework and making notes as I see different tangents popping up in my head.” As a professor, Monica often uses a less structured, more iterative process: “My outlining process can be very explicit, but often it isn't. Particularly, I kind of like to get into the exploration of the ideas as I'm writing. I actually find that the process of writing itself is how I organize my ideas.” Monica was the only participant to explain how her writing process differs depending on what her purpose and genre are:

I didn't make an organizer for this one because I had been thinking about it for a long time and I was in a sense chronicling the narrative of my experience in the classroom. When it's narrative, I don't need an organizer. If I am presenting a proposal for a symposium or some kind of workshop, I definitely present it in a much more organized

fashion and it's less experiential and more research-based and purpose-based: What's the research agenda, where the research lies, what we would do, what we would use. So I use organizers and templates, but it depends entirely on the purpose and genre.

Executing the plan.

Participants described the execution of their plans—their online research writing—in terms of how they summarized, synthesized, and at times synthesized generatively; using rhetorical strategies; using linear versus recursive writing processes; writing collaboratively; and using digital remixing.

Summarizing, synthesizing, and generative synthesizing.

Summarizing involves briefly stating the main points from one document while synthesizing involves generating the main ideas from multiple, sometimes conflicting, documents. Many participants, novices in particular, emphasized the role of paraphrasing when summarizing online information. Susan spoke about how when she found online information, she would “write it down from my head” when she used that information in her project. Judy, a registered dietician, described summarizing when she described her process writing the wiki for the ORCA: “In my mind, I think, okay, what was the overall message from all of these articles?”

When synthesizing, novices tended to ignore conflicting information: “I chose information that fit with the other pieces of information. So if there was a piece of information but it didn't fit with the other information, then I just wouldn't use it” (Cecelia). Meanwhile, Judy described how she considered opposing viewpoints on the ORCA: “I recorded points for and points against and then synthesize that into one overall message. It's not necessarily about being balanced, by looking at the weight of both sides. In this case, there was a lot more risk than benefit so that should come out.”

Only experts spoke about what could be described as generative synthesizing. Generative synthesizing takes those processes further by going beyond the texts to generate new knowledge (Deschryver, 2014). Judy described how she does this at work: “I think that's the hardest thing to do because, to be honest, I do not remember doing that in school. It's probably only after working for 15 years that I've gotten better at reading a paper, critically analyzing it, and then reading a bunch of information on controversial things—there is so much controversy in nutritional reporting—and trying to come up with one statement to say to the residents that *this* is the best diet. Or research shows that *this* pattern of eating is superior. And that's a really hard thing to do. It's trying to balance everything you read plus your own professional opinion. Doing this, everyone can come to different conclusions because you can read it with your own personal bias and see it in a different way.” Although she works with numbers, Robin had a similar experience in her position as a financial management analyst: “In trying to reconcile, I figure out why there's a discrepancy between our financial records and what's been made available publicly, then I chase the data until I figure out what is it that I'm missing or whether there was an error that was made.” The major difference for Robin is that she is making her argument with numbers rather than alphabetic text, but the process is still the same: taking texts that are at times discrepant and making sense of them by trying to account for these discrepancies.

Using rhetorical strategies.

When ORW either in traditional or multimodal contexts, participants spoke about using strategies to reach their rhetorical aims. Both novices and experts tended to emphasize diction when discussing monomodal rhetorical strategies. Novice participant Darryl, for example,

compared writing “cosmetic contact lenses can pose serious problems with your eyes” to “they might hurt your eyes” and how this choice of words might affect readers: “that seems it's actually a danger to you, instead of just something that can be waved off.” Darryl also added that the use of examples, such as how these lenses can lead to “bacterial infection, ulcers, and blindness” makes the information more believable. Experts also spoke about diction, though with a clear eye to a particular community or audience. Here Colin explains about the word choices he makes when interacting with colleagues and the public:

I try to avoid the use of jargon unless it's already established. If it's going to make life easier for a detailed email, than I may introduce a particular word that may appear as jargon to the lay person, although if you're familiar with the regulations wouldn't seem like jargon anymore. With colleagues, there may be a lot of jargon and acronyms that we use. That's a typical thing in government. With the public, we'll avoid too many acronyms or jargon, or at the very least, we'll introduce them.

In terms of using multimodal rhetorical strategies, novices exclusively focussed on using pictures to achieve their rhetorical aim. Meanwhile, experts had a far greater arsenal of multimodal rhetorical strategies that included using multimodal means of representation beyond still images, using hyperlinks, elements of web design, and aesthetic choices in online writing. In one example, Catherine spoke about placing a hyperlink to a YouTube video of an online video game rape in the introduction of her thesis in order to get her readers' attention regarding gendered violence in gaming: “My goal was to shock [readers]. I want them to think about what they have just read and think what the heck is she doing? And then I slowly show them, this why you're here. This is what you need to learn.” Many of the expert participants also spoke about using hyperlinks in ORW. Catherine explained how this is an aesthetic choice, “so as not

to overwhelm a digital reader. Instead, you can provide hyperlinks to provide additional information and easily connect to those.” Other aesthetic choices in ORW included breaking texts up into smaller chunks and breaking texts up with graphics (Margaret and Monica). Finally, Trina referred to the non-linear nature of hypertexts: “I have a friend who says [referring to authoring Web pages] every page is page one. You never know where someone is going to go. You never know what they're looking for. And every sentence I write is written so that it might be the only sentence a person reads.”

Using a recursive process.

Novices exhibited linear writing processes while experts engaged in more recursive processes. Ninth-grader Darryl described it this way:

I just find that when I'm writing a report or researching stuff online, I jot all my notes down and maybe only put them into two different groups before I summarize my notes. I get all my stuff down before I start working, which I find helps me focus because I don't have to go from writing, and then back into research, and back into writing again.

Conversely, experts described a more prolonged, recursive process:

My process is much more iterative and recursive and it's much more driven by the questions of the particular moment that I need to answer. I'm very aware of when I have a question and then I go find it quickly, go back, and go again. The process typically takes much longer because often times I'm writing things that require deep and careful consideration of facts and there's also the professional responsibility piece. When I think about the communities for whom I write, I have to remember that anything I write could have an impact on kids in classrooms. That's a piece that I have to consider. It's professional due diligence. (Monica)

Writing collaboratively.

Both novices and experts described writing collaboratively, particularly with the aide of online technologies. For novices, this typically meant Google Docs and Slides to do group projects and presentations, while experts also used the Google platform in addition to [SharePoint](#), [Slack workspaces](#), wikis, Dropbox, and reference managers.

Experts particularly described writing collaboratively in instances where no one person had sufficient expertise to complete the whole project, which as Robert described, is often the case in the nuclear industry. Judy explained that where she works, a team of health care providers (i.e., general practitioners, nurse practitioners, physiotherapists, psychologists, registered dieticians) contribute to patient records to provide a holistic picture of the patient, particularly in complex health situations (e.g., what diet will help a patient with lupus who is experiencing headaches?). Laura explained how writing collaboratively involves a fair amount of project management so that she and her colleagues are “both not adding to a document at the same time—that would be messy” and so as to meet deadlines that are months away.

Experts also described the challenge of matching voice and style when writing collaboratively. Referring to editing the wiki entry on the ORCA, Trina said, “I dropped my summary in and then I looked for the voice and then fitted my own writing to that voice. I checked to see if they [others who authored the wiki on the ORCA] quoted any experts. They didn't, so I paired that out. Then I made it two sentences to fit with the length of the other sections. I made a heading and bolded it. I wrote it in a way that would blend in.” Trina also described how organizations she has worked for doing technical writing use style guides (e.g., *The IBM Style Guide*) to keep writing conventions standardized.

Digital remixing.

When it came to ORW, novices repurposed online material by copying pictures from a website and pasting them into their projects, or by paraphrasing research they had read and writing it into their projects. Experts variously spoke about repurposing their own online material and that of others, and also transforming online information by applying it to new contexts. For example, Robert noted that when he writes nuclear industry standards, he would often look at external and U.S. standards and do a gap assessment in order to determine what the differences are between the nuclear facility where he works and others. Colin discussed how he and his colleagues have shared, organized folders in their inboxes with emails about recurring topics. Then, when a client emails him an inquiry, he can reference these emails and not “reinvent the wheel.” Also, he says, the government wants “to have a consistent approach to our responses to the same type of issues.” Similarly, in the legal profession, Laura explained how “you don't want to be writing advice again on something that has already been opined on. If someone else has already written a legal opinion on a certain topic, then you don't want to reinvent the wheel. We use the online tools to help us do the research and see what else is out there on the topic.” However, Laura cautioned, each legal situation is unique meaning that she needs to examine situations on a case-by-case basis.

Evaluating the plan and its execution.

While novices tended to focus on local revision, experts focussed on global revision and reflection.

Recursively monitoring and evaluating.

Fiona's response about revision was typical of the novice participants: “Since I did it on the computer, I got the computer to edit my spelling mistakes and then I got my sister and some

friends at school to edit it. I also read it out loud so that I could hear how it sounded. I want to hear if the sentences made sense.” Time and time again, novices noted that they were revising for spelling, grammar, and punctuation, and occasionally for meaning. The writers’ self-knowledge also affected what and how they revised. Take this interview excerpt with Cody, a novice participant:

JC: Did you have to make a lot of revisions along the way?

Cody: Not a lot since I only wrote down something that I was sure about.

Conversely, expert participant Catherine knows that at times she needs to write about challenging subject matter, but she had a strategy to achieve that: “I think it's very valid to know what your weaknesses are and make sure you surround yourself with people who have opposite strengths. For example, if I'm not strong theoretically, I'm better with methodology and analysis, I always turn to a particular colleagues with those strengths.”

Finally, novices spoke about a very simple, linear review process wherein a teacher or peer revised their writing prior to submitting it; for experts, on the other hand, these processes were much more complex. Here Robert describes the review process for writing nuclear industry standards:

We have a pretty elaborate review and comment process. Even before you start writing, you have to get input from a lot of people. There are going to be subject matter experts. And if you know it's going to apply to a certain segment of the workers, than you talk to people in those facilities to get input. As one author, you can't possibly be expected to know all of those things. And that's just the preparation. In the review and comment process, then you basically determine a list of people that you think need to be on the review. That could be anywhere from two to a dozen people, depending on how complex

it is. It's usually because these people have different areas of expertise that need to be covered or different responsibilities that need to be covered. They provide all these comments. It's all done online now. They get a notice that here's the document and provide your comments.

In a similar manner, Laura described getting feedback from her boss at the Department of Justice three levels up from her before she sent her client the legal opinion she had written to ensure that “Justice speaks with one voice,” meaning that the legal opinion Laura was writing needed to be aligned with the practices and policies across the Department of Justice.

Transfer.

Novice participants neither spoke about how they would apply what they learned from current writing tasks (i.e., editing the ORCA wiki) to a new situation (i.e., forward-reaching transfer) nor did they discuss taking what they have learned from previous writing experiences and applying that learning to the current writing task (i.e., backward-reaching transfer; Brent, 2011). Experts, however, did discuss what could be discussed as transfer. For example, Margaret reflected on how she would like to improve future blog posts after considering the one she shared with me during her interview: “With blogs being a visual medium, having that excessively dense text makes it challenging for readers to stay with it and follow. So I think I would separate it with pictures. I would also make paragraphs much more condensed and if people wanted further information, they could click on a link and go to it, rather than having it all on this one page.”

Discussion

This study used a mixed methods approach to explore the cognitive and metacognitive processes engaged with during ORW. Secondly, it sought to explore the differences between

novices and experts, specifically student writers (Grades 7 – 9) and knowledge workers. Finally, it sought to examine whether or not the empirical evidence from this study supported the validity of the bioecological model of writing. The results suggest that there are complex and sophisticated cognitive and metacognitive processes underlying ORW and, tentatively, that some of these processes are unique to online contexts. Secondly, both quantitative and qualitative results suggest distinctive differences between novice and expert groups. Finally, evidence from this study appears to support the validity of the bioecological model of writing. This study will conclude with a discussion surrounding the implications these findings have for theory, research, and practice.

Cognitive Processes Enacted during ORW

Regarding the quantitative cognitive domain results, both experts and novices found locate to be the most challenging domain while synthesize was the easiest domain. In my previous study of representative sample of $n = 1129$ seventh grade students, locate ranked as the easiest of the domains (Corrigan, 2014b), while in this study, that result was inverted. The most likely explanation for this is that, as there were technical difficulties during the ORCA with the some of the locate tasks, this part of the data was not reliable. Specifically, one quarter of my participants (mostly experts) received an error code during the assessment when they attempted a search within the ORCA's internal search engine. This study did, however, confirm results from the previous study that ranked synthesize as the easiest domain.

Qualitative results from this study showed numerous cognitive processes enacted during ORW (see Table 4-4 and Appendix D – G) many of which seem to be unique in or more significant to online contexts. I will begin with processes surrounding locating online

information, followed by those regarding evaluating, synthesizing, and communicating online information.

Locating online information. In this study, I distinguish between primary location (locating sources) and secondary location (locating information within the source). With regards to primary location, results from this study corroborate what Pirolli (2007) described as information foraging. In online environments where information is seemingly infinite, it is necessary for online research writers to triage constantly—that is, to assess the importance of the research problem to themselves and their stakeholders, while taking into consideration the availability of time and energy. Knowing that spending more and more time on locating online information would result in increasingly diminished returns, writers in this study foraged or triaged to determine whether or not more searching was worthwhile. Though this process is evident in offline environments (say, in a library), it would appear to take on a greater and more important role in online environments. While foraging is a process specific to locating information, what I call triaging was evident across all ORW processes—more particularly with experts in the study. The process of foraging is perhaps further pronounced in online environments not only because of the sheer volume of information, but also due to click bait from other commercialized and tabloidized links competing for online readers' attention (Blom & Hansen, 2015).

Efficiently and successfully locating online sources of information also required participants to query search engines, online databases, and search tools embedded within websites, which was followed by navigating the architecture of websites to further locate information. Expert participants used more sophisticated search strategies such as using Boolean operators, using more specific key word searches, and were more deliberate when looking

through search results. That is, they were less likely to automatically click on the top search engine hit, and instead looked for more clues (date of publication, publisher, reading the resume) within the search results before deciding which sites were worth visiting. This finding supports Pan and colleagues' (2007) finding that students place substantial trust in the ability of a search engine to rank results in terms of relevance.

Other ways in which experts distinguished themselves from novices included their less traditional means of locating online information. One key example was that experts sought out information from a greater diversity of online sources including microblogs (e.g., Twitter), social media (e.g., Facebook), listservs, and academic/professional blogs. They then used mobile and web applications to manage this mountain of information (e.g., Mendeley). Further, the results showed that when locating online information, experts would often seek out information with an eye to the audience for whom they were writing. Recall how Judy, the registered dietician, would search for information from different sources based on her intended audience (be they other dieticians, physicians, or patients). That experts would locate information giving consideration to their rhetorical purpose appears to be a finding unique to this study; however, this finding is not surprising considering that many studies have supported the notion that experts most distinguish themselves from novices in writing tasks by means of their rhetorical awareness (Bryson, Bereiter, Scardamalia, & Joram, 1991).

In terms of locating information within the document, or what I have termed secondary location, experts spoke of skimming the text for relevant information while novices tended to read the whole text and became distracted by irrelevant information and click bait. This corresponds to Coiro and Dobler's (2007) hypothesis that proficient online readers require active self-regulation and flexible metacognitive strategies. The need for flexible metacognitive

strategies arises from the fact that hypertext and hypermedia are open-ended and constantly in flux. No two readers read these texts in the same way, depending on where they enter and exit from. Further, the same person may even read the same text differently by taking a different path, or even as the information is updated by the moment in terms of content, form, and structure (Reinking & Colwell, 2015). A limitation (and simultaneously an affordance) of this study was that the ORCA operated within a closed system and thus there was less opportunity for participants to follow a virtually infinite number of links, broken links, and/or become overly distracted by click bait. Thus, there was less opportunity for participants to require active self-regulation while completing the ORCA in comparison to open Internet tasks. This was offset to a degree by participants being interviewed about their writing artefacts and the how their ORW processes differed when composing these in comparison to the ORCA. Another complexity of locating information in online contexts is due to the need to continuously switch between scrolling and skimming (in the case of primary location, for example, when reading search engine results), to knowing when to engage in *slow reading* (e.g., in the case of secondary location, for example, when reading a longer passage on a website; Coiro, 2015).

Evaluating online information. One of the domains in which, according to the quantitative and qualitative data, experts and novices differed the most significantly was evaluation. In general, experts showed scepticism when evaluating online information while novices tended towards naiveté; this pattern was apparent across both primary evaluation (evaluating argument credibility) and secondary evaluation (evaluating source credibility). As such, when experts questioned arguments (including their claim, evidence, and warrants) from a source, they employed sophisticated strategies including backward reference searching (returning to the primary source of information; Levy & Ellis, 2006) and even second-level backward

reference searching (consulting the references of the reference; Levy & Ellis, 2006) to establish credibility or lack thereof. During secondary evaluation, all of the experts clicked on the hyperlink of the author's name to find out more about the author's education, expertise, and purpose for writing, while only two of the novices did. When presented with information about the author, experts noted that an author can fake their credentials and experts said that, when circumstances warranted it, they would conduct further background research about the author and/or publication by cross-referencing the information with other sources of information unconnected to the source in question. This was in stark contrast to novices who tended to view information at face value and, in general, did not corroborate information with external sources. These findings align with those of other studies that have shown that critically evaluating online information is a complex cognitive process that tends to present difficulty for student writers at all levels of education (Forzani, 2016; Kiili et al., 2008; Wiley et al., 2009).

It was also of note that experts tended to view credibility on a spectrum (i.e., to what extent is this information credible?) while novices viewed it as a binary (i.e., is this information credible or not?). This seems to coincide with experts engaging in more nuanced ontological perspectives (i.e., recognizing that, because truth is a social construction, what is true for one group or person may not be so for another).

Synthesizing online information. While experts weighed the credibility of evidence from contradictory sources when synthesizing online information, novices tended to ignore information that contradicted their position. Further, experts constructed a range of diverse connections among their present text and texts previously written by themselves and others (i.e., knowledge of intertextuality) in ways that contributed to new knowledge formation. By contrast,

novices were more likely to paraphrase, parroting information from across texts that coincided with their position.

The expert processes led to greater incidences of what is known as generative synthesis (Deschryver, 2014)—using extant information to generate new ideas. Experts complicated the problem in their effort to “transform knowledge rather than merely report about it” (Leijten et al., 2014, p. 288). This is not a new problem. Bereiter and Scardamalia's (1987) work emphasized the importance of moving students from knowledge telling to knowledge transforming models. They take up Whitehead's (1929) notion of inert knowledge, calling it “propositional knowledge that the student could express but not use” and “the passive reception of disconnected ideas” (Chapter 7, Preface, para. 3)—or what is often referred to as parroting, in the colloquial sense. While it has long been a goal of educators to move their students from passive, disconnected learning to critical, connected learning, perhaps this challenge takes on greater significance in a PostFordist / FastCapitalist society wherein “more and more workplaces are opting for a flattened hierarchy” (New London Group, 1996, p. 66). In the knowledge economy, critical and divergent thinking is revered while widget makers and button pushers need not apply.

Where generative synthesizing was especially evident was in incidences requiring distributed intelligence (Pea, 1993); that is, in some instances, ORW tasks were so complex that no one individual possessed the expertise to accomplish the task. Recall the example of Robert writing safety standards for a nuclear facility necessitating an elaborate review process by numerous subject matter specialists. Or, recall the example of Judy working within an interdisciplinary team of health care professionals to address a patient's complex health needs. One of the affordances of ORW in comparison to its analogue counterpart is the ease by which authors can easily and effectively engage in multimodal collaborations across time and space

(Leijten et al., 2014). In an increasingly globalized world where multinational corporations often employ workers from around the world, collaborative ORW will necessitate writers who not only possess content and rhetorical expertise (Geisler, 1994), but also the ability to negotiate multiple sociocultural and linguistic contexts (New London Group, 1996) and rapidly-changing communication technologies (Leu et al., 2004). Further, this intelligence is not only distributed across the minds of individuals and groups, but rather “across minds, persons, and the symbolic and physical environments, both natural and artificial” (Pea, 1993, p. 47). Digital and online technologies enable us to do create, design, communicate, and share knowledge across time and space in ways we could have never previously imagined.

Communicating online information. Is ORW really that different from its analogue counterpart? If so, how? From the data in this study emerged processes that appear to be quite unique to online environments. One telling example was when Trina, a technical writer and computer coder, described how when constructing online texts, she must consider that every page, in fact, every sentence, could be the first or last that a reader reads. Whereas print-based texts are (generally) read from cover to cover, the nonlinear nature of online texts presents unique challenges to the readers and writers of online information. Additionally, hypertext allows for the author to make intertextuality explicit. Using hypertext, online writers are able to embed hyperlinks in their current text in order to direct the reader to other texts within or beyond a given website. Concomitantly, while hypertext authors provide their hyperlinks, readers draw their own connections, by (a) choosing whether or not to follow a hyperlink and in what sequence and/or by (b) using search tools and navigational menus to link to material of their own determination. As such, reading hypertexts requires a great deal of metacognitive skill (Kamil, 2015) and a “much more active role in determining the quality and coherence of the texts they

read” (Coiro & Dobler, 2007, p. 219). This has important implications for ORW because this hybrid practice necessitates both the reading and writing of hypertext and hypermedia. When locating, evaluating, and synthesizing online information, the online research writer must “plan and monitor choices about where to go in a particular text and in what sequence to move” (Coiro & Dobler, 2007, p. 243).

When communicating online information, participants in this study spoke about their careful consideration of the aesthetics of online writing, taking care not to overwhelm the reader. Thus, the writer might decide to pare down online texts by using hyperlinks to direct the reader to non-essential and/or additional information. Further, a number of the expert participants mentioned how they chose to make their online texts less dense by writing shorter, more condensed paragraphs, and by breaking up sections with graphics and other hypermedia. They spoke about capitalizing on the affordances of online environments by interleaving their texts with multimedia components. This finding provides empirical evidence to support the call for multiliteracies, that is, the need for citizens to learn and engage with designs of meaning beyond alphabetic text (Kalantzis, Cope, & Harvey, 2003; New London Group, 1996). It also highlights the importance of understanding multiple grammars (e.g., the grammar of visual design) and their affordances and limitations (Kress, 2003; Kress & van Leeuwen, 1996).

Another process that took on particular importance in this study was that of remixing (Lankshear & Knobel, 2011), which refers to the recombination or repurposing of pre-existing texts or elements of texts into something new. In this study, there were examples of participants repurposing content (e.g., using an image taken from a website in one’s class project on polar bears) and form (e.g., when writing a blog post for an organization, looking to see the form and style used by other writers within that organization). Similarly, another study found that

professionals remix in order to reuse content, structures/layout, phrasings, and update visuals and layout (Leijten et al., 2014). According to Swarts (2010) remixing is a ubiquitous practice in professional communication. Experts in the current study instinctively sought out models in their writing—particularly when writing in a new context or with unfamiliar subject matter—adapting the form of another text for their own purposes. This was evident when the experts in this study took note of how the ORCA wiki that they were editing had been previously written by their collaborating authors. Experts modelled their own revision of the wiki after the existing text in order to make their voice fit the text; in other words, they remixed taking into account their rhetorical situation.

Experts were also seen remixing via a process of reverse engineering, that is, constructing their own text by deconstructing a pre-existing text to see how it was made. In this study, for example, Robert sought out American nuclear standards as he worked to construct a standard for a Canadian nuclear reactor. He described doing a gap analysis to determine what the American standard had to offer and what the standard at the Canadian facility still needed. These two remixing practices—remixing while considering the rhetorical context and remixing via reverse engineering—corroborate findings from an earlier ORW study (Leijten et al., 2014).

Another process of communicating ORW that I will highlight is that of communicating critical discourse knowledge. Elsewhere, my colleague and I theorized that critical discourse knowledge (Corrigan & Slomp, 2017) is knowing how writing is socially and historically constructed within specific power relations, and knowledge of how writing reinforces discourses of both dominance and marginalization. Though this theme did not occur frequently in my analyses of the transcripts, it was nonetheless an important one. One participant in particular, Catherine, spoke about the need to recognize minoritized voices in her writing. She also spoke

about, in her role creating digital educational resources, being cautious of the prevailing ethos that Western learning styles are superior. More generally, several of the participants spoke about making their writing accessible. Both theorists of New Literacies (Leu et al., 2016) and multiliteracies (New London Group, 1996) espouse the importance of new social practices to communicating in our global village thereby fostering socioculturally and linguistically relevant communication. Indeed, an important means of fostering a socially just and democratic society will be through the recognition of the power of writing as a symbolic act. This is particularly timely here in Canada where we grapple with legislation such as the [Indian Act](#) that made possible policies such as residential schooling, which led to the erasure of indigenous identity and a cultural genocide for our First Nations. Recently, the Canadian government has attempted to redress these historical transgressions by means of the [Truth and Reconciliation Commission](#). The findings of the Commission serve a powerful reminder that the written word, though it be abstract, can and will—without careful consideration and deliberation—lead to actual human rights violations, particularly for minoritized groups.

Finally, I wish to note that, when communicating ORW, more expert participants tended towards recursivity. That is, while these cognitive and metacognitive processes are described linearly in this article (a limitation of communicating via a print-based text), they were enacted recursively. In other words, during the CRR, participants described moving among numerous cycles of these processes. The more actively engaged participants were with metacognitive processes, the more they realized their need to, for example, re-assess the rhetorical situation or pose new problems. It should be noted that the use of CRR as opposed to concurrent think-aloud methods posed a limitation in this study in the sense that participants tended to recall high-level EF processes but were less aware of their low-level EF processes. I suspect that concurrent think

aloud methods would prompt more in the moment monitoring and call greater attention to low-level EF. That being said, this study did show how writers, particularly from the expert group, were evaluating their plan and its execution not just towards the end of the written task, but throughout, which corroborates other studies of online reading and writing (J. L. Coiro & Dobler, 2007; Hacker et al., 2009).

Metacognitive Differences

How were these high-level EF—these metacognitive differences—enacted between novices and experts? In general, experts employed more sophisticated and efficient metacognitive processes, which in turn supported more sophisticated and efficient cognitive processes. Evidence from this study suggests that experts were able to drive their transaction (the process wherein writers construct meaning; Figure 4-1), and not be driven by it. In other words, while experts are dually influenced by their rhetorical situation and context (intrapersonal and ecological) just as novices are, experts are more cognizant of this and use that knowledge to shape their writing. The results pointed to several instances of this.

Quantitative results showed that experts' mean rank for analyzing the rhetorical situation was 15.40 while for novices it was only 5.60. In other words, there was considerable evidence from the CRR and interview data of experts effectively analyzing the rhetorical situation while for novices, there was little. Qualitative results corroborated this. During the interview, Laura spoke about how her decision to communicate legal advice to her clients via email rather than via a formal legal decision. Laura came to this decision upon analyzing the rhetorical situation and seeing the need for her client to have “the bottom line,” coupled with the need for efficiency in the Department of Justice for whom she works. In analyzing the rhetorical situation (blue circle; Figure 1), she has considered her discourse community (she is a professional lawyer

communicating to a lay person); her purpose for ORW (to provide legal advice to her client); and the ecological context in which she is writing (she is writing an email to a client on behalf of the Department of Justice). In more complex and/or precedent-setting contexts, however, Laura chose to write formal legal opinions. Such was the case with one of the [writing artefacts](#) she shared with me involving a complex case in which for many years, incorrect legal advice had been given as a result of faulty legal analysis. As a result, Laura chose to write a lengthy legal decision setting out the facts, providing the legal framework, doing a legal analysis, applying the facts to the framework, and then arriving at a conclusion. From these examples, we can also see how Laura was influenced, though not necessarily controlled by, her intrapersonal and ecological context. Laura's intrapersonal context demonstrates that she was motivated, inquisitive, and resourceful. She went to the effort of digging up older case files and using her legal training to recognize that previous legal analyses had been wrong. She understood her ecological context when she anticipated the challenge she would be facing regarding her legal analysis. To preempt this challenge, she had her boss three levels up review her decision. This transaction—her writing of both formal and informal legal opinions—was supported by a high degree of metacognitive processing. Not only do we see Laura analyzing the rhetorical situation, but the evidence also shows that she posed problems (e.g., she asks herself whether past legal analyses of this case were correct); planned (e.g., she described setting goals such as reviewing old case files, organizing her ideas, writing her legal decision, and having her decision revised by her boss); executed the plan (e.g., she wrote a compelling legal decision that won her accolades from her department); and evaluated the plan and its execution (e.g., she discussed how she monitored herself throughout this process; she discussed what she learned from this process to apply to new situations).

Juxtapose the example of Laura with novice participants' accounts of their ORW. Instead of driving the transaction as Laura did, they were driven by it. In terms of analyzing the rhetorical situation, for example, novices mostly communicated the purpose of their text as being a means to an end, one of a series of assignments that they were required to submit to their teacher. Also, novices generally defined their audience as their teacher, even in situations where fellow students were also their audience (as is the case with a class presentation). Additionally, for the most part, novices failed to acknowledge the ecological context of their texts and the affordances and limitations thereof (e.g., how ORW differs in online and offline contexts). In terms of their intrapersonal contexts, however, novices did cite tapping into prior knowledge, personal experiences, and personal interests as a way to increase their motivation to write. Also, novices portrayed an emerging range of metacognitive strategies to support their ORW processes. For example, they were able to identify (to an extent) their purpose for writing the ORCA (e.g., to communicate whether or not decorative contact lenses are harmful or not). They planned by generating ideas, setting short- and medium-term goals, and organizing their writing. And, they executed their plan, generating some thoughtful revisions to the ORCA wiki, and by constructing their own creative and informed ORW artefacts. These included texts such as a spoken word poem about living with Asperger's syndrome, an essay about why there are too many rules at recess, and projects ranging on topics such as the endocrine system, polar bears, and how to make chocolate.

While novice artefacts were compelling in their own right and appropriate for their developmental level, in general, novice artefacts tended to correspond to a knowledge telling model while expert ones variously corresponded to knowledge transforming and knowledge crafting models (Bereiter et al., 1988; Kellogg, 2008). Based on the results of this study, I

hypothesize that the primary indicator of whether a writer will follow a knowledge telling versus a knowledge transforming or crafting model is a writer's ability to control the transaction—that transaction being driving by recursive metacognitive processes found in the bioecological model of writing (i.e., analyzing the rhetorical situation, posing the problem, planning, executing the plan, and evaluating the plan and its execution).

Limitations

This study continued important research into digital writing yet in its infancy. As an exploratory study, there were a number of limitations. First, this study investigated only one form of digital writing, that of ORW. Not only that, but it focussed on investigating conflicting claims ORW by using the Cosmetic Contact Lenses version of the ORCA. For subsequent research, it would be interesting to use the learn more about versions of the ORCA. Indeed, previous writing research (Leijten et al., 2014) has shown that writers adapt their cognitive and metacognitive processes according to their rhetorical purpose.

Furthermore, I observed participants in a controlled setting, that of a university computer lab. While some preliminary research has been conducted that examined digital writing by people in natural environments such as their workplace (Leijten et al., 2014), we have much to learn. Writing in natural environments would indubitably introduce a number of variables that could lead to the discovery of any number of new cognitive and metacognitive processes.

An additional limitation is posed by my choice of participants. I selected Grades 7-9 students for my novice sample, and knowledge workers for my expert sample. Further research might examine participants from both a broader (e.g., a greater range of occupations) and more specific range of contexts (e.g., studies focussing on particular employment contexts) and from a broader and more specific age range. This would help us to understand the developmental range

of ORW in a more nuanced way. Finally, it is impossible to know from this study whether experts truly had better metacognitive ability or whether they were simply better at articulating it. It would prove interesting to explore this further.

Implications for Theory, Research, and Practice

This study explored the cognitive and metacognitive processes inherent to ORW and how those processes manifest differently between novices and experts. The findings of this study have important implications for writing theory, research, and practice. In terms of theory, a stated aim of this study was to investigate whether empirical evidence supported the previously theorized bioecological model of writing (Corrigan & Slomp, 2017). The results of this study do appear to support this model. Moreover, this model may help to explain why novice writers rely on knowledge telling while experts engage in knowledge transforming and crafting (Bereiter et al., 1988; Kellogg, 2008). Of course, caution is needed when interpreting the results from this study due to its small sample size. However, as I stated in the Methods, this study was designed to elicit analytic generalizations (to generalize from one's findings to theory; Yin, 2009), not external generalizations (to generalize from one's results to a population).

This study also contributes to the development of lower-case new literacies theory (Leu, Kinzer, et al., 2013), namely, the new literacies of online research writing. This study frames ORW as a socially and ecologically situated problem-based inquiry supported by a range of cognitive and metacognitive practices. This research contributes yet more evidence in support of upper-case New Literacies theory (Leu, Kinzer, et al., 2013) suggesting that, indeed, new Internet-based technologies and new social practices redefine what it means to be literate in our global village—one which has unprecedented access to the Internet.

This study also extends writing theory. While Flower (1989) argued against a single, integrated writing theory—“writing is too complex a phenomenon,” she wrote (p. 286)—I respectfully disagree. Rather, it is precisely because writing is such a complex phenomenon that an integrated theory (i.e., the bioecological construct writing) is required. Though there are indisputably limitations to such an integrated theory, I feel that the time has come for writing researchers from all persuasions—from cognitive and metacognitive to sociocultural and post-structural—to unite in contributing to our understanding of this complex, psychosociocultural-meta/cognitive practice known as writing. An integrated theory capitalizes on the depth of singular perspectives, thereby increasing our breadth of understanding, in turn painting a more holistic picture of writing theory. This study provides empirical evidence seeming to validate one such integrated theory, specifically, the bioecological model of writing.

In terms of research, this study contributes to our understanding of the cognitive and metacognitive processes—and their development across a novice-expert continuum—in ORW. To date, this hybrid practice has been under-researched, even though previous research suggests that it is a practice that is vitally and increasingly important in our knowledge economy. Furthermore, while this study investigated ORW particularly, there are ramifications for writing in online contexts more generally. Additionally, this research serves as a model for future researchers seeking to explore how writing is situated within intrapersonal and ecological contexts, is influenced by its rhetorical situation, and is driven by a metacognitively-supported process of transaction.

Regarding practice, although it is outside the scope of this paper to address the myriad of implications this study has for the curriculum design, instruction, and assessment of writing, I will address some key implications. Owing to the fact that this study and preceding ones (D. S.

Davis, Huang, & Yi, 2016; de Milliano, van Gerlderen, & Slegers, 2012; Drijbooms et al., 2015; Ellis, Denton, & Bond, 2014; Kent & Wanzek, 2016b) provided clear evidence in support of the important role of metacognitive processes in ORW (and for writing more generally), it seems clear that metacognition needs to be at the forefront of curriculum design, instruction, and assessment. In fact, in a meta-analysis of 91 studies by Wang, Haertel, and Walberg (1993) identified metacognitive knowledge as one of the most important factors in academic achievement. Integrating metacognition into writing curriculum is becoming increasingly common in jurisdictions around the world (Australian Curriculum, Assessment, and Reporting Authority, n.d.; Common Core State Standards Initiative, 2010; Ontario Ministry of Education, 2007). However, it is seldom assessed, particularly on large-scale assessments of writing (Leu et al., 2016; Slomp et al., 2014). Two likely explanations for why it is rarely assessed on large-scale writing assessments are that it is subjective (thus creating reliability issues) and time-consuming to assess (Hillocks, 2002). Until metacognition takes on a greater role in large-scale assessment, it is likely to have a less significant role in the classroom—a concept known as washback in assessment (i.e., the idea that a test influences what is or is not taught in the classroom; Messick, 1996). Thus innovative assessment tools are needed that recognize the importance of metacognitive practices in writing (Deane, 2015), particularly in online contexts.

While large-scale assessment practices may take time to catch up to classroom practice, it will be important for teachers to have models regarding how to implement evidence-based instructional practices regarding metacognition in online writing contexts. Although more empirical work is needed, evidence from this study suggests promise in looking at writing through the bioecological model of writing. An interesting line of inquiry might be to explore how teachers can teach students to drive their transaction during writing, and not be driven by it.

The evidence from this study suggests that experts—with more sophisticated and elaborate metacognitive strategies—are better able to drive their transaction compared to novices. The bioecological model of writing might prove a useful model for the classroom as teachers challenge students to examine their rhetorical situation and context (intrapersonal and ecological) and reflect on how these can be used to shape—and not control—their transaction in writing.

Finally, this study also revealed a broad range of cognitive and metacognitive processes involved in ORW. As such, it will be important for teachers to develop assessment tasks that are authentic, rich, essential, active, and engaging (Stenmark, 1991) so as to challenge students to develop the diversity and complexity of skills, strategies, dispositions, and social practices that they will need to be successful in ORW in their personal, social, political, and economic lives. Further, because no teacher could ever prepare students for the vast array of communication technologies, discourse communities, genres, and contexts that students may encounter—some of which do not yet exist—it will be necessary for teachers to teach with transfer in mind (Schwartz, Bransford, & Sears, 2005). Yet, a paucity of research exists in this area (Slomp, 2012; Yancey, Robertson, & Taczak, 2014).

In conclusion, as online research writing continues to play an important role in our private, public, and professional lives, it will be important for writing theorists and researchers to continue to explore this complex phenomenon. Furthermore, it will be important to study further this phenomenon in classroom and professional contexts. Although this study was limited by a small sample size as well as various methodological considerations, it represents an important starting point for this area of inquiry. This study began the work of elucidating the cognitive and metacognitive processes underlying the construct of online research writing. It also revealed important differences between novices and experts with regards to these processes. Finally, it

provided empirical evidence to support a model that may be useful in both research and classroom contexts.

Appendices

Appendix A Interview Protocol

Semi-structured Interviews (~20-30 minutes)

I will explain to the participants that the final phase of their participation in this study is to answer a few questions regarding how and why they do online research and writing in their professional and personal lives. I will have asked them in advance to email me samples of their online research writing, or bring them in, if they have any. These samples could be a participant's homework assignment, report for work, links to blogs, emails, Twitter feed, and so fourth. These samples will be used to generate discussion regarding how and why the participants do ORW. The samples can also be used as artefacts I can analyse in terms of the products people create when ORW.

The following are potential interview questions that will help guide the process:

1. [The first question I will ask concerns construct remodelling, a form of evidence used in the model for collecting consequential validity evidence. My goal will be to discover whether participant response processes demonstrate that the writing task is measuring the same construct for different groups—in this case, novice and expert groups.] Questions will vary from participant to participant, based on their responses to ORCA prompts, which I will be monitoring on a secondary screen while the participants are taking the ORCA. Some questions might be, “I noticed that you left this question blank. Can you explain why?” Or, “Here you wrote [...]. Can you explain how you arrived at that answer?” Etc.
2. Can you compare and contrast the tasks in the test you just took to those that you do in your day-to-day life? These can be tasks you do at work/school, or on your own time. Are they similar or different to the ones on the test? Are the tasks on the test realistic? Would you change anything about the test?
3. In your personal or professional life, do you ever share information that you've found online? What types of media or technology do you use, if any? Whom do you share it with? Why? How? For what purpose?
4. Please tell me about the artefacts you brought with you/mailed to me. What are they? Why did you write them? For whom? Can you describe the process you used to research and write them?

5. Do you ever collaborate with people when doing online research writing? Does the process change from the one you use when you do it individually?
6. [For expert group] Can you explain the role that online research writing plays in your job? What type of training or education did you receive, if any, to learn how to do this? What do wish you had learned in your education to prepare you for what you're doing now? What did school do well to prepare you? Where could it be improved?
7. [For novice group] Do you do online research and write about it in school? Which classes? What types of assignments do you do? Did you learn to do online research writing in school or on your own? What lessons do you learn in school, if any, to help you do online research writing? What is missing, if anything, from what you need to learn about online research writing?

Appendix B Rubrics Used for Scoring ORCAs

Table 1.

Analytic Rubric

ORCA Constructs and Score Points		
<i>CATEGORY AND ITEM</i>	<i>ITEM</i>	<i>SCALE</i>
LOCATE: Reading to Locate Online Information		
Can students:		
<ul style="list-style-type: none"> locate the correct email message in an inbox or the correct section of a wiki? 	Locate 1	0-1
<ul style="list-style-type: none"> use appropriate key words in a search engine? 	Locate 2	0-1
<ul style="list-style-type: none"> locate the correct site in a set of search engine results? 	Locate 3	0-1
<ul style="list-style-type: none"> identify correct website addresses in two different search tasks? 	Locate 4	0-1
EVALUATE: Reading to Evaluate Online Information		
Can students:		
<ul style="list-style-type: none"> identify the author of a website? 	Evaluate 1	0-1
<ul style="list-style-type: none"> evaluate the author's level of expertise? 	Evaluate 2	0-1
<ul style="list-style-type: none"> identify the author's point of view? 	Evaluate 3	0-1
<ul style="list-style-type: none"> evaluate the reliability of a website? 	Evaluate 4	0-1
SYNTHESIZE: Reading to Synthesize Information		
Can students:		
<ul style="list-style-type: none"> summarize an important element from one website? 	Synthesize 1	0-1
<ul style="list-style-type: none"> summarize important elements from two websites? 	Synthesize 2	0-1
<ul style="list-style-type: none"> summarize important elements from a second set of two websites? 	Synthesize 3	0-1
<ul style="list-style-type: none"> summarize important elements from the websites in the research task to develop an argument? 	Synthesize 4	0-1
COMMUNICATE: Writing to Communicate Information		
Can students:		
<ul style="list-style-type: none"> make a wiki entry in the correct location? 	Communicate 1	0-1
<ul style="list-style-type: none"> use descriptive voice in an informational wiki? 	Communicate 2	0-1
<ul style="list-style-type: none"> include an appropriate heading for a new wiki entry? 	Communicate 3	0-1
<ul style="list-style-type: none"> compose and post a well-structured, short report of their research in a wiki? 	Communicate 4	0-1

Table 2

Holistic Rubric

https://www.ets.org/s/toefl/pdf/toefl_writing_rubrics.pdf

Appendix C Interview and CRR Excerpts from Cross-Case Analysis of Metacognitive Processes in ORW

Metacognitive Domain	Expert	Novice
<i>Analyzing the Rhetorical Situation</i>		
Knowing the purpose of a text	“I kept reminding myself of the original task, which was are you going to argue that decorative contacts are risky or not?” (Laura)	“So, I’m asking myself, are cosmetic contact lenses dangerous or not?” (Cody)
Anticipating the needs of the audience	“Because it’s a student wiki, and students are going to be drawn by celebrities and model their behaviour, I wanted to get students to think critically about the risk they put themselves at when they use these circle lenses that aren’t regulated by the FDA.” (Margaret)	“I know [my teacher] wanted a lot of pictures.”
Anticipating objections of the audience	Laura described how she went to great lengths to write a legal opinion that was “bullet proof” because she “anticipated this being challenged and reviewed at the highest levels.”	n/a
Analyzing the rhetorical situation multimodally	“I have a friend who says [referring to authoring Web pages] every page is page one. You never know where someone is going to go. You never know what they're looking for. And every sentence I write is written so that it might be the only sentence a person reads.” (Trina)	“Pictures are really powerful for the human brain, especially colour. It will draw you in more than text. When people see text and pictures, most of the time they will look at the picture first and then read the text.” (Darryl)
Analyzing the rhetorical situation	“I went back to the top [of the article] and started looking at why [the author] was	n/a

in order to read online texts	writing the story. I clicked on the hyperlink to the reporter's name, and she's a health reporter, so she's got the health angle, right? She's trying to write to, I can assume, parents mostly, but not strictly, about how to keep your kids safe. She's taking the safety angle. She's not trying to sell contact lenses, but rather she has a broad interest in keeping us safe." (Trina)
-------------------------------	---

Problem Posing

Posing the right problem	"I asked myself about her credibility. I clicked on the author's name to get details about her. I read the description about her. At first glance, it sounded like she was a credible source, but I realized she only worked as an assistant at an optometric clinic, which is not an ophthalmologist. An ophthalmologist is a doctor. An optometrist gives glasses. Whereas the author, she's primarily a journalist. To me she's not totally credible. If you're talking about health and harm to health, she does not have the medical training to say whether or not those are harmful." (Margaret)	"We were asked, 'Is Ms. Dilday an expert?' Yes, since there were a lot of quotes from experts in eye care. She was making a lot of good points. Since she was a specialist. She knew a lot about that topic." (Anna)
Complicating versus simplifying the problem	"Just trying to find the right source—that's probably the hardest part. Not necessarily in the ORCA because it gave you the specific website or article to find, but if you're giving a talk and you have to find out about lupus and headaches for example, where do I even start?"	"So I just typed zebras in the search and opened a whole bunch of links." (Cecilia)

That's probably the hardest thing to do is to say, okay, first I'm going to look in medical literature database and then I'm going to look at a nutrition database put together by other dieticians in the field that's evidence based. Then I might look at [Up to Date](#), which is a database for the physicians. Then I might look at blogs and websites of other health professionals and then I'll Google just to see what the average person is saying just to know what my patients are reading about too.”
(Judy)

Planning

Goal setting	Experts described simultaneously being involved in several writing projects at various states of completion; research, writing, goal setting, and reviewing goals were regular activities	Cecilia described how in the first week she will get books from the library or find online information, over the next few weeks she will write the material, and in the final week she will do editing
Triaging	“The other thing is, because I knew this was a child's wiki, I was less concerned about the credibility of the authors in those online research articles than I might have been if I were purposefully doing academic research. And even though this is academic, this is at a different level.” (Margaret)	“If it's not a school report, if it's something I do for fun, like air guns, I just look up where's the cheapest place to buy air guns. Things that are more low key, I just look it up online [not paying particular attention to the source]. But if I have to write an essay or do a book report or whatever, then I'll look up sites that are 99% reliable and find the sites that they [he gave Encyclopaedia Britannica as an example] recommend to branch off.”
Generating ideas	All of the novice processes plus annotating source material, thinking over a period of	Generating ideas by reading source material, free writing, using graphic organizers, using mediating

Organizing	<p>time, having a daily writing routine, referring to texts that one previously wrote or read (texts often stored in reference managers), following social media on subjects of interest, using technology</p> <p>“[How I organize my writing] depends on what the purpose is. When it's narrative like my blog post, I don't need an organizer as I was chronicling the narrative of my experience in the classroom. If I am presenting a proposal for a symposium or some kind of workshop, I definitely present it in a much more organized fashion and it's less experiential and more research- and purpose-based. What's the research agenda, where the research lies, what we would do, what we would use. So I use organizers and templates, but it depends entirely on the purpose.” (Monica)</p>	<p>documents, talking to others, writing about interests, and writing about experiences.</p> <p>“Our teacher had us do a spider web organizer. I figured out what the three paragraphs would be and what I would talk about in each, plus which facts I was going to use. We already had an outline for how we had to do it: we had to have a beginning, first, second, and third argument, and a counter argument, and conclusion.” (Jill)</p>
<i>Executing the Plan</i>		
Paraphrasing		<p>“With some articles we found, we had to copy some ideas and write it down from our head.” (Susan)</p>
Synthesizing	<p>“I wasn't actually willing to take a one side or the other given the information that I had. In each of the articles that I read, I found information that could help me develop a more complex understanding of [the safety of cosmetic contact lenses], not a yes it's harmful or no it's not.”</p>	<p>“The most important idea was that cosmetic contact lenses can hurt your eyes, so I looked for things that said how it could hurt your eyes and why.” (Jake)</p>

Generative synthesizing	<p>“It's probably only after working for 15 years that I've gotten better at reading a paper, critically analyzing it, and then reading a bunch of information on controversial things—there is so much controversy in nutritional reporting—and trying to come up with one statement, let's say to the residents: <i>this</i> is the best diet. Or research shows that <i>this</i> pattern of eating is superior. And that's a really hard thing to do. It's trying to balance everything you read plus your own professional opinion. I think that takes a lot of experience to do, but it's definitely an important skill. It's that process of evaluating the sources and summarizing information and then reflecting what the evidence says. Doing this, everyone can come to different conclusions because you can read it with your own personal bias and see it in a different way.” (Judy)</p>	n/a
Using a linear/ recursive process	<p>“My process is much more iterative and recursive and it's much more driven by the questions of the particular moment that I need to answer. I'm very aware of when I have a question and then I go find it quickly, go back, and go again. The process typically takes much longer because often times I'm writing things that require deep and careful consideration of facts and there's also the professional responsibility piece.” (Monica)</p>	<p>“I just find that when I'm writing a report or researching stuff online, I jot all my notes down and maybe only put them into two different groups before I summarize my notes. I get all my stuff down before I start working, which I find helps me focus because I don't have to go from writing, and then back into research, and back into writing again.” (Darryl)</p>

Revising locally vs. globally	<p>“We have a pretty elaborate review and comment process. Even before that, you have to get input from a lot of people. There are going to be subject matter experts. And if you know it's going to apply to a certain segment of the workers, than you talk to people in those facilities to get input. As one author, you can't possibly be expected to know all of those things. And that's just the preparation. In the review and comment, then you basically determine a list of people that you think need to be on the review. That could be anywhere from two to a dozen people, depending on how complex it is. It's usually because these people have different areas of expertise that need to be covered or different responsibilities that need to be covered. They provide all these comments. It's all done online now. They get a notice that here's the document and provide your comments.”</p> <p>Robert</p>	<p>“After the rough draft, we did peer editing. Once I type it I read through it and make sure there's nothing giant that I spelled something wrong or something.” (Jill)</p>
Reflecting	<p>“With blogs being a visual medium, having that excessively dense text makes it challenging for readers to stay with it and follow. So I think I would separate it with pictures. I would also make paragraphs much more condensed and if people wanted further information, they could click on a link and go to it, rather than having it all on this one page.” (Margaret)</p>	n/a

Appendix D Interview and CRR Excerpts Regarding Cognitive Processes Involved in Locating Online Information

Process	Expert Example	Novice Example
<i>Primary Locating: Locating Sources of Information</i>		
Searching by keyword		So I typed zebras in the search and opened a whole bunch of links. (Cecilia)
Searching by title	I was given the title to search for, so my preferred way would be to copy and paste that into the search bar or address bar. It's faster and more accurate. When I'm told to find a specific page, it should achieve that. (Colin)	Helen: "I wrote the title "Coloured lenses could trick, not treat, for Halloween" JC: "Why did you write the title?" Helen: "Because I wanted to look for what they (the assessment prompt) were asking exactly."
Being specific	I chose the one that most closely resembled the one I was given by Brianna [the avatar in the ORCA who delivers the assessment prompts]. Then I went to it and the title was a precise match. I was looking for a New York based article. This one said Seattle Times, so I skipped it. (Colin)	They told us to find the newspaper article. I put in the title and all the information, including the month, the date, location (Fiona)
Using trial and error	N/A	I would flip through some of the different websites and see which one had the best information. (Anna)

Using Boolean searches and operators	I put the name of the article in quotes when I search of the article in the search engine. (Laura)	N/A
	At work I look at court decisions or other tribunal decisions through online databases. These use Boolean search operators. That helps me search online references, laws, regulations, and policies. (Laura)	
Clicking on the top hits	I always look at the first one first because usually it's the most connected to what you're looking for. But in the [ORCA] instructions it said a New York newspaper and the first one was from the <i>Seattle Times</i> , so I knew the first one wasn't it. The second one had the same title and it was from the <i>NY Times</i> , so I knew it was that one. (Judy)	I usually take the first one because the first one is usually the one you're looking for. (Natasha)
Reading the resumé	I knew because I keyed into the title and as soon as I saw the title and the description underneath, I knew that was the right one. (Margaret)	Most people look at the titles that you click on. A lot of times if you're looking at something common or even anything at all, you find a bunch of the same ones and one that might not be what you are looking for. But, if you're looking for something specific, under the blue it will have little piece of what's written there. If you read that, you'll have a better idea of what's written there compared to just knowing the title of the website. (Darryl)

Locating divergent sources

During elections, if I wanted to find out a particular piece of information or have an argument with someone, I would search #canadianpolitics and #elections42 and I would go in and I could find any information I wanted about particular platforms, what was the hot topic at that moment. This is the place that I go to must effectively to find things and experts too. It's not just people on the streets. There are lots of interesting people in the Twittersverse. Also, if I'm conducting research and need help on a particular topic, because I work in digital pedagogy, I can Tweet #digped and say I'm working on this, and, 'Hey, does anybody have any sources for me?' Within a minute, someone will have answered me. (Kelsey)

Using the right search engine / mechanism

Work is a little different because I'm always trying to find government sources of information, what's been published to the public. One of the challenges with the government websites is that a lot of the links, even internally where I work, it's faster for me to Google to find it as opposed to using our own website. (Robin)

You can also look at the titles [in the search results]. The blue [hyperlinks]. And also what's underneath. It writes part of what is in the website. If you read it might tell you about the website. (Fiona)

I did use some books, but most of the books I found at the library were either really childish—it would mention there's this thing called the endocrine system and it produces hormones. That's no help at all. Or there were scientific manuals, but I'm not going to read all of that to do an eight-page project. So I did most of the research online because you can usually find more specific things, more manageable pieces. (Darryl)

I looked up how to make chocolate. A lot of things came up like bean to bar. Then I looked at some helpful videos, articles, and pictures. These were online. We didn't have any books. We looked online and asked friends and teachers. (Susan)

All novice participants with the exception of Darryl mention using Google as their only search engine.

First I looked up the article on Google [internal ORCA search engine]. Then I went to the Consumer Affairs website

		because that was who published it. If they didn't have it I would have tried searching Bing or Yahoo. (Darryl)
Backward / forward reference searching	I mean, I will use Wikipedia, but I'll usually use their references if I can find them. But Wikipedia gives you a good snap shot. (Robert)	At school and home, if I'm researching stuff, or even my own personal stuff, I usually use Wikipedia first, and then sites that Wikipedia recommends or Encylopedia Britannica, and the sites they recommend to branch off instead of starting from scratch. (Daniel)
Backward author searching	I began [my research] with some known researchers [...] because that's not my field of expertise (Catherine)	N/A

Secondary Locating: Locating Information within the Source

Skimming the text	I scan for information that I'm looking for. I look for headings and I skim over opening sentences. When I'm searching through a company's website or I'm searching for specific technical information that I might find online, I need to bypass the irrelevant information. (Colin)	I read through all of it first. (Susan)
	If I'm told to find something specific to whether the contact lenses do or do not harm your eyes, then I can scan each paragraph each paragraph or that group of paragraphs is going to be relevant.	I read the whole thing and thought about what were they trying to tell me. (Fiona)
		Then I skimmed the article because I thought it would help me find the answer. (Jake)

I'm focussed on a very specific task. I'm not asked to read and understand the whole article. In the case of the Miss Dilday's article, it's covering stuff like glow sticks and other treats and things. I knew that I could skip over that. I'm not asked to analyze the whole article or to understand everything she is saying about Halloween safety. I narrowed it down to a paragraph where there's a quote from an optometrist and I basically just reiterated that sentence in my summary. (Colin)

Locating what is relevant

I was linking back to the theme of the original task which was 'Are decorative contacts risky or not?' So, on one of the articles, that piece of it was just a minor piece. The article was really more about what a fashion trend this is, but I was linking back to the theme that I was working on. So all of the fashion stuff was blah blah blah, given that I was supposed to be supporting this one argument. To me that's going to be your one key piece because everything else is just fluff that doesn't link to the task you're supposed to be focussed on. (Laura)

If I'm told to find something specific to whether the contact lenses do or do not harm your eyes, then I can scan each paragraph each paragraph or that group of paragraphs is going to be relevant. I'm focussed on a very specific task. I'm not asked to read and understand the whole article.

That's the name of the article. You can click the name of the article and read about it. Also, those are ads, so I wouldn't pick those. (Susan)

I read this so there was nothing interesting or important there so I went to this paragraph [...]. (Cody)

In the case of the Miss Dilday's article, it's covering stuff like glow sticks and other treats and things. I knew that I could skip over that. I'm not asked to analyze the whole article or to understand everything she is saying about Halloween safety. I narrowed it down to a paragraph where there's a quote from an optometrist and I basically just reiterated that sentence in my summary. (Colin)

Metacognitive Practices: Using Metacognitive Practices to Locate Information

Narrowing down the search

I don't get overwhelmed by too much information anymore. Part of it is that I'm very comfortable in an online environment. [...] In truth, in my professional life, I'm typically reading on topics I know a lot about. So that's the other piece. There's a comfort level with the content. I often think we ask kids to read on these topics that they don't know anything about, so no wonder they're overwhelmed and don't know what to choose. How do you know what to pick if you don't know what to pick? Right? You don't know what you don't know; you don't know what to pay attention to. It's very analogous to why do toddlers put everything in their mouth. Why do they touch everything? Why can't they ignore anything in their environment? It's because they don't know what to pay attention to. They don't know what's important and what's not. It's like that for novice readers, too. Don't you know not to put that in

That's the name of the article. You can click the name of the article and read about it. Also, those are ads, so I wouldn't pick those. (Susan)

your mouth? It's analogous. Don't you know not to click on that? (Michelle)

Triaging

That's another constraint: nobody has infinite time. Ultimately the goal is to get [our resources and] and get on with the research. (Monica)

If it's not a school report, if it's something I do for fun, like air guns, I just look up where's the cheapest place to buy air guns. Things that are more low key, I just look it up online [not paying particular attention to the source]. But if I have to write an essay or do a book report or whatever, then I'll look up sites that are 99% reliable and find the sites that they recommend to branch off. (Darryl)

Taking advantage of technology

I use Endnotes to manage my references, but also to share references with my research teams so that we can all collectively use the same space to keep track of the same articles because I'm working on a team with 15 people right now. There's lots of overlap. I want to be able to see what's being used where. (Catherine)

N/A

Twitter has become my new research tool. I use pocket application when I find a really good article but don't have time to read it. I save it to my pocket. It's amazing. It's a curator of great information that's available on the Internet. (Margaret)

I'm working with a colleague at another university and we're working constantly in a

collaborative environment to hash out our thinking, our ideas. We're using several different channels for communicating. We're using a [slack channel](#). We have Google folders and Google Docs. (Monica)

Using shortcuts

I was given the title to search for, so my preferred way would be to copy and paste that into the search bar or address bar. It's faster and more accurate. When I'm told to find a specific page, it should achieve that. (Colin)

Brianna said please find the website 'Trick not treat for Halloween,' which was a news article. So I just typed that in because that was get me the most direct result. (Jake)

Remixing information (one's own)

I'm probably not researching the company's information online. Some things are routine enough that I know them. If I want to put hyperlinks or make references, I will also go back to previous emails I've written. We don't want to reinvent the wheel. It's been said to people before, let's have a consistent approach to our responses to the same type of issues. We may copy paragraphs from other emails in Outlook—we have a shared account. Or I'll go back into my own folders in Outlook to find those. I'll search through my emails to find relevant and similar responses I'm looking to use. (Robert)

N/A

Remixing information (from others)

You don't want to be writing advice again on something that has already been opined on. If someone else has already written a legal opinion on a certain topic, then you don't want to reinvent the wheel. We need the online tools to help us do the research and see what else is out there on the topic. (Laura)

N/A

Then it's also looking at what we [the nuclear agency] already have and doing a gap assessment of what the differences are. Also using other standards and using external standards from the U.S. (Robert)

Locating the primary source At work I would go to the initial sources. For example, what is the FDA [Food and Drug Administration] doing or saying? I wouldn't make the assumption that it's illegal—that's someone's interpretation of the law. I would want to see the actual law. (Laura) N/A

Locating with an eye to the audience to whom the researcher is communicating Just trying to find the right source—that's probably the hardest part. Not necessarily in the ORCA because it gave you the specific website or article to find, but if you're giving a talk and you have to find out about lupus and headaches for example, where do I even start? That's probably the hardest thing to do is to say, okay, first I'm going to look in medical literature database and then I'm going to look at a nutrition database put together by other dieticians in the field that's evidence based. Then I might look at [Up to Date](#), which is a database for the physicians. Then I might look at blogs and websites of other health professionals and then I'll Google just to see what the average person is saying just to know what my patients are reading about too. (Judy) N/A

Mid reading, I probably would have gone out to find diverse perspectives on particular things. I would have gone to see Lady Gaga's video. I would have gone to see what's the big fuss about her eyes. I would have done that because it helps to build some background knowledge that will enable me to see, get into the minds of the fans. (Monica)

Using a recursive process

I have questions, it drives a process of search, I gather some information, I read it, I'm always annotating as I read, whether I'm reading that document online or whether I'm reading a hard copy. And of course my writing flows from there. The reason for the inquiry is often because I'm writing some scholarship. Then I go read some more and I come back and do some writing and then I read some more. You know, there's that cycle. It's absolutely fundamental to my life as a scholar. (Monica)

N/A

Appendix E Interview and CRR Excerpts Regarding Cognitive Processes Involved in Evaluating Online Information

Process	Expert Example	Novice Example
<i>Primary Evaluation: Evaluating Argument Credibility</i>		
Evaluating the Argument (typically, arguments include a claim, evidence, and warrant)		
<ul style="list-style-type: none"> • Evaluating the claim 	<p>I also felt like that particular piece was actually offering a balanced perspective, so that was the other thing that I was looking for as well. To me, a balanced perspective is journalism that engages in complexity and multiple viewpoints, and it's an indicator of it being more valid. At least it's not providing a single perspective. (Monica)</p>	<p>The information [claims] seemed to make sense. Also, if they contradicted themselves, if they had opposing views on the same page, then it wouldn't be very reliable. Also, whichever one made more sense and had better points [claims]. (Cecilia)</p>
<ul style="list-style-type: none"> • Evaluating the evidence 	<p>I might read the same report or article, but often people will find what they want to in an article. Right? They can come to their own conclusions about the data—often incorrect ones. (Judy)</p>	<p>Since there were a lot of quotes from experts in eye care. She was making a lot of good points. Since she was a specialist. She knew a lot about that topic. (Anna)</p>
<ul style="list-style-type: none"> ○ Evaluating the degree of the evidence (i.e., is the evidence from a primary or secondary source?) 	<p>I would look at a newspaper to gather information, but I would never quote from a newspaper as a reliable source. At work, I would go to the initial sources, for example, what is the FDA (Food and Drug Administration) doing or saying. And I wouldn't make the assumption that it's illegal—that's someone's interpretation of</p>	N/A

	the law. I would want to see the actual law. (Laura)	
○ Evaluating the level of generalizability of the evidence	Also, what type of article is it? A meta-analysis? A randomized controlled trial versus just an observational study? How many other people found the same results? I would evaluate it on those terms versus somebody writing it from their own experience, you know, I try to evaluate these articles with an open mind but just because one person had the experience, it doesn't mean it's going to be the same for everyone. (Judy)	N/A
• Evaluating the warrant	Here [the journalist offers] some pieces of advice to deal with the various categories of risk [to wearing cosmetic contact lenses]. She basically said why these things can be okay. She talked to at least one person who was touted as an expert on all those things. (Robert)	When she [the reporter] said stuff, she had something to prove what she said. There were quotes from organizations and she said what they said, and she said what doctors said. (Natasha)

Secondary Evaluation: Evaluating Source Credibility

Evaluating the author / organization (collective author)

• Evaluating the author's expertise/authority	I clicked on the author's name to get details about her. I read the description about her. At first glance, it sounded like she was a credible source, but I realized she only worked as an assistant at an optometric clinic, which is not an	Cody: She's not an expert, per se. She is just someone who researches [eye health] and writes down her opinion. Author 1: Who would be an expert? Cody: A scientist or someone like that.
---	--	---

ophthalmologist. An ophthalmologist is a doctor. An optometrist gives glasses. Whereas the author, she's primarily a journalist. To me she's not totally credible. If you're talking about health and harm to health, she does not have the medical training to say whether or not those are harmful." (Margaret)

I gravitated towards that association there because it's not just one person—it's a collective. So I'll give a collective more value than an individual. (Trina)

- Evaluating the author's purpose

I went back to the top and started looking at why she was writing the story. So the story was broader. She's a health reporter, so she's got the health angle, right? She's trying to write to, I can assume, parents mostly, but not strictly, about how to keep your kids safe. She's taking the safety angle. She's not trying to sell contact lenses, but rather she has a broad interest in keeping us safe. (Trina)

I think government agencies would be reliable, like the poison control ones. It would not be reliable if the article didn't say where the author was from and it didn't give the name of the author. (Jill)

N/A

Evaluating the publisher

- Evaluating the authority of the publisher

I looked for the 'About' page to evaluate the validity, the trustworthiness, of the website. My first question was who created the website. So I went to the 'About' and learned that it was published

I think government agencies would be reliable, like the poison control ones. It would not be reliable if the article didn't say where the author was from and it didn't give the name of the author. (Jill)

by graduate students in Journalism at Northwestern University. I was like, okay, well obviously they're students in graduate school, they're engaged in the practice of journalism and disseminating information. (Monica)

- | | | |
|---|--|---|
| <ul style="list-style-type: none"> • Evaluating the modality of publication | <p>To be digitally critically literate, a person would look at the sources that we used today and think more deeply about that. They might think that the author is an expert and not look further. Just clicking on the person's name, and saying okay, this is what she says her background is. In this day and age, any person can have access to say whatever they want on the Internet. (Catherine)</p> | <p>If you put a book out there, it has to go through so many revisions and you have to get someone to publish it. Even for a newspaper article. But anyone can publish something online. (Darryl)</p> |
| <ul style="list-style-type: none"> • Evaluating the Uniform Resource Locator | <p>N/A</p> | <p>To know it's secure, the web address has to be https—the 's' stands for secure. (Natasha)</p> |
| <ul style="list-style-type: none"> • Evaluating the general impression of the electronic source (e.g., the page layout, the presence or absence of advertisements, the presence or absence of typos, etc.) | <p>I was scanning the site and it looked authentic enough to me. There wasn't 'click here to scan your disk for free.' There was none of that crap. It looked reasonably professional and that was convincing enough. I didn't see any obvious typos or grammatical scary things that give you that feeling. It was clean. (Trina)</p> | <p>Sometimes if [the source] isn't reliable, there would be lots of mistakes in the quotes. There could be grammar mistakes and stuff like that. Errors like if she says something at the beginning and she says something different at the end—that's something that isn't reliable. (Natasha)</p> |
-

Metacognitive Practices: Using metacognitive practices to evaluate information

Focusing on the task

I was linking back to the theme of the original task which was ‘Are decorative contacts risky or not?’ So, on one of the articles, that piece of it was just a minor piece. The article was really more about what a fashion trend this is, but I was linking back to the theme that I was working on. So all of the fashion stuff was blah blah blah, given that I was supposed to be supporting this one argument. To me that's going to be your one key piece because everything else is just fluff that doesn't link to the task you're supposed to be focussed on. (Laura)

I added the stuff that I read and found interesting. Who uses them [cosmetic contact lenses]? (Susan)

Background reading/research

- General background reading/research

Websites are a source of information. When I'm learning about something initially, that's usually where I start. I do that superficial read to just get an idea and a perspective on it. (Robert)

At school and home, if I'm researching stuff, I usually use Wikipedia first. (Darryl)

- Backward reference searching

I mean I will use Wikipedia, but I'll usually use their references if I can find them. But Wikipedia gives you a good snap shot. You can [use it to] find that information in academic journals. (Robert)

At school and home, if I'm researching stuff, I usually use Wikipedia first, and then sites that Wikipedia or Encyclopedia Britannica recommend, and branch off instead of starting from scratch. Once I've read that if I want more information I'll go to the bottom and see all the references that

they use and I'll use those, because they are most likely reliable. Wikipedia does have people who—anyone can change it—but they do have people who scan it for mistakes regularly. (Darryl)

- Backward author searching I began [my research] with some known researchers [...] because that's not my field of expertise (Catherine) N/A
 - Second-level backward reference searching So I read all of the previous legal opinions that had been given. Things started jumping out at me like this opinion is wrong, it is wrong in law. This lawyer did not do the proper analysis. Or basically this lawyer provided the wrong legal framework when she did her analysis. Now we've got a problem on our hands because [...] we've got advice that's gone out the door years ago that's wrong. (Laura) N/A
- If in trying to reconcile, I figure out why there's a discrepancy between our financial records and what's been made available publicly, then I chase the data until I figure out what is it that I'm missing or whether there an error that was made. (Robin)

Triangulating evidence / corroboration

- Corroboration with self (i.e., prior knowledge / personal experience)

In truth, in professional life, I'm typically reading on topics I know a lot about. So that's the other piece. There's a comfort level with the content. I often think we ask kids to read on these topics that they don't know anything about, so no wonder they're overwhelmed and don't know what to choose. How do you know what to pick if you don't know what to pick? Right? You don't know what you don't know, you don't know what to pay attention to. It's very analogous to why do toddlers put everything in their mouth? Why do they touch everything? Why can't they ignore anything in their environment? It's because they don't know what to pay attention to. They don't know what's important and what's not. It's like that for novice readers, too. Don't you know not to put that in your mouth? It's analogous. Don't you know not to click on that? (Monica)

She says a lot of other health related to information that I already knew just from school and my parents, like don't chew on glow sticks. If those things that I know to be true are true, than likely her other stuff is true too. (Cecilia)

- Corroboration with others

That way you can cross reference stuff. If you have two sites that are possibly unreliable but not associated with one another, you can go and read one and you read the other and it has pretty much the exact same information, it's likely that it's a fact. The more sites that you have to cross reference with, it's more likely that it's going to be true. (Darryl)

Self-regulation

Managing oneself in the fire hose of information is something I've learned to do. I don't get overwhelmed by too much information anymore. Part of it is that I'm very comfortable in an online environment. I understand, I can predict how things are going to go, I know how the technology should work, I'm not flustered or flapped when it doesn't go the way I want because I know that I've got strategies for figuring that out and I also know that I've got lots of other places I can go if this particular place doesn't give me what I need. (Monica)

It's really annoying. They [the websites] give us the complete opposite of what I'm asking and I can't find what I'm looking. Then I keep looking until I find the answer that I want. (Natasha)

Triageing

The other thing is because I knew this was a child's wiki, I was less concerned about the credibility of the authors in those online research articles than I might have been if I were purposefully doing academic research. And even though this is academic, this is at a different level. (Margaret)

If it's not a school report, if it's something I do for fun like air guns, I just look up where's the cheapest place to buy air guns. Things that are more low key, I just look it up online wherever, but if I have to write an essay or do a book report or whatever, then I'll look up sites that are 99% reliable and find the sites that they recommend to branch off. (Darryl)

Consulting a critical friend

I would rarely post like this without talking to someone who is an expert or without validating this somehow or having someone else check my work over, read it over, you know? I just wouldn't drop it in and be done with it. Part of it is that a lot of stuff I was writing about was

Abby: We had to edit it ourselves and then get a friend to edit it too.

JC: Did your friend give you suggestions?

Abby: She found a spelling mistake

nanotechnology and toxicology [participant is a technical writer, and so not an expert in this area] so you can't just throw your opinion out. I would always have someone look it over [because stakes are high] because you get your own context and sometimes when you're an outsider, you're missing a zone. You get the details, but you're missing connective tissue. Sometimes a third party can help you with that. To get it to go where you need it to go, you might need a bit more than that. (Trina)

Gaining multiple / diverse perspectives

Catherine spoke about seeking out Indigenous scholars, scholars of colour, and female scholars to gain a more diverse perspective on her research: "The thing that bothers me the most is the colonial adaptation, the thinking that Western learning is best."

N/A

Appendix F Interview and CRR Excerpts Regarding Cognitive Processes Involved in Synthesizing Online Information

Process	Expert Example	Novice Example
<i>Summarizing</i>		
Paraphrasing	<p>“When I was putting them all together in my head, I didn't use much of the wording that I had used in the notepad [paraphrased text]. At that point, based on the other three sites, I had come up with a view in my mind about what was important about the safety of these lenses. I put that into the chat thing as well, and I copied it and pasted it into the wiki afterwards.” (Colin)</p>	<p>“With some articles we found [for a school project], we had to copy some ideas and write it down from our head, so basically the same thing as what we did [on the ORCA].” (Susan)</p>
Finding key ideas	<p>And then I was really looking for information to either make a point or refute a point if I needed to. As I read the article, I kept referring back to the prompt [Are cosmetic contact lenses safe for your eyes?].” (Monica)</p>	<p>“I took the main idea from each [website]. The main idea is what the article is focused on. Sometimes there's a sentence, and the sentences after it explain about it. The first sentence is the main idea, I guess.” (Anna)</p>
Reducing information	<p>“If I'm told to find something specific about whether the contact lenses do or do not harm your eyes, then I can scan each paragraph each paragraph or that group of paragraphs is going to be relevant. I'm focussed on a very specific task. I'm not asked to read and understand the whole article.” (Colin)</p>	<p>“Since they [the ORCA] gave us a topic, it made it easier. Some of the information on the website isn't needed for the question that we're answering.” (Anna)</p>
	<p>“In this case, I knew the first two articles were saying that cosmetic contact lenses were dangerous and the last two were saying that they</p>	

were safe. I was basically just looking for the paragraph where it said that. There was a bunch of other information, but I was basically looking for what part of the article was saying that.”
(Robert)

Considering voice and style

“I dropped my summary in and then I look for the voice and then I fit my writing to the voice. I checked to see if they [authors of wiki] quoted any experts. They didn't, so I paired that out. Then I made it two sentences to fit with the other sections. I made a heading and bolded it. I wrote it in a way that would blend in.” (Tina)

“I just looked through my notes until I could find a couple of points that could stick together and then I wrote them down.” (Cecilia)

Synthesizing

Weighing the evidence

“I recorded points for and points against and then synthesize that into one overall message. It's not necessarily about being balanced, by looking at the weight of both sides. In this case, there was a lot more risk than benefit so that should come out.” (Judy)

n/a

Considering singular versus multiple perspectives

“I wasn't actually willing to take a one side or the other given the information that I had. In each of the articles that I read, I found information that could help me develop a more complex understanding of the issue, not a yes it's harmful or no it's not.” (Monica)

“I chose information that fit with the other pieces of information. So if there was a piece of information but it didn't fit with the other information, then I just wouldn't use it.” (Cecelia)

“The most important idea was that cosmetic contact lenses can hurt your eyes, so I looked for

things that said how it could hurt your eyes and why.” (Jake)

Generative Synthesizing

“I think that's the hardest thing to do because, to be honest, I do not remember doing that in school. It's probably only after working for 15 years that I've gotten better at reading a paper, critically analyzing it, and then reading a bunch of information on controversial things—there is so much controversy in nutritional reporting—and trying to come up with one statement to say to the residents that *this* is the best diet. Or research shows that *this* pattern of eating is superior. And that's a really hard thing to do. It's trying to balance everything you read plus your own professional opinion. Doing this, everyone can come to different conclusions because you can read it with your own personal bias and see it in a different way” (Judy). n/a

Appendix G Interview and CRR Excerpts Regarding Cognitive Processes Used to Communicate Online Information

Process	Expert Example	Novice Example
<i>Communicating Critical Discourse Knowledge</i>		
Writing accessibly	“If I'm writing for a community, I want to make sure what I'm writing is accessible and engaging for them.” (Monica)	n/a
Communicating inequity	“The thing that bothers me the most is the colonial adaptation, the thinking that Western learning is best.” (Catherine)	n/a
Communicating the historical perspective	“In this case, there were previous legal opinions that came up with different conclusions so I was pulling together a whole bunch of history in this file in order to give the client an interpretation that was compelling, that they could live with, and that they could use to make an informed decision.” (Laura)	n/a
<i>Communicating Discourse Knowledge</i>		
Using language appropriate to the discourse community	“I try to avoid the use of jargon unless it's already established. If it's going to make life easier for a detailed email, than I may introduce a particular word that may appear as jargon to the lay person, although if you're familiar with the regulations wouldn't seem like jargon anymore. With colleagues, there may be a lot of jargon and acronyms that we use. That's a typical thing in government. With the public, we'll avoid too	“My language was kind of informal. This was on a classroom page. I saw one of the other topics was video games, so obviously people aren't going to be super serious about things. This is going to be talking to other students who are using [the lenses]. I find that when I'm researching, it doesn't matter to me personally whether they use

many acronyms or jargon, or at the very least, we'll introduce them.” (Colin)

“The language varies, but it's supposed to be simple, yet technical standard terms. Industry and regulatory terms. I mean often you're using the language from a regulation, so you use those words exactly so that's it's clear that you're following that. Depending on who your audience is, it can be very technical at a low level. You know, if it's a technical person who is going to be doing a lot of these steps. Sometimes it's for a broad audience—it could literally be read by anybody in the company—and then it's less technical.” (Robert)

“The type of language I use depends on my audience. In this case, I know my audience is going to be teachers or parents who aren't necessarily going to be engaging with academic language.” (Catherine)

formal or informal language, but I know some of my friends find that they absorb information better if it's informal language, language that they use more often. That's why I didn't make everything super formal. I tried not to use a lot of slang, but if it were on Wikipedia I would have checked it out three times and got a thesaurus out. I've never posted on Wikipedia, but that's what I have done.” (Darryl)

Communicating Rhetorical Knowledge

Using rhetorical strategies in traditional and multimodal contexts

“If you read the introduction of my dissertation, I open it with the story of an online video game rape in real-time and listening to the YouTube video of what that's like. My goal was to shock [readers]. I want them to think about what they have just read and think what the heck is she doing? And then I slowly show them, this why you're here. This is what you need to learn.” (Catherine)

“I have a friend who says [referring to authoring Web pages] every page is page one. You never know where someone is going to go. You never know what they're

The conclusion was one of the things that I found really interesting. Because it was really interesting I put it at the end so people could leave the website with a pretty good exciting fact about eyes. (Susan)

“It's more convincing and it seems more professional if you use language that is more serious and concrete. I think making it seem like an actual threat, using words like 'cosmetic lenses can pose serious problems with your eyes,' that seems like—

looking for. And every sentence I write is written so that it might be the only sentence a person reads.”
(Trina)

Wikis have certain affordances and one of the affordances is to be able to communicate information using graphic images or video. (Monica)

When I post text online, I like to break it into smaller chunks with images that help anchor meaning.
(Margaret)

I pull up a couple of websites that I know I'd like to engage with via hyperlink throughout my text. It's an aesthetic choice so as not to overwhelm a digital reader. Instead, you can provide hyperlinks to provide additional information and easily connect to those.
(Catherine)

instead of 'they might hurt your eyes'—that seems it's actually a danger to you, instead of just something that can be waved off. Also, I find if you mention actual, like if you say 'they can cause harm to your eyes,' and don't say how they can, people will not accept it as much. People might just think, oh, my eyes might get itchy. So if you add they've been reported to cause bacterial infection, ulcers, and blindness, people will believe it more, especially if it's a serious effect, like blindness. The greater the effect is, the more likely it is that people will pay attention.” (Darryl)

Our teacher said that I really needed to make the reader feel like he was there, so I added things in about what it felt like, smelt like... Sammy got sprayed by a skunk in this particular retell, and my brother and I had to clean up. I used exaggeration... it smelled like thousand-year-old eggs! My first draft was pretty straightforward but then I revised it to add sensory details and exaggerations and it turned out pretty well.

One of my ideas was that the ice was melting and the polar bears live and fish on the ice. This picture shows the ice melting and the polar bear. (Helen)

Pictures are really powerful for the human brain, especially colour. It will draw you

Building an argument	For me, the first step is building a valid argument by building what sources I will use to connect my audience to the discussion. (Catherine)	more than text. When people see text and pictures, most of the time they will look at the picture first and then read the text. This picture I chose has text in it so you can go look and little arrows lead you in and you get an idea, almost a teaser because you don't know what most of this stuff means, and so you'll want to look in for an explanation. (Darryl)
		My slides have a lot of pictures and not a lot of text because I show [the pictures] and then explain to them orally. (Susan)
		Sometimes people wrote that it affects your eyes, but they don't say what effects those are or how it affects your eyes. If they say, oh, it affects your eyes... well, why should I believe it? So I decided to put all the things it can do to your eyes so that they know that, oh, it would actually have those effects so I won't do it [wear the decorative lenses]. It's making them [the reader] more sure about it. (Natalie)

Communicating Genre Knowledge

Choosing an appropriate genre for the discourse community	Our clients are seeking advice, and we give the answer in typically emails. We've really gotten away from writing formal written opinions. The reality is that resources are scarce and you're trying to be efficient. Meanwhile, the client just wants the bottom line. They don't care about all of your analyses. Not really.
---	--

Often, they just want the bottom line and they can get on with their business. [...]From the client's perspective, they want their options, the risks that go with different options, and then they want your recommendation as to how to proceed or your recommendation as to how to solve their problem.
(Laura)

Communicating in collaborative genres

Wikis exist for documentation. Teams of writers will document big suites of things. Like if there are tools built by hundreds of people, one writer is not enough. You all have to share a voice. You have to agree upon style guidelines. We will write to grade 8, we will use colloquialisms, we will not use humour. There are some industry standards but some of the bigger companies will have a style guide. Microsoft used to have a guide that was really popular.

They didn't use 'you' anywhere, but I decided that directness was okay. They're just spewing facts. Oh, here it says 'your eyes' so using second person is okay. You're addressing that question in your answer. How did I choose that voice? I really wanted to match what was there. I'm used to dropping stuff in other people's containers. For vocabularly, polymacron ? was the biggest one on the page, but it was defined. I didn't choose huge words, which I think was appropriate. I write to a grade eight level a lot. I went with that voice and kept out big words

Communicating genre features

In a formal legal opinion, you're going to set out the facts, you're going to provide the legal framework, and then you're going to do a legal analysis where you're applying the facts to the framework and then you're going to have a conclusion.

We use electronic medical records in our family health teams. So no matter who sees the patient, everyone would document the encounter or the appointment that they had with the patient afterwards. As a dietician, we have certain standards for things we have to put in our notes to make sure that everything is there. We follow an international dietetic terminology standard consisting of the SOAP format: Subjective, Objective points, Assessment, and Plan. (Judy)

It's a blog post. That implies narrative, story telling, experience, it can be evidence-based experience, evidence-based observations, but the style is conversational, even though it's within an academic realm, it's less conversational than I would have with my friends and family, and more conversation that I would have with fellow professionals who have a body of knowledge and that's why I was able to reference someone like Marzano. (Margaret)

The first paragraph [of an inquiry response] often thanks them for their inquiry and refers to what the subject of the inquiry was, in some general form. And the last paragraph is usually some basic salutation like thank you. We hope this was helpful. Contact us here if you have more questions. Those things are fairly standard, especially the last one. (Colin)

I actually don't think about what should come first and what should come last, but I find that I put the more necessary information at the beginning. The best information I put at the first, because that's what I know for sure. And then when I don't have anything else, I was like, okay. The other information I'm okay with I put at the end. I decided in my head without even realizing it. (Natasha)

Using Writing Process Knowledge

Organizing thinking	<p>“I start pencilling out a framework and making notes as I see different tangents popping up in my head.” (Laura)</p> <p>“There's definitely a lot of background reading and background thinking and I come to a point where ideas are synthesized in my head. I usually frame it with subtitles first. I have this big idea in mind, the idea that there's this revolution going on in text forms, and I'd already thought of this metaphor of a bridge. I use metaphor a lot in my thinking and so it comes out in my writing.” (Margaret)</p> <p>“I kind of like to get into the exploration of the ideas as I'm writing. I actually find that the process of writing itself is how I organize my ideas. I have to get into the communicating of ideas for my audience and iteratively come back always to what have I written and where do I need to go next.” (Monica)</p>	<p>“We already had an outline for how we had to do it: we had to have a beginning, first, second, and third argument, and a counter argument, and conclusion.” (Jill)</p> <p>“I jot all my notes down and maybe only put them into two different groups before I summarize my notes.” (Darryl)</p>
Using mediating documents	<p>“It's only recently that I got two computer screens set, which makes a big difference in how you work, being able to have reference documents on one screen and your writing notes on another screen. Otherwise you have your reference documents, your cases, and your law in paper on your desk, and you're writing on your computer screen.” (Laura)</p>	<p>“I looked up how are polar bears being affected by climate change. I found all these websites and then I put down the main ideas and made it into my own paragraph. I did take notes on paper. I would copy it and then put it in my own words, I put it down way different, but kept on referring to what it said.” (Helen)</p>

Using a linear
/recursive process

“I have questions, it drives a process of search, I gather some information, I read it, I'm always annotating as I read, whether I'm reading that document online or whether I'm reading a hard copy. And of course my writing flows from there. The reason for the inquiry is often because I'm writing some scholarship. Then I go read some more and I come back and do some writing and then I read some more. You know, there's that cycle. It's absolutely fundamental to my life as a scholar.” (Monica)

“I just find that when I'm writing a report or researching stuff online, I jot all my notes down and maybe only put them into two different groups before I summarize my notes. I get all my stuff down before I start working, which I find helps me focus because I don't have to go from writing, and then back into research, and back into writing again.” (Darryl)

Using a collaborative
process

“Wikis exist for documentation. Teams of writers will document big suites of things. Like if there are tools built by hundreds of people, one writer is not enough. You all have to share a voice. You have to agree upon style guidelines. We will write to grade 8, we will use colloquialisms, we will not use humour. There are some industry standards but some of the bigger companies will have a style guide. Microsoft has a guide that was really popular.” (Tina)

“It was easy because for the question they gave, I basically wrote that down [for the heading]. I basically follow what they say. For how long it should be, I looked at the other ones and I thought I can't do it too big and I can't do it too small. So, I basically did it medium and put all the information that I needed down.” (Darryl)

“It's more taking turns, like we're not both adding to a document at the same time—that would be messy. You try and have a bit of project management, like I'll take this piece, you take that piece. Or I'll take the first pass on it, and then you review what I've done and add to it.” (Laura)

“There's a comment box underneath the article I wrote [for an online journal]. It was published last February, but then I used it in a class and then her students engaged with it. I had a whole new round of

Twitter discussions. There's definitely a collaborative nature to it, there's also the continual learning process that accompanies it. The writing process is never finished.” (Catherine)

“I do a lot of collaborative writing. As a for instance, I'm working on a protocol for ethics review. I'm working with a colleague at another university and we're working constantly in a collaborative environment to hash out our thinking, our ideas. We're using several different channels for communicating. We're using a slack channel. We have Google folders and Google docs. And of course we have the documents that our institutions require us to fill out, so we're collaborating on building PDF forms that aren't necessarily shareable. My co-author gave a first pass on a lot of the content and then I spent a few hours that week going back in to, say, he'd write in parentheses, we need citations here. Or putting in little comments. There's the text that gets written in the documents, and then there's this constant stream of comments along the side that are metacognitive. Here's the text, and here are the questions that as a writer that I'm thinking about that maybe you as my co-writer would have something to say about. What do you think about this? Or how should we say this? Or that?” (Monica)

Triaging

“It depends on the significance of the issue and the risks and the profile of the issue. In this case I was getting my boss, three levels up from me was reviewing this at the end of the day. Other things that

“If it's not a school report, if it's something I do for fun, like air guns, I just look up where's the cheapest place to buy air guns. Things that are more low key, I just look it up online [not paying particular attention to

are lower risk, lower profile, I would read it through once and hit send.” (Laura)

“With my colleague we decided this is good enough--writing can always be better—let's just, there's a form that needed to be filled out and it's good enough. So we decided to share it with the adjudication committee, the review board, we're going to get feedback from them, and reintegrate it into the next version. It's never a 'this is done,' it's more we've done what we can in the amount of time that we have, I made that's another constraint: nobody has infinite time, so we're going to send it out and get feedback, that starts a conversation, so we can address those issues. Ultimately the goal is to get is approved so we can get on with the research.” (Monica)

When I come up with ideas, it could be on my phone or on my note pad. Sometimes on my phone, I'll email it to myself. It's usually in point form. Like I wouldn't write a paragraph. It's just something quick that I write down so that my brain remembers that I had that conversation with myself in my head the day before. I'll look at it and remember, oh yeah, that thing! And then I'll write it.

the source]. But if I have to write an essay or do a book report or whatever, then I'll look up sites that are 99% reliable and find the sites that they [he gave [Encyclopaedia Britannica](#) as an example] recommend to branch off.” (Darryl)

Communicating Subject Matter Knowledge

“We have regulations that have been published and we have information online and we have people and companies that are subject to those regulations, and often they're looking for additional information or they're aware that the regulations exist but they're not

really aware of them. It's complicated legislation.”
(Colin)

“Things started jumping out at me like this opinion is wrong, it is wrong in law. This lawyer did not do the proper analysis. Or basically this lawyer provided the wrong legal framework when she did her analysis. Now we've got a problem on our hands because we give advice as the Department of Justice, and Justice speaks with one voice, you know, this disembodied entity that is the Department of Justice giving advice, and we've got advice that's gone out the door years ago that's wrong.” (Laura)

“In one job I worked for a program making company wide documents, processes, and procedures and standards and things like that. It was the Nuclear Criticality Safety Program. We would have a lot of internal, higher, overarching procedures on things that had to be followed. This would have to meet all those things. So you would have to look up those things and see where it falls under those. We had some external standards from the U.S. that we had to follow, so we incorporated those.” (Robert)

**CHAPTER 5 - USING THE INTEGRATED DESIGN AND APPRAISAL
FRAMEWORK TO APPRAISE A LARGE-SCALE ASSESSMENT OF
NEW LITERACIES**

Using the Integrated Design and Appraisal Framework to
Appraise a Large-Scale Assessment of 21st Century Literacies

Julie A. Corrigan
University of Ottawa

Author Note

Julie A. Corrigan, Faculty of Education, University of Ottawa.

This research was generously supported by a Social Sciences and Humanities Research Council of Canada Bombardier Canada Graduate Scholarship, Ontario Graduate Scholarships, and internal funding from the University of Ottawa.

Correspondence should be addressed to Julie A. Corrigan, Faculty of Education, University of Ottawa, 145 Jean-Jacques-Lussier Private, Ottawa, ON K1N 6N5, Canada

Abstract

This study used a holistic approach—wherein the interconnected issues of validity, reliability, and fairness are considered within one model—to appraise a large-scale assessment of New Literacies. Specifically, this study appraised the Online Research and Comprehension Assessment (ORCA), with a particular emphasis on the writing domain. The ORCA challenges seventh graders to solve research problems (e.g., Does playing video games harm your eyes?) by locating, critically evaluating, and synthesizing online information in order to communicate their results in online genres such as email and wiki. For this study, I invited 10 novice and 10 expert participants to complete an ORCA. Including participants from across the novice-expert spectrum enabled me to observe a wider range of response processes. I used Cued Retrospective Reporting, semi-structured interviews, Venn diagrams, surveys, and writing artefacts to investigate the response processes elicited by the ORCA and to compare and contrast those to the writing practices that participants used in their school, work, and/or personal lives. I also completed an extensive analysis of the sample of observations permitted by the ORCA juxtaposing those with the target domain. Results of this study indicate that the ORCA provides an important form of assessment data regarding 21st century literacies previously neglected on traditional assessments. Limitations of the ORCA such as construct-irrelevant variance and construct underrepresentation are also explored.

Key words: New Literacies; writing assessment; validity; standardized assessment

Using the Integrated Design and Appraisal Framework to Appraise a Large-Scale Assessment of 21st Century Literacies

Increasingly, standardized literacy and writing assessments (e.g., the Organisation for Economic Co-operation and Development's [OECD's] Programme for International Student Assessment [PISA]; the Ontario Secondary School Literacy Test [OSSLT] in Canada; and the National Assessment of Educational Progress [NAEP] in the United States) are moving from pen and paper to computerized and even Internet-based testing formats. As online versions of large-scale literacy assessments emerge, there is an increasing need for validity studies to ensure that scores from these tests are used and interpreted correctly, and to examine the consequences of these tests, both intended and unintended. One such assessment is the Online Research and Comprehension Assessment (ORCA) developed by the University of Connecticut's New Literacies Research Lab and funded by the Institute for Education Sciences. The ORCA challenges seventh graders to solve research problems (e.g., Does playing video games harm your eyes?) by locating, critically evaluating, and synthesizing online information in order to communicate their results in online genres such as email and wiki. Although the ORCA is a hybrid reading-writing assessment, in this study I focus on the writing dimension of the assessment, considered through the lens of a bioecological construct of writing (Corrigan & Slomp, 2017). Finally, not only do I present a validity argument for the ORCA, but I use a much-needed systematic model known as the Integrated Design and Appraisal Framework (IDAF; Slomp, 2016; Slomp, Corrigan, & Sugimoto, 2014) to do so, thus serving as a model for future validation research.

What is the ORCA?

According to the test designers, the ORCA was created to be a reliable, valid, and practical assessment of online research and communication ability among seventh-grade students, particularly in the subject of science (Leu et al., 2012). As such, like other literacy assessments, it is a hybrid reading-writing assessment. It should also be noted that the ORCA is not a high-stakes test. Rather, the ORCA is unique in that it could be considered a large-scale formative assessment. In fact, teachers and school administrators do not even receive reports of individual student performance on the ORCA. Rather, they are given class and/or school-level reports (Appendix H) so as to inform school- and system-level program decisions. According to Gareis and Grant's (2015) definition, the ORCA can be considered a formative assessment because it informs teachers' near term instructional decisions and it is considered low stakes. Of course the distinction between formative and summative assessment creates a false dichotomy. Assessment can be both formative and summative simultaneously. And so, the ORCA can also be considered a summative assessment because no feedback is directly provided to the student to improve learning and because the results are used to communicate the nature and degree of learning to stakeholders, primarily, other than the students themselves. Thus, depending on how the data is used, the ORCA is at once both a formative and summative assessment.

The current iteration of the ORCA contains four research scenarios (one per assessment version) that challenge students to investigate different life science topics through doing online research and communicating their results using online writing tools, particularly email or wiki. In each of these scenarios, students were presented with a problem that focused on the domains of health and human body systems, an area common to many seventh-grade science curricula. An example of a research question presented to participants from one of the ORCA scenarios is,

“How do energy drinks affect heart health?” The scenarios were framed around two types of research: Learn More About (LMA) and Investigate Conflicting Claims (ICC). Half of the scenarios presented the research problem to students via an email message from the school board president (LMA scenarios) and half via a class wiki with a message from the teacher (ICC scenarios). LMA scenarios ask students to learn more about the research topic and to form a main idea about what they learn. ICC scenarios, on the other hand, ask students to investigate two sides of an issue and to take a position.

Table 5-1 presents the four LESC (i.e., Locate, Evaluate, Synthesize, and Communicate) scenarios used in the ORCA.

Table 5-1 The Four LESC Scenarios by Topic

Topic	Student Research Question	Type of Research	Communication Tool
Energy Drinks	How do energy drinks affect heart health?	Learn more about	Email
Heart-Healthy Snacks	How do snacks affect heart health?	Learn more about	Email
Contact Lenses	Do decorative contact lenses harm your eyes?	Investigate conflicting claims	Wiki
Video Games	Do video games harm your eyes?	Investigate conflicting claims	Wiki

This study uses a format of the ORCA known as the ORCA-Virtual, while past iterations of the ORCA have included formats called the ORCA-Open and ORCA-Multiple Choice (Leu et al., 2012). The ORCA-Open represented the least restricted information space as assessment tasks used an open Internet space and the responses were in an open format (i.e., the task was a

performance-based assessment, as opposed to a multiple choice one). The ORCA-Open format was eventually discontinued due to the instability of the assessment context. Unlike print-based texts, the form, content, and structure of online texts is in a constant state of flux. This presented a problem for the ORCA-Open because target websites used in the assessment were altered or would disappear altogether. In response to this challenge, the ORCA-Virtual was developed using a restricted Internet environment (such as ones used in online banking in the sense that, once you are on the bank's website, you cannot link to outside websites without logging off). This assessment format was designed to provide a more stable assessment context in that the while it takes place on the Internet, it was not open to the Internet at large and was stable over time. Another current format used, although not investigated in this study, is the ORCA-Multiple Choice. The ORCA-MC is the most restricted format in terms of its assessment context (it used screenshots of websites, but was not open to the Internet at large) and its response format (multiple choice).

In the ORCA-Virtual, students conducted their research and selected information from pre-determined websites from the project's search engine named "Gloogle" (**Error! Reference source not found.**). As the tasks were in a constructed-response format, the ORCA-Virtual is a performance-based measure. The LESC research problems appear within a Facebook-like environment via avatars named Brianna and Jordan (Figure 5-2) who are introduced as students from another school. The questions did not appear in a linear sequence according to their domain (i.e., Locate, Evaluate, Synthesize, and Communicate). Rather, a more natural and logical sequence was used according to the nature of the research task. Students were guided to engage in their online research and writing tasks via requests and questions from Brianna and Jordan. The ORCA culminates in a task requiring students to communicate their finding

regarding the research problem via either writing an email or editing a wiki entry. A video of a high-performing student taking the ORCA-Virtual, “Do decorative contact lenses harm your eyes?”, may be viewed by clicking on the video link below (Figure 5-5).

Figure 5-1. ORCA-Closed format with 'Gloogle' search engine

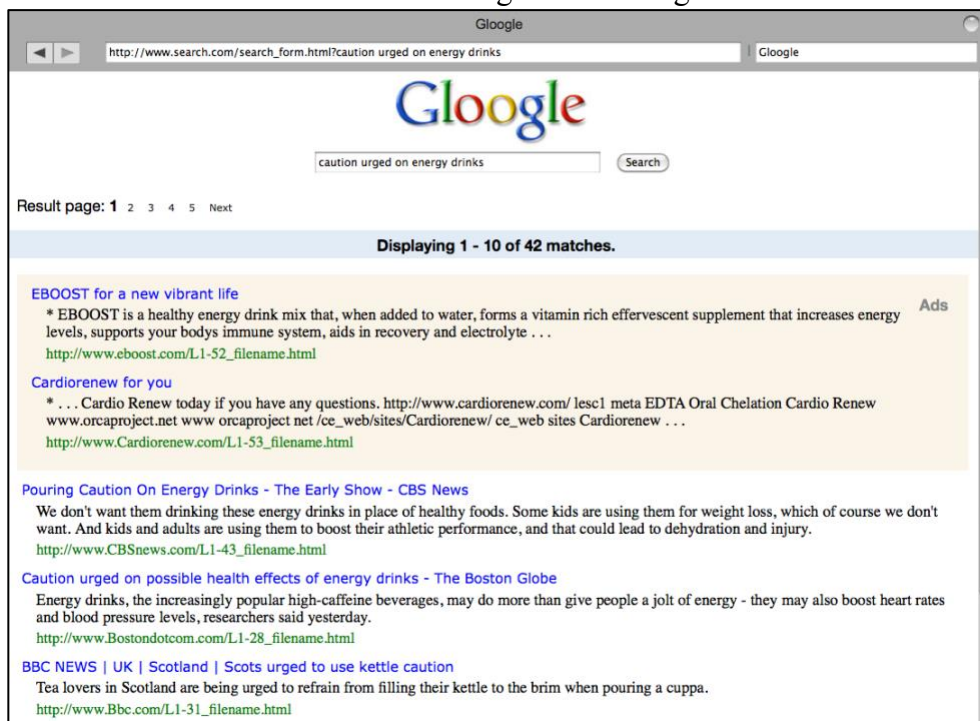


Figure 5-2. Energy Drinks, Introduction: The introduction portion of an email LESC, along with the email the student clicks on to learn about her research task.

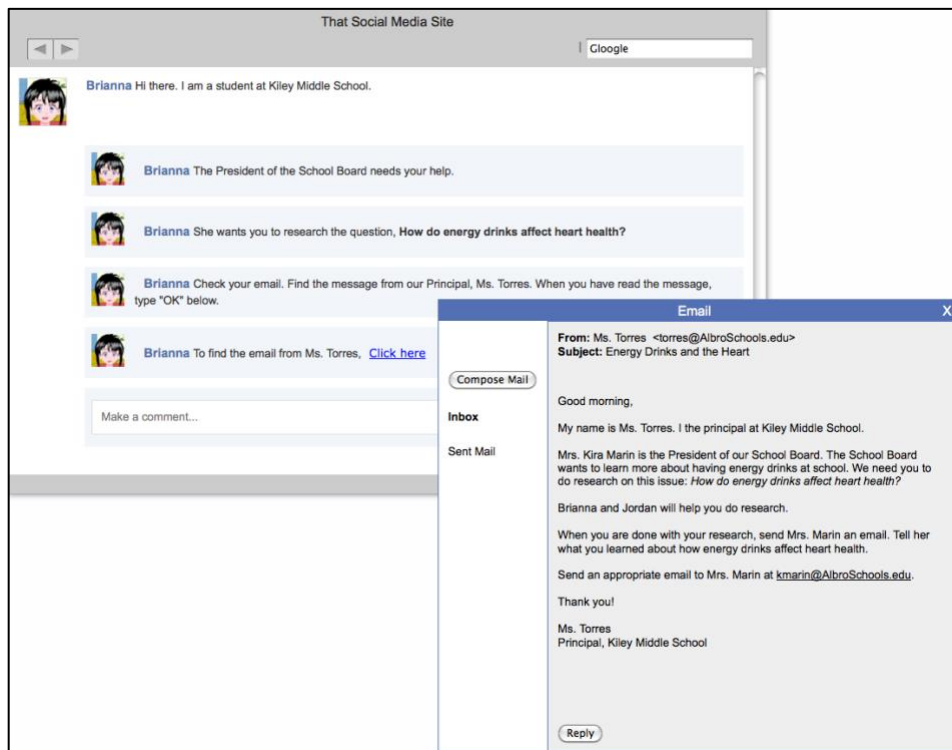


Figure 5-3. The final task on the ORCA challenges students to communicate the results of their research findings either in an email or wiki format.



Figure 5-4. A student edits a wiki page in the cosmetic contact lens LESC scenario.

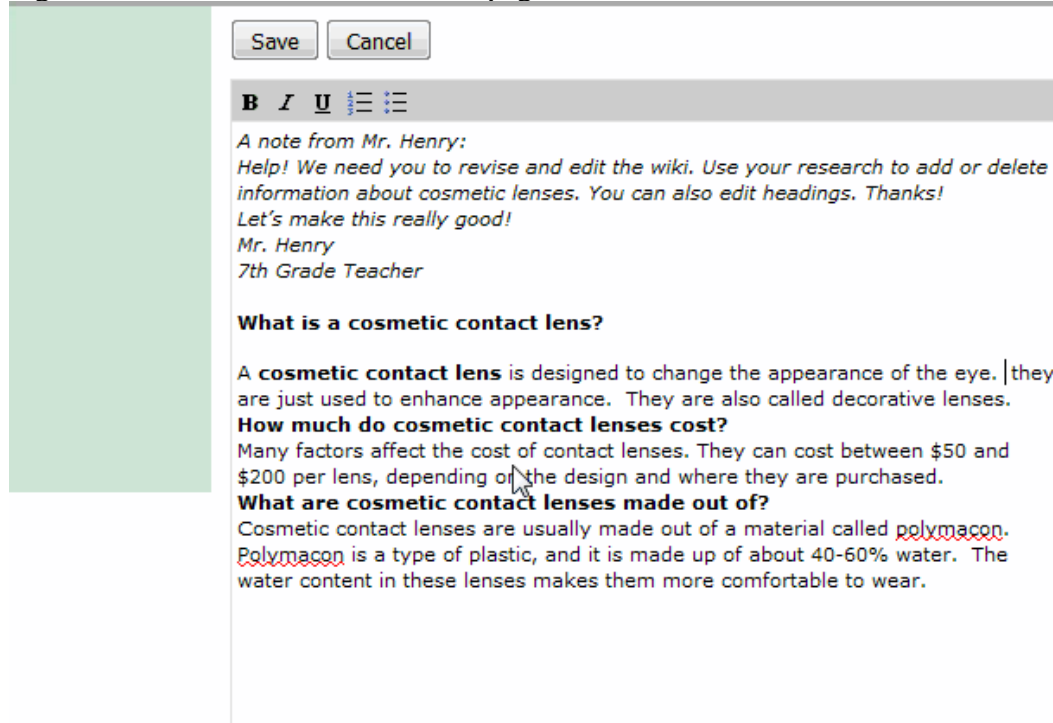
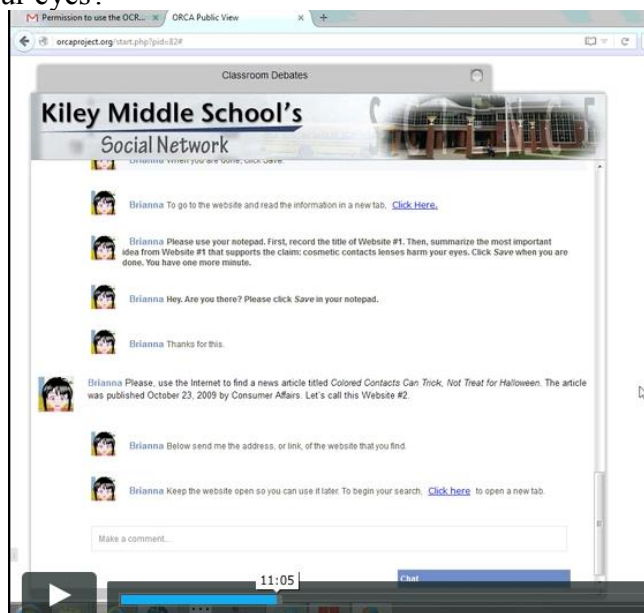


Figure 5-5. Video link to a high-performing student taking the ORCA-Virtual “Do cosmetic contact lenses harm your eyes?”



ORCA responses are assessed using a multidimensional scoring model for a total ORCA score of 16 points with 4 points per dimension each of locate, evaluate, synthesize, and communicate (LESC). Scoring procedures will be explained in greater detail later in this article. The following series of four tables outline the scoring criteria for the Investigate Conflicting Claims ORCA tasks (i.e., Cosmetic Contact Lenses and Video Games) for each of the domains of Locate, Evaluate, Synthesize, and Communicate. While the tables are presented in this order, as mentioned, the questions did not appear in a linear sequence according to their domain.

Table 5-2 Locate Scoring Criteria for the Cosmetic Contact Lenses and Eye Health Task

Item	L1	L2	L3	L4
Purpose of Stem	Can the student locate the correct email message in an inbox or the correct section of a wiki?	Can the student use appropriate key words in a search engine?	Can the student locate the correct site in a set of search engine results?	Can the student correctly identify and share the website addresses from two different search tasks?
Item Stem	First, go to the wiki. Then, read the section on cosmetic contact lenses. When you are done, tell me you are finished. To go to the homepage of the wiki, click here .	Please use the Internet to find a news article titled <i>Colored Contacts Can Trick, Not Treat for Halloween</i> . This article was published October 23, 2009 by Consumer Affairs. Let's call this Website #2. To begin your search, click here to open a new tab [opens Google].	Please use the Internet to find another website that says cosmetic contact lenses <u>do not harm</u> your eyes. It is titled <i>What Big Eyes You Have Dear, but Are Those Contacts Risky?</i> This article was published on a website from a New York newspaper. Let's call this Website #4. To begin your search, click here to open a new tab [opens Google].	For both the first and second URL: Below, send me the address, or link, of the website that you find.
Scoring Criteria	On the first search task, the student locates the correct email message (email) or wiki section (wiki) on the first click.	On the first search task, student uses appropriate key words based on information provided by Brianna.	On the first search task, the student selects a correct site from the search results on the first click when the correct site is not in the	The student provides two correct website addresses from the two search tasks (scored on exact URL entry).

<p><i>Restricted and Unrestricted Tasks: Same criterion for both tasks.</i></p>	<p><i>Restricted Task (Energy Drinks and Contacts):</i> On the first search task, the student uses appropriate key words, entering the article's title as the search term.</p> <p><i>Unrestricted Task (Video Games and Heart Healthy Snacks):</i> On the first search task, the student uses appropriate key words, entering both topic and claim as search terms.</p>	<p>first position.</p> <p><i>Restricted:</i> For the first search task, did the student select the correct, target website from search results on the first click?</p> <p><i>Unrestricted:</i> For the first search task, did the student select one appropriate site related to the topic from search results on the first click?</p>	<p><i>Restricted:</i> Did the student correctly provide both correct website addresses from the two search tasks? (scored on exact URL entry)</p> <p><i>Unrestricted:</i> Did the student correctly provide two topically relevant website addresses from the two search tasks? (scored on exact URL entry).</p>
---	---	--	--

Table 5-3 Evaluate Scoring Criteria for the Cosmetic Contact Lenses and Eye Health Task

Item	CE 1	CE 2	CE 3	CE 4
Purpose of Item	Can the student identify the author of a website?	Can the student evaluate the author's level of expertise?	Can the student identify the author's point of view?	Can the student evaluate the reliability of a website?
Item Stem	Can you tell us who is the author or creator of this website, " <i>Avoid the Tricks that Come with the Treats: Tips for a Healthy, Happy Halloween</i> "?	Is Julia Dilday an expert on cosmetic contact lenses or eye health? How do you know?	What is the author's point of view? What words or images does the author use to support that point of view?	Is the information at this website reliable? How do you know?

Scoring criteria	The student correctly identifies the author of the website (First, last, or first and last names accepted. Spelling variants accepted.)	The student judges the author's level of expertise AND provides an appropriate supporting detail about the author's level of expertise (or lack thereof).	The student identifies the author's point of view AND provides one accurate and specific piece of evidence from the webpage about words or images that supports this determination of the point of view.	The student evaluates the reliability of the website (or lack thereof) AND provides one logical and accurate explanation to support their conclusion about either: author expertise, publisher trustworthiness, or research findings.
------------------	---	---	--	---

Table 5-4 Synthesize Scoring Criteria for the Cosmetic Contact Lenses and Eye Health Task

Item	S1	S2	S3	S4
Purpose of Stem	Can the student summarize an important element from one website?	Can the student summarize important elements from two websites?	Can the student summarize important elements from a second set of two websites?	Can the student summarize important elements from the websites used in the research task to develop an argument?
Item Stem	Summarize the ONE, MOST IMPORTANT, idea you found from Website #1 to support this claim: Cosmetic contact lenses harm your eyes.	Below, send me a summary of what you learned from BOTH Website #1 and #2. Include important details	Below, send me a summary of what you learned from BOTH Website #3 and Website #4. Include important details	Now, take a position. Do cosmetic contact lenses <u>harm</u> your eyes? Send me a summary of what you think after reading all FOUR sites.

	Use your own words.	from both websites.	from both websites.	Include important details from the websites that explain your thinking.
Scoring Criteria	Does the student provide a summary of one important element from the first website using their own words (at least 3) that supports the given claim ("Cosmetic contact lenses HARM your eyes"). Can be exact same as notepad because it is the notepad.	Using notes from the notepad or information from the sites themselves, can students use their own words (at least 3) to integrate one detail from each of the first two websites that supports the given claim ("Cosmetic contact lenses HARM your eyes")?	Using notes from the notepad or information from the sites themselves, can students use their own words (at least 3) to integrate one detail from each of the second two websites that supports the given claim ("Cosmetic contact lenses HELP your eyes")?	Student uses own words (at least 3) to provide a claim/argument and one supporting detail from each of two different websites. Student uses own words (at least 3) to provide an argument (that takes a position) AND two pieces of evidence, one detail from each of two different websites.

Table 5-5 Communicate Scoring Criteria for the Cosmetic Contact Lenses and Eye Health Task

Item	C1	C2	C3	C4
Purpose of Stem	Wiki Task: Does the student make a wiki entry in the correct location?	Wiki Task: Does the student use descriptive voice in an informational wiki?	Wiki Task: Does the student include an appropriate heading for a new wiki entry?	Wiki Task: Does the student compose and post a well-structured, short report of their research, including sources, in a wiki?
Item Stem	Now, write a short report on	Now, write a short report on	First, add a new heading in a	Be certain to explain your

	the class wiki. Use what you learned to add to the wiki for Mr. Henry and his students.	the class wiki. Use what you learned to add to the wiki for Mr. Henry and his students.	location that makes sense. Then take a position: Do cosmetic contact lenses harm your eyes?	thinking. Use evidence from the websites that supports your position about the question: Do cosmetic contact lenses harm your eyes? Please include your sources. Use website titles.
Scoring Criteria	Student types a wiki entry in the correct location that includes any relevant content in the correct location, either: 1) below the other sections if no heading; 2) anywhere in the wiki below the Mr. Henry message if they do have a heading on a separate line	Student composes an informational wiki entry in a descriptive voice.	Student composes an appropriate heading in their wiki entry, containing both aspects of the topic (e.g., snacks and heart or health; video games and eyesight or eye health). An appropriate heading must be on a separate line.	Student composes and saves a wiki post with one relevant claim and two pieces of supporting evidence. They also include at least two sources, listing the names of the websites or the URLs.

Previous Validity, Reliability, and Fairness Research on the ORCA

Previous validation work regarding the ORCA completed by the New Literacies Research Team included the following: cognitive labs over two years with approximately 300 students; pilot testing with approximately 1600 students; and, a panel review of the ORCA by measurement and online research and comprehension experts. Several modeling procedures, including classical test score analysis and latent variable modeling strategies, were used to

evaluate the reliability and validity of scores (Leu et al., 2012). Exploratory Factor Analysis was used to determine which items were too easy or too difficult based on the criterion of factor loadings $< .20$, and such items were subsequently revised. Further, previous reliability estimates by measurement experts determined that the KR-20 for the full scale ranged from .86 to .90 for the ORCA Virtual format, depending on the version of the ORCA. The KR-20 was .80 for the cosmetic contact lenses version used in this study (Leu et al., 2014). For each of the separate component values, the KR-20 values were low to moderate, which would be expected with only four items per component. These KR-20 values were .56 for Locate, .43 for Evaluate, .67 for Synthesize, and .51 for Communicate (Forzani, 2016). From the publicly available documentation on the ORCA, little has been written about its fairness. I will expand the scope of the preceding validation work to include a holistic and interconnected examination of validity, reliability, and fairness as analyzed through the IDAF (explained in a subsequent section) in order to better understand how ORCA test score data should be used and interpreted.

The Need for Large-Scale Assessments of New Literacies

This study is framed by New Literacies, a dual-level theory that encapsulates the changing natures of literacy, instruction, and assessment (Leu, Kinzer, et al., 2013). Upper-case New Literacies is an overarching theory that describes the ways in which the Internet and its associated technologies have profoundly disrupted the researching, teaching, and assessment of literacies. Multiple lower-case new literacies theories account for the more specific, and constantly evolving nature of New Literacies. For example, the ORCA is based on the new literacies of online research and comprehension (described forthwith; Leu, Forzani, et al., 2013).

The development of large-scale New Literacies assessments is in its nascence. In the meantime, writing continues to be assessed using traditional pen-and-paper tests that are

increasingly removed from the day-to-day practices of communicating in a knowledge society (Stagg Peterson et al., 2011). Stagg Peterson, McClay, and Main (2011) remark that large-scale assessment designers have a unique opportunity to make multiliteracies (and New Literacies, for that matter) mainstream in the classroom. Indeed, logic suggests that if it is tested, it will be taught. In assessment, this refers to the concept of washback, the extent to which a test influences what is, or is not, being taught in the classroom (Hamp-Lyons, 1997; Messick, 1996; Morrow, 1986).

Current research suggests that writing practices associated with New Literacies and multiliteracies are frequently ignored on large-scale testing (Slomp et al., 2014; Stagg Peterson et al., 2011). For example, when my colleagues and I (Slomp et al., 2014) completed a systematic analysis of publicly available grey literature and peer-reviewed studies on large-scale writing assessment across Canada, we discovered that standardized writing assessment constrained the construct of writing in a number of ways, including by ignoring multimodality and multiliteracies (Slomp et al., 2014). A logical extension of these findings would be that, until large-scale assessments attend to the increasingly diverse social, multimodal, and technological practices of online writing, they will be given less priority in the classroom. Further, the more contextualized and authentic these assessments can be, the greater the possibility that they will mirror the actual literate practices of literate people using technology to communicate in their personal, public, and professional lives—and thus, the greater the validity of these tests.

The ability to write effectively in online contexts is growing increasingly important as the labour market moves toward knowledge-based industries (Yan, 2005), thus necessitating highly-skilled workers to "successfully use and adapt to the rapidly changing information and communication technologies and contexts that continuously emerge" (Leu et al., 2013, p. 1163).

Not only are writers challenged to adapt to new communication technologies, but also to the multiple linguistic and sociocultural writing practices that they will encounter as they communicate within our global village (Cope & Kalantzis, 2009; Corrigan & Slomp, 2017). Further, aside from their professional success, those who can use online writing practices effectively will be better empowered to participate in the political, social, and economic practices that are increasingly merging across analogue and digital spaces. Everyday examples of this are apparent as people participate in online learning, banking, shopping, dating, employment seeking, citizen journalism, trip planning, and the list goes on. More radical examples of this are demonstrated by revolutions such as the Arab Spring, Occupy Wall Street, and Idle No More where citizens extensively relied on social networking tools like Twitter and Facebook to mobilize supporters and organize protests (Bastos, Travitzki, & Raimundo, 2012; Donkin, 2013).

The Need for Validity Studies of Large-Scale Assessments of New Literacies

Test designers are beginning to respond to the need for large-scale digital/online writing assessments. As previously mentioned, such assessments have been developed for the OECD, NAEP, and for the OSSLT. Concomitant with the development of these tests arises the need for validity research examining the uses, interpretations, and consequences (both intended and unintended) of data from these tests. While peer-reviewed validation studies of large-scale writing assessments abound in journals such as *Assessing Writing*, there is a paucity of these studies regarding large-scale digital/online writing tests. This is likely due to the fact that very few large-scale writing assessments evaluate students on digital, and especially online, writing practices. One notable exception is a special issue of *Assessment in Education* containing four articles examining the validity challenges associated with computer-based PISA such as

“sampling, language, item difficulty and demands, as well as the secondary analyses of students’ reported experiences of formative assessment in the classroom” (Hopfenbeck, 2016, p. 423).

While there is a paucity of peer-reviewed literature on this topic, it is somewhat more common to find grey literature regarding particular digital/online assessments. For example, the American-based NAEP makes publicly available documents that describe how its assessments are designed to measure student writing proficiency. The document *Writing framework for the 2011 National Assessment of Educational Progress* (National Assessment Governing Board; U.S. Department of Education, 2010) defines the construct of writing being measured (construct validity evidence), the design process, the context of the assessment (particularly how the new computer-based assessment compares to previous pen-and-paper versions), as well as scoring and reporting procedures. Other reports compare and contrast NAEP assessments to CCSS based curricula and assessments (Bohrnstedt & Stancavage, 2015) and in doing so presents content validity and (the potential for studies examining) concurrent validity evidence. Similarly, in Ontario, Canada, the Education Quality and Accountability Office (EQAO) makes publicly available the framework used for the design of the computer-based OSSLT. This document describes the construct assessed (construct validity evidence); the alignment between the OSSLT and the curriculum as well as other national and international assessments (content validity evidence); the design process; and scoring procedures (EQAO, 2007). Another EQAO document, *EQAO’s technical report* (EQAO, 2016) presents validity evidence in the form of descriptions regarding “test development, test alignment, test administration, scoring, equating, item analyses, reliability, achievement levels and reporting” (p. 80) and a concluding chapter, which provides an argument-based approach to the validation of EQAO tests.

Literature Review and Conceptual Framework

Three conceptual frameworks were key in framing this study: the Integrated Design and Appraisal Framework (IDAF; Slomp, 2016); bioecological construct of writing (Corrigan & Slomp, 2017); and, online research writing (Corrigan, 2017).

A Framework for Examining Validity Evidence from the ORCA

In this study I used a systematic approach for collecting validity evidence that can be used during the design and appraisal of assessments known as IDAF. Unlike other technocentric approaches, IDAF approaches validation from a more holistic perspective, considering validity, reliability, and fairness in one model. A strength of IDAF is that it highlights the interdependent nature of validity, reliability, and fairness. That is, each of those facets on its own is insufficient for the ethical design and validation of assessment. Evidence from each of those facets considered together, however, presents a more comprehensive and ethical validation argument. Because the definitions of validity, reliability, and fairness differ, I pause here to define these terms. I borrow these definitions from the most recent *Standards for Educational and Psychological Testing* (a.k.a., the *Standards*). When discussing validity, I am referring to “the degree to which evidence and theory support the interpretations of test scores for proposed uses of tests (AERA et al., 2014, p. 11). By reliability, I mean “the consistency of scores across replications of a testing procedure” (AERA et al., 2014, p. 33). According to the *Standards*, there is no single technical meaning for the term fairness (AERA et al., 2014). However, recent conceptualizations of fairness refer to both common standards and equality of opportunity (Nagy, 2000); I will discuss both.

Central to the IDAF process is the identification and appraisal of inferences, building on the work of Kane (2006, 2013), in order to understand how assessment data should be used an

interpreted. While IDAF can be used both in the design and appraisal of an assessment, I will be using it to appraise the ORCA. The description below focuses on that aspect. For a detailed description of IDAF, please see Slomp (2016). This model has been receiving positive feedback: “Slomp et al. (2014) provide a framework that should be carefully considered as a prototype for conducting consequential validity studies” (Behizadeh & Engelhard, 2015, p. 47). The IDAF involves six iterative phases that should be considered during the appraisal of an assessment:

1. *Define assessment aims, context, and principles of fairness:*
 - a. Aims: What inferences are stated or implied by the test designers that will inform how test scores will be used and interpreted?
 - b. Context: Who developed the assessment and for what purposes (e.g., To improve student learning? To assess a writing program? As an accountability measure?)? Who is the audience for the test score data? Who are the stakeholders impacted by this data?
 - c. Principles of fairness: How will the use of this assessment promote equality of opportunity?
2. *Identify elements foundational in the assessment design:*
 - a. What construct(s) is/are being represented by the assessment? How well is the construct understood? Is the construct stable across social, racial, and cultural contexts?
 - b. What is the target domain⁶ defined by the test designers? Does the target domain demonstrate content validity?

⁶ While the target domain is the full range of possible observations that could be measured with relation to the construct, the universe of generalization is a subset of representative observations that the test designers claim to measure. Meanwhile, the sample of observations refers to the

- c. How well does the assessment map onto the intended construct?
3. *Appraise the assessment program:* What method was used to collect the assessment data? What were the affordances and limitations of this approach? Did the method used capture a representative sample of observations from the construct? Does either the method used or sample of observations provided avoid construct underrepresentation and construct-irrelevant variance?
 4. *Appraise the scoring system:*
 - a. What are the scoring criteria? How do the scoring criteria capture the construct, universe of generalization, and sample of observations? How do the scoring criteria capture the domains that comprise the construct? How are domain scores combined into a total score? Does the total score over- or under-represent aspects of the construct?
 - b. What scoring procedures are used? What are the affordances and limitations of these scoring procedures? How will the scoring procedures achieve reliability? Is validity sacrificed in doing so?
 5. *Appraise the assessment results:*
 - a. Scoring Inference. What evidence supports the scoring inferences (e.g., interrater reliability scores, quality control data, statistical models)? What evidence undermines the scoring inferences (e.g., scoring rubrics that fail to capture the

cognitive/metacognitive processes that the assessment enables the test takers to demonstrate (Kane, 2006).

construct)? What claims about performances on the test items does the assessment data enable?

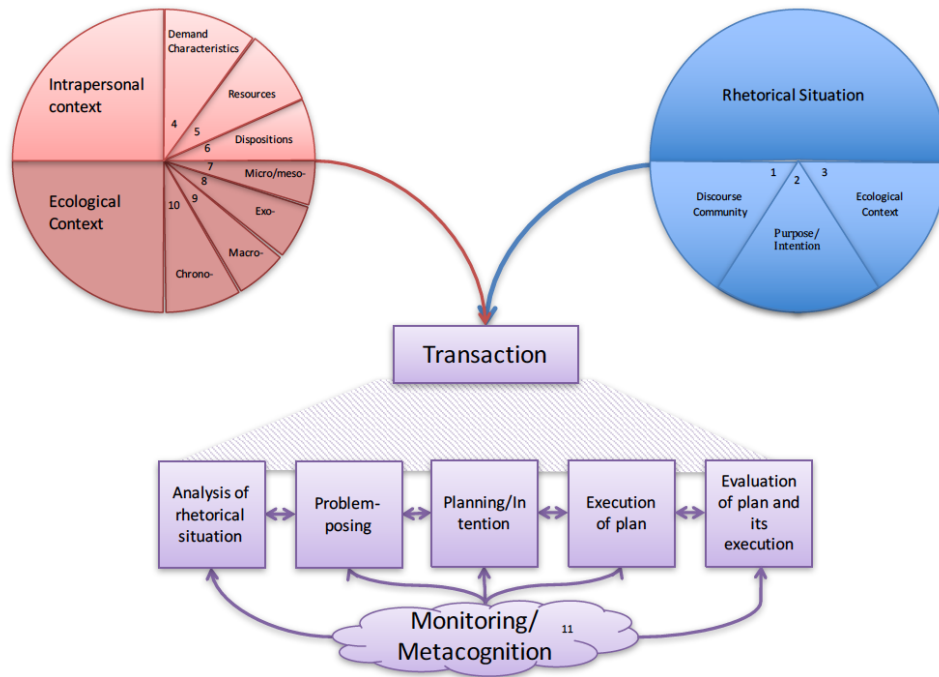
- b. Generalization Inference. What claims about the construct does the assessment data enable? Has a serious effort been made to draw a representative sample from the universe of generalization? Have serious attempts been made to mitigate systematic and random error? Is the ceteris paribus assumption undermined?
 - c. Extrapolation inference. What claims about performance beyond the specific test context, construct, or set of test tasks does the assessment data enable? Is the universe of generalization representative of the target domain? Are the processes involved in responding to the test tasks similar to those involved in other tasks in the target domain?
 - d. What arguments support each of the inferences stated above?
 - e. How have different groups been differentially impacted by the assessment? Are differences in performance among different subgroups due to construct relevant or irrelevant variance?
6. *Appraise the assessment consequences:*
- a. Taken collectively, has the assessment served the purpose and audience for whom it was designed?
 - b. What are the intended (positive and negative) consequences? What are the unintended consequences? Do these consequences vary by population?

A Bioecological Model of Writing

In the previous section, I described the six iterative IDAF phases that I used to appraise the ORCA. Central to IDAF is the process of describing the construct and how it was measured.

With a focus on the writing, I examined the ORCA using the bioecological model (Corrigan & Slomp, 2017). A bioecological model differs from other writing models in that it considers the profound ways in which the Internet has influenced the practice of writing. Further, a bioecological model is useful for analyzing construct validity evidence in that it considers writing through sociocultural, cognitive, metacognitive, and a range of disparate lenses. This approach heeds the call from those within the writing community (Abbott & Berninger, 1993; Graham et al., 2013; Hacker et al., 2009) who have advocated for a more integrative approach to researching writing in all its complexity, rather than remaining siloed by research or paradigmatic tradition. Accordingly, the bioecological is so named because it considers, among other things, how writers' biopsychological characteristics and ecological (i.e., sociocultural) context affect writing and writing development (see the red circle in Figure 5).

Figure 5-6. The bioecological model of writing.



(Source: Corrigan & Slomp, 2017)

I will briefly explain the bioecological model here, although for a more thorough description, please read Corrigan and Slomp (2017). In the figure above, the red circle represents those factors that influence a writer's development. The red circle is broken into two categories: intrapersonal and ecological context. A writer's intrapersonal context are the factors that influence a writer from within, specifically biopsychological characteristics such as personality, motivation, long-term memory, self-efficacy, skills, and experience. A writer's ecological context, conversely, refers to those factors that influence a writer from without. These influences can be proximate to a writer such as micro (e.g., family) and mesosystems (e.g., school/work) or more distant though still very influential, macrosystems (e.g., broadband internet access, funding for the arts, standardized testing policy). While the red circle represents the

factors that directly influence the writer, the blue circle represents those factors that influence the writing context—and the writer too, if she has an awareness of her rhetorical situation.

The purple rectangle below the circles represents the writer's transaction—by which I mean the cognitive and metacognitive processes wherein writers construct meaning through their writing. Undergirding the transaction are the metacognitive processes that occur during writing. Thus, the bioecological model suggests that the success of the transaction is presupposed by the writer's analysis of the rhetorical situation, problem posing, planning, execution of the plan, and an evaluation of the plan and its execution. These processes are not linear, but rather occur recursively throughout the writing process. Also, it is worth mentioning that the processes that are engaged are dependent on the rhetorical situation. For example, writing a grocery may elicit fewer processes and for a shorter period of time than composing a professional blog post. Collectively, this construct provides a description of how a writer's context and rhetorical situation influence a writer and his writing. It also recognizes the important role that metacognition plays in moving from novice to expert proficiency. While expert writers possess greater metacognitive awareness of how their context (intrapersonal and ecological) and rhetorical situation shape their writing, novices tend to lack this awareness (Corrigan, 2017).

Key to IDAF is the process of mapping the assessment onto the construct of interest. A bioecological model is an ideal construct on which to map large-scale literacy assessments such as the ORCA because it considers how factors from both within (intrapersonal) and without (ecological) the writer affect student performance. For example, many large-scale assessments collect data beyond academic performance such as data around socioeconomic status, English language learner status, gender, motivation, self-efficacy, or access to computers and the Internet. Using IDAF also puts a greater emphasis on understanding students' metacognitive

processes. Without this data, it is challenging if not impossible to understand where the breakdown in learning occurred for a given student. Solely examining the final writing product (say, an essay) on a literacy assessment gives stakeholders very little information about how to improve student learning. By looking at an essay response and its associated score, for example, we might be able to tell that a student did not make effective arguments nor include sufficient supporting evidence. But, is that because the student struggled with reading comprehension on the accompanying stimulus material? Or because the stimulus material was biased towards students with knowledge of Canadian symbols (e.g., a text about hockey)? Or does not understand the genre of argumentative writing? Or struggled to generate and organize ideas? Or had insufficient background knowledge about the writing topic? Or got distracted? Or ran out of time? Data from final written products are simply too reductionist to provide this information. All these data enable are comparisons (e.g., among students or writing programs) for accountability purposes or gatekeeping (e.g., a cut-off grade for admission). When I appraise an assessment using a bioecological model, I am compelled to consider data beyond the final written product because this model is predicated on the importance of metacognitive processes that undergird the transaction. Further, if one of the key purposes in assessment is to improve student learning—and it should always be—then IDAF provides a systematic method for appraising assessments by considering the assessment against a range of robust and comprehensive methods that might more fully account for the construct.

Online Research Writing

Of course it is impossible for any one assessment to account for the range of varied and complex practices that occur in writing. In this study I use the bioecological model to refer to the construct of writing broadly while I use online research writing (ORW; Corrigan, 2017) to

refer to the specific practice of writing found in the ORCA, and indeed across many work, home, and school contexts. While the practices of ORW are quite diverse, what those practices have in common is that they are socially and ecologically situated problem-based inquiries supported by a range of cognitive processes (i.e., locating, evaluating, synthesizing, and communicating online information) and metacognitive processes (i.e., analyzing the rhetorical situation, posing the problem, planning, executing the plan, and evaluating the plan and its execution). Further, ORW spans the use of a range of information sources (e.g., databases, search engines, Twitter); a range of genres and platforms (e.g., websites, Tweets, blogs, online academic journals); and takes place in a range of contexts (e.g., academic, professional, political).

The Current Study

The current study is the final phase of a three-phase validation study of the ORCA, centred on this overarching research question: How should test score data from the ORCA be used and interpreted? The research questions for each of the phases were as follows:

1. Based on a systematic review of the literature, what is the construct underlying the ORCA?
2. Based on an empirical investigation of response processes, what is the construct underlying the ORCA?
3. Using the IDAF, how should ORCA test scores be used and interpreted, considering the results of the preceding literature review and empirical study?

Phase one of the study involved a systematic literature review of writing research over the last 50 years and resulted in the articulation of an integrated writing construct (Corrigan & Slomp, 2017). This construct is timely in that it acknowledges the profound ways in which the Internet has revolutionized the construct of writing. Phase two explored differences in response

processes between novices and experts during ORW, and particularly during the ORCA (Corrigan, 2017). In the third phase of this validation study, I synthesize data collected from the previous two phases; previous validation work done on the ORCA by my colleagues; and new data collected during phase two of this study but presented here for the first time. I did so in order to build a comprehensive validity argument to demonstrate the ways in which ORCA test scores should be used and interpreted.

Method

I will briefly outline the method used to collect and analyze the data newly presented in this study. For a more detailed description, please refer to Corrigan (2017).

Overview

After obtaining ethics approval, I invited novice and expert participants to a university research site to complete an ORCA. Once they had completed the ORCA, I guided participants through a think-aloud protocol known as cued retrospective reporting (CRR, also known as stimulated recall). While there are two major types of think aloud methods—concurrent and retrospective—the latter was used so as not to tax the participants' cognitive load, which is already challenged by the completion of the ORCA. CRR is a type of retrospective think aloud wherein “participants are instructed to report retrospectively on the basis of a record of observations” (van Gog, Paas, van Merriënboer, & Witte, 2005, p. 238). Participants' memories were augmented by playing video screen capture of their performance on the ORCA. Because retrospective think alouds are prone to fabrication as they are based on memory, CRR is thought to help mitigate this shortcoming (van Gog et al., 2005). During CRR, participants watched a screencast of their performance while being prompted to describe the processes they used to complete tasks found on the ORCA. Following this, I asked participants to complete a Venn

diagram, comparing and contrasting their ORW practices between the ORCA, and those used at school or work. I opened the interviews with a discussion of these Venn diagrams. Finally, I interviewed participants about their ORW backgrounds and asked them to share and discuss an artifact from work or school that demonstrated their ability to conduct ORW. This was important because it helped me to identify issues of construct underrepresentation prevalent in the ORCA by highlighting practices used for ORW beyond the test itself.

Stimulus

The ORCA is an assessment designed to measure the online research and writing abilities of seventh graders (Leu, Kulikowich, Sedransk, & Coiro, 2009). For the purposes of this study, the ORCA was used to observe participants' ORW practices in a controlled environment. For this study, I used the ORCA-Virtual "Do cosmetic contact lenses harm your eyes?"

Participants

Overall, there were $n = 20$ participants, including $n = 10$ novices and $n = 10$ experts. An overview of the demographic characteristics of participants is provided Appendix I. Participants in Grades 6 to 9 were considered novices, as the ORCA was designed for seventh graders and thus this age range was considered appropriate. To be considered an expert, participants were required to be employed full-time in the knowledge economy (i.e., service/information producing industries). Knowledge workers were chosen for the expert sample because their jobs entail working primarily with information by using it or developing it to solve problems, typically in online environments (Davenport, 2005; Reinhardt et al., 2011). Participants were screened with a preliminary questionnaire in order to ensure that working with and critically evaluating online information was a routine part of their employment. Fluency in English and not having a major visual impairment were additional inclusion criteria for all participants. Replication logic was

used to establish the sample size in which, according to Yin (2009), 6-10 cases (per group) are required to make analytic generalizations.

Analysis

I first transcribed and then uploaded data from CRR and interviews into [Dedoose](#) for coding. First cycle coding involved process coding, or the use of gerunds (“-ing” words), “to connote observable and conceptual action in the data (Miles, Huberman, & Saldana, 2014, p. 75). Process coding aligned with my research objectives of identifying response processes (i.e., cognitive and metacognitive processes), which are indicative of the construct being represented on the ORCA and beyond in ORW in general. During second cycle coding, I used pattern coding (Miles et al., 2014) to cluster codes from the first cycle into the response process categories of Locate, Evaluate, Synthesize, and Communicate. Next, I compared and contrasted the sample of assessment observations permitted by the ORCA with the target domain (i.e., the full range of possible observations) of ORW that I identified during Phase 2 of this study. Finally, I used IDAF to appraise the ORCA in terms of its aims, foundational elements, assessment program, scoring system, assessment results, and assessment consequences.

Results

In the following section, I used the six-stage IDAF to appraise the ORCA in order to determine how data from ORCA test scores should be used and interpreted, and also how the assessment might be redesigned to improve the validity of inferences made from these data.

Defining Assessment Context, Aims, and Principles of Fairness

Aims.

In phase 1 of the IDAF appraisal, a central goal is to identify the inferences provided or implied by test designers. These are the inferences upon which stakeholders might make

decisions regarding the use and interpretation of assessment data. While there are no stated inferences from the ORCA documentation, the following inferences are implied:

- Scoring inference: The scoring procedures for the ORCA are appropriate and they are administered correctly, reliably, and free from overt bias.
- Generalization inference: The ORCA and its associated inquiry-based tasks, assessment prompts, and scoring rubric capture the construct of online research and communication in the context of seventh grade science specifically and seventh grade literacy more broadly.
- Extrapolation inference: The ORCA accurately represents the range of literacy practices that students might encounter in their personal, academic, and eventual work lives in a global information age.
- Decision inference: Stakeholders can use ORCA data to assess accurately students' ability to conduct and communicate online research so that stakeholders can determine next steps for learning, research, professional development, and curriculum development.

In later phases of IDAF, I will appraise these inferences in terms of their appropriateness

Context.

The ORCA was developed by a [team of researchers](#) associated with the New Literacies Research Lab with expertise in New Literacies, reading comprehension, statistical analysis, and psychometrics. The genesis for the project began with research that demonstrated that additional cognitive and metacognitive practices are required to participate in online literacies beyond those required in traditional, print-based literacy (Coiro & Dobler, 2007; Leu et al., 2004). The team's goal was to design a large-scale assessment that could measure the complex construct of literacies (i.e., New Literacies) that are "increasingly important for any nation determined to lead a global information economy" (ORCA Project Overview, n.d.).

The ORCA was created to be a reliable, valid, and practical assessment of online research and communication ability among seventh-grade students, particularly in the subject of science (Leu et al., 2012). More specifically, the goals of the ORCA project were described as follows:

1. [to] develop three different types of valid and reliable assessments to measure online reading comprehension;
2. [to] evaluate the internal assessment characteristics for each instrument type to inform decisions about which is most useful and practical for schools; and,
3. [to] estimate both the utility and practicality of each instrument in the eyes of key decision makers: teachers, school administrators, and Chief State School Officers.

(ORCA Project Overview, n.d.)

Interestingly, the ORCA was originally named the Online *Reading* Comprehension Assessment, but the test designers soon recognized the important role that writing played in the assessment and changed the name to its present version, the Online *Research* Comprehension Assessment. Further, fully 25% of the score accounted for in the ORCA scoring rubric is directly related to writing (i.e., Communicate; refer to the rubric on Table??). Thus, using a bioecological model of writing (Corrigan, 2017) to examine the ORCA will help identify aspects of the construct that are both represented and/or ignored or underrepresented.

Currently, the ORCA is not a high-stakes test. Rather, it is unique in that it is a large-scale formative assessment. Data from the ORCA has been used for the following: [to inform research](#); to inform the creation of [professional development resources](#); and, to provide feedback to teachers and other stakeholders regarding how students are performing with regards to online research and communication. Teachers and administrators from participating schools are provided with a report (Appendix A) with regards to how their students performed within a quick

(approximately two week) turn-around period. There are no reports on individual student performance. The class reports include the overall average score, average score by LESC domain, as well as the scores across each of the 16 individual score points. Furthermore, the report also identifies the four strongest and weakest skills.

Principles of Fairness

In addition to measuring students' ability to conduct and communicate online research, the ORCA also elicits data regarding students' intrapersonal and ecological context. Prior to completing the ORCA, the following data were collected:

- school variables (e.g., name of state, city, and school; percentage of students in each district who qualify for free or reduced price lunch; ratio of students to a computer; percentage of computers in the school with Internet access);
- student variables (e.g., name of student; teacher, and school; gender; language spoken at home; standardized reading and writing scores from state tests; Individual Education Plan; prior knowledge particular to version of the ORCA that the students took [i.e., knowledge about the eye and vision for the ORCA version on cosmetic contact lenses]); and,
- student and teacher internet use variables (e.g., number of computers at home/school; access to Internet at home/school; time spent on the Internet for particular subjects, doing particular practices, and using particular technologies; Internet use at home versus school; and, self-efficacy scores for a variety of new literacies practices).

These data, in combination with student scores on the ORCA, provide a more holistic interpretation of student performance. Importantly, they can help to identify subgroups who perform poorly on the ORCA in order to identify possible achievement gaps and provide

additional resources. This acts as a mechanism for providing fairness through equality of opportunity.

Identifying Elements Foundational in the Assessment Design

As large-scale assessments are designed with the purpose of measuring a particular construct, defining the construct is an important phase in IDAF. Therefore, in this section, I will describe the construct intended to be measured by the ORCA focussing particularly on the writing domain of the construct. Then, I will describe whether and how construct validity evidence supports this. Finally, I will conclude by presenting content validity evidence, with specific attention to construct underrepresentation and construct-irrelevant variance.

The test designers articulate the construct being measured by the ORCA in their [framework document](#). In the document, they articulate the new literacies of online research and comprehension (J. L. Coiro & Dobler, 2007; Leu, Forzani, et al., 2013) as a problem-based inquiry involving the processes of locating, critically evaluating, synthesizing, and communication online information. They stress that these processes combine new online and traditional offline literacies in complex, interrelated ways.

I analyzed the ORCA using a bioecological model (Corrigan & Slomp, 2017) in order to focus on the affordances and limitations of the assessment, particularly the writing domain of the construct. This allowed me to consider both how the writer's context (i.e., intrapersonal and ecological) and the rhetorical situation influenced performance on the ORCA. Further, it allowed me to consider both the cognitive writing practices, as well as the metacognitive practices undergirding writing.

I will begin with an analysis of the rhetorical situation elicited by the ORCA. Arguably, one of the ground-breaking features of the ORCA is that it is designed to measure competencies

in online—not just computerized—literacy practices. While other large-scale assessments measure literacy competencies in computerized contexts (e.g., NAEP and PISA), the ORCA is unique in that it measures those in online contexts (Leu et al., 2014). Due to the dynamic nature of online environments and the need for stable assessment contexts in large-scale assessment, measuring an online literacy construct is challenging. As previously mentioned, the designers of the ORCA responded to this challenge by creating a stable assessment context through which to measure online literacies in the form of a closed Internet environment. This closed Internet environment is so authentic that the majority of participants in this study did not realize that the ORCA was not taking place in an open Internet environment. For example, Natasha, one of the novice participants, noted that she strove to write an accurate wiki entry because "they're [those who will read the wiki] actually going to use this information and maybe it will be wrong." Monica, one of the expert participants, had this to say about the authentic feel of the ORCA: "I did find it very normal and familiar to write in the chat windows that resemble a Facebook messenger environment." What makes the ORCA feel so authentic are features such as a search engine, social networking platform, email, chat features, websites, and editable wikis. The ORCA is unique among large-scale assessments in that the sample of observations taken from the universe of generalization⁷ includes literacy practices typically ignored on other assessments. To name a few, these include identifying key words to input into a search engine, locating a text from a search engine results page, the ability to click on hyperlinks to discover more

⁷ While the target domain is the full range of possible observations that could be measured with relation to the construct, the universe of generalization is a subset of representative observations that the test designers claim to measure.

information, and communicating via varied networked and digital communications technologies (i.e., instant messaging, social networking, email, and wiki).

In the bioecological model of writing, both the rhetorical situation and the writer's (intrapersonal and ecological) context shape the transaction, whether the writer is conscious of it or not (Corrigan & Slomp, 2017). As previously mentioned, data from a number of variables concerning the writer's intrapersonal and ecological contexts (school, teacher, and student variables) are collected alongside the ORCA. This helps to further situate writers in their intrapersonal and ecological contexts, providing a more complete picture of a student's performance. Further, these data can be used to analyze whether or not the construct is stable across social, racial, and cultural contexts. I will elaborate on this later.

The transaction described within the bioecological model refers to the processes wherein writers construct meaning through their writing. The transaction is made manifest through cognitive processes and (ideally) supported by metacognitive processes. In phase 2 of this study, I analyzed the response processes (i.e., metacognitive and cognitive processes) used by participants completing the ORCA and during ORW more generally. The results of that study showed that participants relied on varied and at times complex processes to locate, evaluate, synthesize, and communicate online information; furthermore, these processes were undergirded by a series of complex metacognitive processes (Corrigan, 2017).

Next, I will juxtapose the response processes elicited by participants taking the ORCA (i.e., the sample of observations) with those that participants use in ORW more generally (i.e., target domain—the full range of possible observations associated with the construct). I will go through each domain of the construct (i.e., Locate, Evaluate, Synthesize, and Communicate) describing which aspects of each domain are represented, underrepresented, or ignored on the

ORCA. In other words, this next section will focus on content validity evidence with particular attention to construct underrepresentation. I will discuss this evidence both in this section and in Scoring System analysis phase. Doing so will help inform how data from the ORCA should be used and interpreted.

Locate.

Students taking the ORCA are assessed on the following four Locate items: locating the correct email message in an inbox or the correct section of a wiki (L1); using appropriate key words in a search engine (L2); locating the correct site in a set of search engine results (L3); and, correctly identifying and sharing the website addresses from two different search tasks (L4). The question is, how do these items map onto the target domain? In other words, do these items constitute a representative sample of observations from the target domain? In

Table 5-6 below, the left side of the table identifies the target domain that emerged from case summaries of a cross-case analysis of response processes in ORW (see Corrigan, 2017). These are juxtaposed with the sample of observations permitted by the ORCA, as analyzed in the current study.

Table 5-6 Locating Online Information: Target Domain vs. Sample of Observations

	Target Domain	Sample of Observations from the ORCA
<i>Primary Locating: Locating Sources of Information</i>	Using conventional search strategies including <ul style="list-style-type: none"> • key word searches • title searches • narrowing search by date, location, and name of publication and • reading the resume in the search engine results 	The ORCA elicited all of the conventional search strategies listed on the left. These processes were captured in the scoring rubric under the Locate domain: <ul style="list-style-type: none"> • Locating the correct email message in an inbox or the correct section of a wiki (L1) • Using appropriate key words in a search engine (L2) • Locating the correct site in a set of search engine results (L3)

		<ul style="list-style-type: none"> Identifying and sharing the website addresses from two different search tasks (L4)
	<p>Using more sophisticated search strategies including using</p> <ul style="list-style-type: none"> Boolean operators using divergent sources of information (e.g., Twitter feed; reference managers; listserv; professional/expert blog posts) backward and forward reference searching second-level backward reference searching backward author searching 	Not assessed on the ORCA.
<i>Secondary Locating: Locating Information within the Source</i>	<p>Skimming texts in order to efficiently identify relevant information</p> <p>Identifying relevant information</p>	<p>Participants engaged in secondary locating, to varying degrees, in order to complete the ORCA tasks. These processes were indirectly captured in the Synthesize domain of the scoring rubric (S1 – S4).</p> <p>Same as above.</p>
<i>Metacognitive Practices: Using Metacognitive Practices to Locate Information</i>	<p>Engaging in a broad range of metacognitively-supported search strategies including</p> <ul style="list-style-type: none"> triaging using technology to manage information remixing information for new purposes locating primary source information (being sceptical of secondary sources) locating information with an eye to the audience for whom one is writing 	Not explicitly assessed or unable to assess on the ORCA.

As the table above illustrates, the ORCA solicited mainly primary locating practices, specifically, conventional ones. In terms of construct underrepresentation, more sophisticated

primary locating practices (e.g., using Boolean operators) and metacognitive locating practices were either ignored or underrepresented. However, given that the ORCA was designed for Grade 7 students, this seems reasonable. Were the ORCA expanded to include secondary-level students, an expanded sample of observations should be considered that would elicit a deeper, broader, and more complex array of locating practices.

In terms of primary locating practices, one feature of the ORCA that some participants remarked about was that in the ORA-virtual, the participants were told which websites to search for (i.e., students are given the article's title and its publication) instead of finding their own sources of information. "I would have preferred if I could have done my own search and found my own way," said Tina, one of the expert participants. Tina also remarked that, in her job as a technical writer and business analyst, she would rarely be given information to locate. Rather, she would need to identify and locate that information on her own. In the video games and heart healthy snacks ORCA tasks (i.e., the unrestricted versions), however, students are given the latitude to search for a website of their choosing. In the validation year of the ORCA, both restricted and unrestricted versions of the ORCA were designed in order to test for differences in reliability and usability. When the search task was unrestricted, it posed greater challenges for scoring and reliability.

Evaluate.

Concerning critically evaluating online information, students are assessed on the ORCA on their ability to identify the author of a website (E1); evaluate the author's level of expertise (E2); identify the author's point-of-view (E3); and, evaluate the reliability of a website (E4). In Table 5-7 below, I describe which critical evaluation processes that describe ORW more generally, and which can be found on the ORCA. As I demonstrate, the ORCA primarily solicits

an evaluation of the source credibility, and only implicitly involves evaluating the argument credibility. This is an important distinction because even a seemingly credible source can make arguments that lack credibility from time to time. Thus, when conducting ORW, it is incumbent upon the writer to evaluate both the source (i.e., author and publication) as well as the argument. The lack of critical evaluation of the argument posed an issue of construct underrepresentation. Another construct underrepresentation issue was created by the closed Internet environment that prevented participants from conducting more sophisticated critical evaluation processes (e.g., triangulating evidence from a variety of sources; background checking; and, backward and second-level backward reference searching). Considering that the ORCA is an assessment designed for seventh graders, once again, this level of critical evaluation seems appropriate—with the exception that evaluating argument credibility should be explicitly assessed. More research will be required to ascertain whether this is an issue or not as research regarding developmental levels for ORW is still in its nascence.

Table 5-7 Evaluating Online Information: Target Domain vs. Sample of Observations

	Target Domain	Sample of Observations from the ORCA
<i>Primary Evaluation: Evaluating Argument Credibility</i>	Evaluating arguments (including their claims, evidence, and warrants)	Indirectly assessed on the ORCA via Evaluate 4: The student evaluates the reliability of the website (or lack thereof) AND provides one logical and accurate explanation to support their conclusion about either: author expertise, publisher trustworthiness, or <i>research findings</i> .
	Questioning the degree of the source of the information (e.g., primary vs. secondary source)	Not explicitly assessed on the ORCA.

	Viewing information credibility as a matter of degree rather than dichotomously.	Not explicitly assessed on the ORCA, although some participants communicated this in their wiki entries.
<i>Secondary Evaluation: Evaluating Source Credibility</i>	Applying scepticism toward sources (author, organization, and/or publisher) such as <ul style="list-style-type: none"> questioning the author's education, expertise, and purpose for writing background checking the author and/or organization 	Participants were prompted to identify the author of a website (E1); evaluate the author's level of expertise (E2); identify the author's point-of-view (E3); and, evaluate the reliability of a website (E4). Due to the closed Internet environment, participants were unable to conduct background checking.
	Using sophisticated strategies to verify source credibility such as following hyperlinks both within the website (clicking on the author's name to learn more) and beyond it (background checking via sources not hyperlinked to the text being verified)	Participants were able to verify source credibility from within the website although not from beyond it due to the closed Internet environment.
<i>Metacognitive Practices: Using Metacognitive Practices to Evaluate Information</i>	Using sophisticated and diverse metacognitive strategies to evaluate online information (e.g., triangulating evidence; considering multiple, diverse perspectives; triaging; backward and second-level backward reference searching)	Participants were mostly unable (aside from considering multiple, diverse perspectives, to some degree) to use metacognitive critical evaluation processes due to the closed Internet environment.

During their interview following the ORCA, participants noted ways in which the ORCA could be improved to make the assessment more authentic in terms of the Evaluate domain. Monica, a professor and one of the expert participants in the study, noted that she would have “cited primary source evidence.” She explained that she would have liked to have “actually gone to the source of the information rather than reporting second hand on what these articles told me.

Part of my professional practice would have been to do that. I never would have shared second hand information without having evaluated that information myself, so that was a limitation.”

The reason that Monica and other participants could not go back to the primary source was because, as explained earlier, the ORCA uses a closed Internet environment

Similarly, Laura, a lawyer who participated in the study, noted how she wished she could have inserted a hyperlink to the Food and Drug Administration (FDA) laws to which the article was referring: “For me it was challenging when I was editing the wiki because I wanted to be citing the law. I wanted to add a reference here to law or regulation, not even quoting from the FDA saying it was illegal, but actually getting to the law or the regulation.” She noted that the websites she was instructed to read during the ORCA regarding the selling of cosmetic contact lenses were but authors’ interpretations of the FDA law (i.e., secondary sources), not the law itself (i.e., the primary source). While Laura’s strategy was obviously framed by her legal training, it is also informed by more basic critical evaluation practices such as those noted in the CCSS at the seventh-grade level, specifically the Reading Standard 8 for Literacy in Science: “Distinguish fact, reasoned judgment, and speculation.” Essentially what Laura is doing is employing scepticism towards a secondary source. Within the ORCA, she said she would have liked to have gone back to the primary source to determine whether the information is fact or otherwise. This is a strategy known as backward reference searching. In interviews about their school and professional ORW practices, participants particularly at the expert end of the novice-expert spectrum cited this as a strategy.

Catherine, an expert participant who has given digital literacies workshops to middle and secondary students noted how she often challenges students to be sceptical of anything they read online. She specifically recommends to students that they go back to the original source of the

data (i.e., backward reference searching), that they see if the information is consistent across a number of websites (i.e., triangulating the data), and that they do background checking on the author (i.e., Are authors really who they say they are? Did they really work there or graduate from that school, for example?). Because the ORCA used a closed Internet environment, this breadth of critical evaluation is not fully possible. However, participants did note that within the ORCA they were able to look across the four websites they were researching to triangulate data.

Synthesize

The Synthesize items gradually increase in difficulty from summarizing ideas from one website (S1), to two websites (S2 and S3), and finally to synthesizing information from all four websites and taking a position (S4). In order to achieve the score point for S4, students are required to use their “own words (at least 3) to provide an argument (that takes a position) AND two pieces of evidence, one detail from each of two different websites.” While this is quite basic synthesis, it is a progression from summarizing. The ORCA challenges students at a rudimentary level to both summarize and synthesize online information (see Table 5-8). As participants from the novice-expert study revealed in phase 2, this process of synthesizing becomes much more complex in real-world scenarios where writers are challenged with synthesizing information from copious amounts of online information from multiple competing perspectives (Corrigan, 2017), not just four pre-determined websites with only two perspectives. However, I would argue that synthesizing information from four websites is an appropriate challenge for the grade level for which the ORCA was designed. Should the ORCA be redesigned for secondary school students, however, synthesizing and weighing evidence from multiple, competing viewpoints should be incorporated.

Table 5-8 Synthesizing Online Information: Target Domain vs. Sample of Observations

Target Domain	Sample of Observations from the ORCA	
<i>Summarizing</i>	Summarizing involves <ul style="list-style-type: none"> • Paraphrasing information using original language • Identifying key ideas 	Summarizing is assessed on the ORCA by the following items: <ul style="list-style-type: none"> • summarizing an important element from one website (S1); • summarizing important elements from two websites (S2); • summarizing important elements from a second set of two websites (S3); and, • summarizing important elements from the websites in the research task to develop an argument (S4).
<i>Synthesizing</i>	Synthesizing involves summarizing plus <ul style="list-style-type: none"> • Weighing credibility of evidence by comparing and contrasting information from a variety of sources, some of which may have contradictory viewpoints • Considering multiple perspectives and the affordances and limitations of each 	This is assessed via S4 via the following prompt: <ul style="list-style-type: none"> • Now, take a position. Do cosmetic contact lenses <u>harm</u> your eyes? Send me a summary of what you think after reading all FOUR sites. Include important details from the websites that explain your thinking. <p>It is assessed again in the Communicate tasks wherein students are prompted to revise Mr. Henry’s class wiki and, in doing so, take a position. The following prompt is used:</p> <ul style="list-style-type: none"> • Do cosmetic contact lenses harm your eyes?
<i>Generative Synthesizing</i>	Used extant information to generate new ideas	Though not explicitly assessed, the open-ended Synthesize and Communicate tasks would enable this.

Communicate.

As mentioned, the ORCA was originally designed as a reading assessment, so perhaps it is not surprising that this is domain of the ORCA that is least developed in terms of its assessment prompts, but particularly in terms of how the domain is scored. The sample of observations enables a fairly robust observation of the writing domain in that the task of composing a wiki entry by taking a position on a controversial issue is quite rich and authentic. However, the scoring criteria, particularly for the Communicate domain, are quite limiting. Students are assessed on four Communicate score points: their ability to make a wiki entry in the correct location (C1); use descriptive voice in an information wiki (C2); include an appropriate heading for a new wiki entry (C3); and, compose and post a well-structured, short report of their research in a wiki (C4). Given the complexity of the writing task, these scoring criteria, for the most part, appear superficial.

Table 5-9 Communicating Online Information: Target Domain vs. Sample of Observations

Target Domain		Sample of Observations from the ORCA
<i>Communicating Critical Discourse Knowledge</i>	<ul style="list-style-type: none"> • Communicating with a sense of promoting equity and social justice; • Writing accessibly • Considering intertextuality and historical perspectives 	Not assessed, although due to the open-ended nature of the wiki task, students could take this into consideration to some extent. For example, they could choose to write in an accessible manner.
<i>Communicating Discourse Knowledge</i>	Communicating with a sense of disciplinary, professional, and/or community norms and values	To a minor extent, this is assessed via the following item: Compose and post a well-structured, short report of their research in a wiki (C4)
<i>Communicating Rhetorical Knowledge</i>	Choosing diction to communicate rhetorical purpose, but also with a sense of discourse community values	To a minor extent, this is assessed via the following item: use descriptive voice in an information wiki (C2)

	Using a range of multimodal rhetorical strategies including using video, using hyperlinks, elements of web design, and aesthetic choices in online writing	Not assessed.
<i>Communicating Genre Knowledge</i>	<ul style="list-style-type: none"> • Choosing genre by considering the rhetorical situation (e.g., deciding to send client an email rather than a formal legal decision) • Using genre conventions as appropriate to the community/profession 	To a minor extent, this is assessed via the following item: use descriptive voice in an information wiki (C2).
<i>Using Writing Process Knowledge</i>	Using a recursive process <ul style="list-style-type: none"> • Using collaborative writing processes necessitated by distributed expertise; • Using complex collaborative writing processes requiring project management 	Not assessed. Simulated asynchronous collaborative writing is assessed in the wiki versions of the ORCA.
<i>Communicating Subject Matter Knowledge</i>	Communicating subject matter knowledge with consideration given to the norms and values of the discourse community	This is assessed to an extent via the following item: <ul style="list-style-type: none"> • Compose and post a well-structured, short report of their research in a wiki (C4)

In terms of the Communicate domain of the ORCA, feedback from participants was mostly positive around the ways in which the assessment mimicked their lived experiences with ORW. However, participants did suggest additional ways in which the ORCA could be designed to be more authentic. For example, four participants noted how it would be useful to be able to insert hyperlinks in their wiki entries on the ORCA. Expert participants including Monica, Laura, and Colin, as well as novice participants Cecilia and Anna got around this by pasting URLs into their wiki entries and putting them in parentheses where they would have inserted a hyperlink instead.

Aside from using hyperlinks to refer to the primary source, participants also described using hyperlinks to acknowledge the sources that they cited. A professor named Monica, indicated on her wiki entry where she would have hyperlinked to the websites she was citing in order to reference them. As opposed to creating a formal reference list, hyperlinking to sources is a common practice in ORW in some genres, such as the wiki. Other participants (Kelsey, Melanie, and Colin) noted in their interviews how inserting hyperlinks is a common part of their professional writing in online environments. These participants noted how hyperlinks are used as part of the aesthetic of online writing so as to not overwhelm their reader with densely written texts. Hyperlinks can be used to refer the reader to additional information.

The tasks on the ORCA demonstrate moderate alignment with the bioecological model of writing as they challenged participants to adapt to online communication technologies and their associated writing processes, some of which were new to them. For example, editing a wiki page was something that only three of the 20 participants in this study had ever experienced. According to the bioecological model, it is important for writers to be able to engage in near and far transfer as they encounter new technologies, new processes, new genres, and new discourse communities (Corrigan & Slomp, 2017). The ORCA also challenged participants to edit the wiki page in way that would match genre features of the previous wiki entries on Mr. Henry's class wiki (e.g., short, simple sentences; third person point-of-view). It should be noted that this was an implied, not a stated challenge. In other words, participants were only asked to post a response to the research question on the class wiki. However, during CRR, one novice (Fiona) and most expert participants noted that they were considering the genre features of the wiki, specifically how to make their writing blend in with the existing wiki page. This overlaps with analyzing the discourse community and rhetorical features of the wiki as participants in the study

spoke about their consideration of audience and how they would write to meet the needs of their audience (i.e., Mr. Henry's class) by selecting an appropriate style (e.g., an informal tone, vocabulary appropriate to seventh graders, and relatively simple syntax) and content (e.g., examples of celebrities who had used cosmetic contact lenses, the health risks associated with these lenses, and suggestions for how to use them safely; see Corrigan [2017] for more). Taken together, the response processes of participants to ORA-virtual suggest that it is enabling an authentic writing context. However, where the ORCA falls short are the Communicate scoring criteria, which limit the ways in which writing assessment is measured. I will describe this in further detail in the upcoming sections.

Evaluating the Assessment Program

In this phase of the IDAF process, I examine the methods used on the ORCA for collecting assessment data considering both the affordances and limitations of this approach. In particular, I will examine content validity evidence paying close attention to construct underrepresentation and construct-irrelevant variance.

Content validity evidence and construct underrepresentation.

Obviously, no one assessment can ever fully capture a construct; however, it is important that large-scale assessments “draw a representative sample from the universe of generalization” (Kane, 2006, p. 35). In other words, large-scale assessments are designed to solicit a representative sample of observations from the construct in order to demonstrate content-related evidence of validity (Cureton, 1951). It should also be noted that the ORCA was designed to align with the CCSS, an initiative seeking to establish consistent curricular outcomes across the U.S. The ORCA does indeed align with numerous seventh-grade CCSS strands such as Reading

Standards for Informational Texts, Reading Standards for Literacy in Science, and Writing Standards (see [Common Core Standards and the ORCA Project](#) for more).

While there has been a noble intent to fully and accurately represent the construct of online research and writing during the ORCA, this representation is by no means complete—nor could it ever be. The following is a summary of instances of construct underrepresentation that were noted previously. First, students were given which sources of online information to locate instead of being tasked with locating their own. Second, while students were prompted to critically evaluate the reliability of the author and/or publication, none of the assessment prompts directly asked students to critically evaluate the arguments presented by these authors. The evaluation of arguments is also not explicitly accounted for in the scoring rubric. Third, students were prompted to summarize and synthesize information with less of an emphasis of generative synthesis. Finally, with regards to the communicate domain, there were also some challenges with regards to construct underrepresentation. Composing a wiki entry in order to take a position on an issue such as whether or not cosmetic contact lenses are safe to wear is a robust task. However, the scoring criteria used to assess this task should be revised to reflect a more meaningful and holistic assessment of students' writing performance, as I will discuss shortly.

Construct-irrelevant variance.

An important step in examining the issue of fairness in assessment is to identify whether variance is construct relevant or irrelevant. I find Zieky's (2016) matrix useful in framing this analysis:

Table 5-10 Relationships among Sources of Variance, Fairness, and Validity

	Unfair	Fair
--	--------	------

Valid		Construct relevant variance
Invalid	Construct-irrelevant variance related to group membership	Randomly distributed construct-irrelevant variance

Source: (Zieky, 2016, p. 13)

In terms of whether variance related to group membership is construct-irrelevant (Zieky's lower-left quadrant in Table 5-10) or relevant (upper-right quadrant), there is as of yet no research on this matter. As an initial step, there has been research done on the ORCA that has confirmed variance among groups; however, whether or not that variance is construct relevant remains uncertain. For example, my own earlier research on the ORCA (Corrigan, 2014b) confirmed a difference between males and females. Specifically, a mixed between-within subject analysis of variance was conducted to assess the impact of gender on participants' ORCA and Communicate scores. There was a significant interaction between gender and score, Wilks Lambda = .96, $F(2, 1127) = 43.78$, $p < .0005$, multivariate partial eta squared = .04. There was a substantial main effect for score type, Wilks Lambda = .09, $F(2, 1127) = 11199.97$, $p < .0005$, partial eta squared = .91. The main effect of gender was significant, $F(1, 1127) = 31.89$, $p < .0005$, partial eta squared = .028, or a small effect size (Cohen, 1988). Aside from gender, previous studies have also shown a significant and substantial achievement gap based on socio-economic status (Leu et al., 2015). In order to determine whether or not the variance among groups is construct relevant, further studies investigating response processes would be beneficial (Goodwin & Leech, 2003)

Aside from construct-irrelevant variance based on group membership that I was unable to confirm, I also explored construct-irrelevant variance related to the assessment itself. These sources of construct-irrelevant variance included time, digital literacies knowledge, and the lack of universally designed assessment features.

The value of timed writing assessments such as are found in the ORCA has long been a contentious issue. To recap from earlier, the ORCA challenges students with an online research task that culminates with students communicating their research findings via an email, or in the case of the ORCA version examined in this study, a wiki entry. In its large-scale format, students were given one class period to take the ORCA, which generally meant that participants had about an hour to complete conduct their research and write the wiki entry. Whether administered in the classroom or as standardized examinations, timed writing assessments offer a “rough and imprecise, but most often fairly accurate, prediction of a student’s performance on longer, traditional writing assignments” (Lau, 2013, n.p.). This begs the questions, what issues of construct-irrelevant variance are specifically introduced by timed assessments? For one, timed assessments emphasize the final product and truncate or—in the case of the ORCA—completely ignore the writing process (Slomp et al., 2014). In the ORCA, there is a place to take notes on each article; however, there is no place to outline or draft the wiki or email entry. Indeed, nowhere in the assessment are there any prompts that indicate that, for example, students should brainstorm, outline, draft, or revise their work. Results from phase 2 of this study indicated that while novices followed a more linear writing process with few revisions, experts took a more iterative, process-oriented approach to ORW (Corrigan, 2017). While writing a grocery list, for example, would not necessitate such a process-oriented approach, the complex nature of ORW tasks generally does. If we wish for students to develop greater writing expertise, teachers must emphasize these iterative, process-oriented approaches. It sends a confusing message to students when they are assessed on product-oriented writing during large-scale assessments on one hand, but on the other hand are told that process-oriented writing is valued. If this type of writing is truly important, assessment needs to reflect that.

Some have argued that the timed, impromptu writing that occurs in standardized assessments such as the ORCA is a vast improvement over previous attempts to assess writing via proxies such as multiple choice questions regarding grammar and vocabulary (Lau, 2013). Undoubtedly, they are. But, the question remains about whether or not the writing in these standardized contexts is generalizable to ORW done in authentic contexts (say, even to writing a wiki entry in a professional or classroom context). If timed, impromptu writing assessments constitute their own genre (i.e., the assessment genre), it is questionable as to whether or not the writing scores resulting from these assessments can be generalized to writing more broadly—no matter how similar the tasks may be to those in real-life contexts. The assessment genre “embodies the idea that timed writing is fundamentally similar to the writing involved in extended composition” (Lau, 2013). As participants such as Monica, Tina, and Laura noted during their interviews, writing a wiki entry in their context as professionals is a high-stakes, protracted process because the writing is for public consumption. They noted that—unlike in the context of the ORCA—before publishing such a wiki entry, they would subject their writing to numerous iterations, revisions, and checks for accuracy. Monica noted that in her role as a professor of educational technology, she is professionally and ethically obligated to ensure that whatever she writes for public consumption is accurate or it might act to misinform practice and policy. While ORW in real-life contexts does indeed include time restrictions, the ones imposed on the ORCA are quite restrictive. While the time restriction for writing a wiki entry might be a couple of days in real life, for example, the one hour or so time limit on the ORCA seems unnecessarily restrictive. Thus, as ORW practices in real-life contexts appear to be fundamentally different from those used on the ORCA where time constraints are overly restrictive, I would argue for caution when interpreting the results of the ORCA.

When asked how the ORCA could be improved, 17 of the 20 participants commented that they felt they did not have enough time and/or felt rushed. While I informed participants in the protocol before they took the ORCA that they could have as long as they needed, there were reminders embedded within the ORCA that indicated to participants how much time they had left. For example, the avatar Briana would remind participants that they had x number of minutes left during various segments of the ORCA via an instant message. Jill, an eighth grader who took the ORCA, commented that it was “kind of annoying how the time kept popping up. It’s really rushed.” Even the expert participants commented on this. “The time crunch is the hardest thing. It’s somewhat realistic because at work there’s never a lot of time, but at work I wouldn’t send an email that was half done or, you know, put half the information and say I was done,” explained Judy.

In addition to time constraints, there were interesting challenges that arose during the ORCA with regards to students’ knowledge of digital literacies. For example, when one of the novice participants named Sue got to the part where she was prompted to write her wiki entry, she could not figure out that she needed to click on the pencil icon to commence the editing. In my study, once it was clear that Sue could progress no longer, I made a note of this, but proceeded to direct her to click on the pencil icon so that she could resume. Otherwise, Sue would have automatically received a score of 0 on all four Communicate items. But, does the fact that she did not know how to click on the pencil icon to commence the editing mean that Sue does not know how to perform any of the four Communicate tasks? As this study demonstrated, the answer is no. Yet, knowing how to click on the edit button is a simple, yet vital part of writing in online contexts such as these where literacy and digitality are intertwined (Stroupe, 2000). When I assisted in the large-scale administration of the ORCA, however, the protocol

directed that test administrators were not permitted to provide such assistance for reasons of reliability. In fact, I recall an instance where a student asked for such assistance and I had to say, “I’m sorry. I’m not permitted to help you with this.” It seems likely that the ORCA would underestimate the writing ability of some students with weaker digital literacies. What might such an underestimation of a student’s writing score do to a student’s self-esteem or to the teacher’s perception of the student’s writing ability? These questions are cause for concern on the ORCA and indeed on any standardized assessment. In the case of the ORCA, this concern is mitigated by the fact that no individual results of students’ performance are reported. Rather, teachers receive a report of the overall class performance only.

Just as lack of digital literacies knowledge caused Sue to struggle with the Communicate component of the ORCA, it also caused all but two of the novice participants (i.e., Cecilia and Helen) to struggle with the Evaluate domain. While all of the expert participants knew to click on the hyperlink of the author’s name when asked to evaluate the author’s credibility, only two out of the 10 novices were aware of this strategy. By clicking on this hyperlink, participants were directed to a page describing the author’s educational and professional background. In follow-up interviews, when I advised participants who had not done so during the test to click on the hyperlink, they provided a much better evaluation of the author’s credibility than they had before seeing this additional information. Clearly, assessing students on researching and writing challenges that are so intertwined with digitality poses unique challenges. So, is the ORCA biased towards students with access to Internet technologies? Yes—but maybe that’s the point?

One final issue connected with construct-irrelevant variance and the ORCA relates to an absence of universally designed assessment (UDA; National Center on Educational Outcomes, 2002) features, which are features designed to accommodate the needs of a diversity of learners.

While clear and simple language was used in the development of items on the ORCA, a number of other UDA considerations might be considered in the ORCA's redesign. For example, the ORCA would benefit from UDA features such as text-to-speech and speech-to-text capability; text descriptions of all graphics; adjustable font size; electronic dictionary, thesaurus, and translation tools; graphic organizers for organizing and planning writing; and extremely generous time limits. Although this was not a major issue in my study, novice participant Natasha who goes to a French school and speaks English and French at home did note that there was some vocabulary used on the various websites on the ORCA that she did not understand. In this case, a translation tool embedded into the ORCA would have been beneficial (National Center on Educational Outcomes, 2002). In countries like Canada where one in five speaks a language other than English or French at home (Statistics Canada, 2012), UDA features such as online dictionaries and translators seem like a logical way to promote fairness.

Appraising the Scoring System

The fourth phase of IDAF involves appraising the scoring system. Earlier in Tables 2-5, I described the scoring criteria used for each of the Locate, Evaluate, Synthesize, and Communicate (LESC) domains. As portrayed in Table 5-11, all four versions of the ORCA are assessed using the same multidimensional scoring model for a total ORCA score of 16 score points with 4 points per each LESC domain. Further, each item is scored using dichotomous scoring procedures (i.e., items are either correct or incorrect) and the score points were scaled by subject-matter experts and psychometricians from easiest to most difficult to facilitate item response function.

Table 5-11 ORCA Constructs and Score Points

CATEGORY AND ITEM	ITEM	POINTS
LOCATE: Reading to Locate Online Information		
Can students:		
• locate the correct email message in an inbox or the correct section of a wiki?	Locate 1	1/0
• use appropriate key words in a search engine?	Locate 2	1/0
• locate the correct site in a set of search engine results?	Locate 3	1/0
• identify correct website addresses in two different search tasks?	Locate 4	1/0
EVALUATE: Reading to Evaluate Online Information		
Can students:		
• identify the author of a website?	Evaluate 1	1/0
• evaluate the author's level of expertise?	Evaluate 2	1/0
• identify the author's point of view?	Evaluate 3	1/0
• evaluate the reliability of a website?	Evaluate 4	1/0
SYNTHESIZE: Reading to Synthesize Information		
Can students:		
• summarize an important element from one website?	Synthesize 1	1/0
• summarize important elements from two websites?	Synthesize 2	1/0
• summarize important elements from a second set of two websites?	Synthesize 3	1/0
• summarize important elements from the websites in the research task to develop an argument?	Synthesize 4	1/0
COMMUNICATE: Reading and Writing to Communicate Information		
Email:		
Does the student:		
• include the correct address line in an email message?	Communicate 1	1/0
• include an appropriate subject line in an email message?	Communicate 2	1/0
• include an appropriate greeting in an email message to an important, unfamiliar person?	Communicate 3	1/0
• compose and send a well-structured, short report of their research in an email?	Communicate 4	1/0

—OR—

Wiki:

Does the student:

• make a wiki entry in the correct location?	Communicate 1	1/0
• use descriptive voice in an informational wiki?	Communicate 2	1/0
• include an appropriate heading for a new wiki entry?	Communicate 3	1/0
• compose and post a well-structured, short report of their research in a wiki?	Communicate 4	1/0

While the ORCA could be considered a robust and authentic assessment for numerous reasons, its scoring rubric is arguably its Achilles heel. Specifically, while using dichotomous scoring improves inter-rater reliability and facilitates item response function, it arguably comes at the expense of validity. Though some ORCA tasks are well suited to dichotomous scoring (e.g., Communicate 1 – Does the students make a wiki entry in the correct location?), others (e.g., Communicate 4 – Does the student compose and post a well-structured, short report of their research in a wiki?) would benefit from a more nuanced scale (i.e., a four-point scale). Using dichotomous scoring for all ORCA tasks does not seem to represent the depth, breadth, and importance of some of the tasks within the ORCA. While it seems reasonable to weigh Communicate 1 with 1 point, the task for Communicate 4 is much more complex and involved. More importantly, Communicate 4 assesses students on more important learning than Communicate 1 because it assesses higher-ordered thinking. Thus, it merits a higher weight. By continuing to use dichotomous scoring for Communicate 4, various stakeholders would be deprived of more detailed information about what in particular needs to be done to improve student performance in this area. Perhaps a compromise would be to continue using dichotomous scoring throughout the ORCA (as it is appropriate for the majority of the assessments tasks), with the exception of the Communicate 4 task, which could benefit greatly from a more nuanced scoring procedure such as the use of a holistic or analytic rubric.

As can be seen from Table 5-11, the use of the current scoring rubric greatly facilitates inter-rater reliability. Additionally, all tasks that contained select-response items (such as in the prior knowledge multiple choice quiz on eye sight at the beginning of the ORCA) were scored automatically to avoid error. For constructed-response items, a team of scorers were trained to a minimum inter-rater reliability level of 90% accuracy by a team of expert scorers. Throughout the scoring process, there were checks to ascertain that inter-rater accuracy levels remained above 90% for each item. If the inter-rater accuracy level dropped, the scorers were retrained by expert scorers and retested before being allowed to proceed. In terms of other measure to improve reliability, to prevent systematic error, test administration was led by a trained test administrator and followed a standard protocol.

On the negative side, the scoring procedures were opaque to the test taker. For example, at no point during the ORCA were test takers made aware of how their responses would be scored. In order to guide test takers to know what counts and where to allocate their time, it would be beneficial to have more transparency about the scoring procedures. For example, rubrics are often included with assessments so that students are aware of the assessment criteria. In sum, a number of procedures were used to ensure that the ORCA had strong reliability, though, arguably, validity was sacrificed due to dichotomous scoring procedures.

Appraising the Assessment Results

In order to appraise the assessment results, I return now to the inferences made during phase 1 and appraise those inferences in light of the evidence presented from the preceding phases. The first inference presented was the scoring inference: the scoring procedures for the ORCA are appropriate and they are administered correctly, reliably, and free from overt bias. As I argued previously, the ORCA possessed high reliability, which has, arguably, compromised its validity.

Many procedures were used to ensure reliability with the ORCA. First, in terms of the internal consistency of the assessment items, experts determined that the KR-20 for the full scale ranged from .86 to .90 for the ORCA, depending on the version used. The KR-20 was .80 for the cosmetic contact lenses version used in this study (Leu et al., 2014). For each of the separate component values, the KR-20 values were low to moderate, which would be expected with only four items per component. These KR-20 values were .56 for Locate, .43 for Evaluate, .67 for Synthesize, and .51 for Communicate (Forzani, 2016). Aside from internal consistency, reliability was shown through the use of standardized protocol for administration procedures, think-aloud sessions to ensure that there was no ambiguity in the assessment prompts, and rigorous rater training and checks for inter-rater reliability (with 90% inter-rater agreement required). However, in terms of validity, the dichotomous scoring procedures used compromise the feedback available to stakeholders as the feedback only states whether the test taker got the question right or wrong, with inaccurate feedback regarding partially correct answers (i.e., those answers received a zero). Whether or not the ORCA is free from overt bias is not known. I personally did not observe any bias. However, in order to determine this for certain, it would be useful to have the ORCA reviewed by a sensitivity committee trained in detecting gender, cultural, socio-economic, urbancentric, or other forms of bias.

The generalization inference was that the ORCA and its associated inquiry-based tasks, assessment prompts, and scoring rubric capture the construct of online research and communication in the context of seventh grade science specifically and seventh grade literacy more broadly. I would argue that a serious attempt has been made to draw a representative sample from the universe of generalization, as previously outlined. For one, the ORCA aligns with numerous CCSS seventh-grade literacy and science strands. Further, while the ORCA is by

no means a perfect representation of the construct of ORW (i.e., there remain issues of construct underrepresentation) at the seventh grade level, many skills are addressed on the ORCA that are certainly not addressed by other large-scale assessments, and to an extent, not even by classroom assessment where the focus is still on literacy (i.e., traditional reading and writing; Lankshear & Knobel, 2007). Many have argued that traditional literacy assessments do not account for the profound ways in which the Internet has revolutionized the practices of reading and writing (Leu et al., 2016; Stagg Peterson et al., 2011). Thus, data from the ORCA will help stakeholders develop a more complete picture of literacy competency. It is important to note, however, that data from the ORCA should not be used to generalize to the writing construct as a whole, nor do the assessment designers claim that it should.

Just as the generalization inferences were for the most part supported, so too was the extrapolation inference: The ORCA accurately represents the range of literacy practices that students might encounter in their personal, academic, and eventual work lives in a global information age. Understandably, no assessment could ever cover the complete range of literacy practices that students might encounter. However, this assessment provides a reasonable representation of ORW tasks, particularly if the data are meant to be used in conjunction with rich and varied assessments in addition to the ORCA.

In this study, I invited participants to complete Venn diagrams describing the online researching and writing practices from the ORCA and to compare and/or contrast those to ones they use in their school, work, and personal lives. In other words, I wanted to investigate whether or not the writing practices on the ORCA could be extrapolated beyond the test, and for that matter, beyond school walls. Through analyzing the Venn diagrams, I noted a number of practices that participants engaged with on the ORCA that carried over to their school, work, and

professional lives. For example, participants described using search engines, researching topics using websites, questioning the credibility of websites, and using online communication tools such as the ones found in the ORCA (i.e., email, wiki, instant messaging, and social networking). Thus, assessing students' researching and writing practices in an online context gave the ORCA strong face validity (i.e., the ORCA appears to measure what it is purported to measure) in addition to helping support the extrapolation inferences.

Where these extrapolation inferences were not supported, however, was with regards to going beyond surface-level researching and writing practices. This begs the question about whether or not seventh graders could be expected to conduct such practices requiring higher-ordered and metacognitive thinking. For example, should a seventh grader be expected to question the degree of the source (i.e., knowing the difference between primary and secondary sources of information, and which are more credible)? The ORCA currently does not require students to question the source degree. Of the 10 novice participants in this study, two participants (Darryl and Connor) mentioned that while they use Wikipedia for online research, they look at the references and the bottom of the page and go back to those. In other words, they question the degree of the source. Thus, while more research is needed with regards to the developmental levels appropriate for critically evaluating online sources, it seems reasonable that students at this level could be expected to question the degree of a source.

A similar issue which compromises the extrapolation inferences concerns the Communicate domain of the ORCA. Large-scale writing assessments have been critiqued for some time now because they constitute their own genre of writing. The so-called assessment genre, according to Lau (2018), is theorized to be a unique genre of writing found only on assessments that possesses no function aside from external evaluation. This leads to superficial,

formulaic writing, according to many in the field of writing assessment (Camp, 2009; Hillocks, 2002; Slomp, 2007). It is often the case with the assessment genre that rhetorical knowledge is under represented as the writing tasks are devoid of an authentic purpose and audience. To compensate for the inadequacies of the assessment genre, the ORCA's test designers formulate an authentic-feeling audience (Mr. Henry's classroom wiki) and purpose (to answer to an inquiry-based problem, e.g., 'Are cosmetic contact lenses dangerous to eye health?'). In this way, the ORCA does have an authentic feel and, to an extent, this minimizes the deficiencies of the ORCA's assessment genre. As such, a number of participants commented that the ORCA felt authentic. Referring to the tasks found on the ORCA, Judy (a registered dietician) noted that,

This is something I do in my job all the time. I find and read online information. Then, in my mind, I think, okay, what was the overall message from all of these articles? How are [cosmetic contact lenses] harmful and why? I try to write that in a succinct manner. Also, I try to reference where I found that information [...] I recorded points for and points against and then synthesize that into one overall message. Rather than being balanced, I look at the weight of both sides. In this case, there was a lot more risk than benefit so that should come out in the summary.

Similarly, all of the novices indicated that the tasks on the ORCA reflected the ORW they do at school, at least to some extent. For example, Anna noted that "going on different websites and finding information is something that I do in history class quite often."

In addition to the issues created by assessment genre (i.e., lack of a real purpose and audience) are coinciding issues regarding the timed, impromptu nature common to large-scale writing assessment tasks. The tasks on the ORCA are impromptu in that students write about an

issue in which they have little background knowledge and are, nevertheless, expected to write intelligently about this topic on the day of the test. They are timed, of course, in that there are generally significant time constraints imposed on crafting a response to what are often complex tasks. The timed, impromptu writing task found on the ORCA hearkens the debate discussed earlier about the utility of this style of assessment and whether it should have any place in writing assessment. As discussed earlier, the highly-restrictive timed nature of these tests poses an issue of construct-irrelevant variance. Furthermore, many participants (experts in particular) noted that their ORW process on the ORCA was severely truncated and simplified in comparison to those processes they would take in real life.

Lastly, I would argue that for the most part, the decision inference was also supported. The decision inference stated earlier was the following: Stakeholders can use ORCA data to assess accurately students' ability to conduct and communicate online research so that stakeholders can determine next steps for learning, research, professional development, and curriculum development. This inference is supported by the fact that both the novices and experts in my study noted that the tasks found throughout the ORCA were a true reflection of their ORW practices and school, work, and beyond. Further, the detailed nature of the reports provided to teachers and administrators would be useful in supporting next steps at the program and classroom level. Finally, the addition of the TIUS, SIUS, and demographic data will also help researchers and administrators recognize barriers and facilitators that support students in acquiring New Literacies.

While the inferences have been considered for analyzing the assessment results as a whole, there has not been a sufficient disaggregated performance analysis (Slomp et al., 2014).

As previously mentioned, at this point, we cannot determine whether group variance is construct-irrelevant or relevant.

Appraising the Assessment Consequences

No assessment is ever neutral. Each assessment carries with it both positive and negative, intended and unintended consequences. Thus, the final stage of IDAF requires an appraisal regarding whether or not the aims of the assessment were met, and whether or not—taken on balance—the positive consequences outweigh the negative ones. I would argue that, taken on balance, the positive consequences of the ORCA do indeed outweigh the negative ones and that the aims of the assessment were met. I will now present each in turn.

In terms of the negative consequences, there are three major concerns. Firstly, the reliability of the ORCA came, to some extent, at the expense of validity. In particular, while the dichotomous scoring used on the ORCA facilitated high-interrater reliability and item response processes, the feedback available to stakeholders was, as a result, less nuanced and less accurate. The most obvious example of this was that, in many cases where participants received a score of 0/1, they had in fact responded with a partially correct answer. For example, in item E2, the student receives a score of 1/1 if “[t]he student judges the author's level of expertise AND provides an appropriate supporting detail about the author’s level of expertise (or lack thereof).” The following response from Judy, however, received a score of 0/1 because she failed to explicitly evaluate the expertise of the author: “She works at an eye care clinic as an assistant but does not seem to have any specific eye training. Her education background is in Health Journalism.” Judy implies that the author of the article is not an expert, but that is insufficient for receiving a point on the ORCA, despite the fact that Judy clearly understands why the author in question is not an expert in eye health.

Secondly, the timed, impromptu nature of the researching and writing tasks on the ORCA promoted more superficial and formulaic researching and writing practices, ones which are often associated with the assessment genre. This is particularly problematic because students begin to associate the processes used on timed, impromptu assessments as representative of the researching and writing processes used beyond the test. This issue is problematized by Hanson (2000) who, using a semiotic lens, argued that, with timed writing assessments, there is an inherent danger of the signified (e.g., a test, which represents writing used in the assessment genre) becoming the signifier (e.g., what it means to write in general). Professors in Lau's (2013) study of timed writing assessments, for example, noted that students often employed the same formulaic and superficial writing style that earned them high marks in an exam to their term papers. With regards to the ORCA, it will be important for teachers to acknowledge this with their students. Just because extensive background and fact checking strategies, in addition to recursive writing processes, were not employed when composing the wiki entry on the ORCA, it does not mean that these practices—among other more metacognitive and higher-order thinking skills—should not be used outside of test taking contexts.

Thirdly, as Behizadeh (2014) argued, there is always a danger in large-scale writing assessment results being used to tell a “single story” wherein “a student writes one timed essay on an assigned prompt, and then this one writing sample is used to evaluate that student’s writing ability” (p. 1). Meanwhile, Behizadeh reminds us, that student who performed poorly on a large-scale writing assessment could have a carefully crafted novel stowed in her backpack. Not only does an emphasis on large-scale assessments have negative consequences for teachers, but for students too. For teachers, over-emphasis on the single story of large-scale assessments leads to teachers taking time away from teaching writing to focus on test-taking skills. It also leads to a

narrowing of the writing construct by focussing on product over the process, focussing on narrow/traditional genres, and by ignoring the importance of multiple assessments over time (Slomp et al., 2014). For students, research has shown that standardized test scores led to students' negative self-image as writers and lowered their motivation to learn (Slomp et al., 2014). While this is not the case for the ORCA because reports of individual student performance are not provided, caution is urged should the ORCA change its role from that of a formative to a summative assessment.

Although these negative consequences should be considered in any future iterations of the ORCA, there were also a number of positive consequences. First, as previously highlighted, the ORCA is the only large-scale assessment to date to assess New Literacies. Furthermore, even classroom assessment is focussed on literacy and seldomly on New Literacies. Thus, the ORCA provides a valuable form of feedback to stakeholders on an aspect of the construct typically ignored on other assessments. Also, as mentioned, the assessment data is formative and not summative in nature. The scores that students receive on the ORCA are not reflected in their final grade. For that matter, only class and school data are made available regarding ORCA scores, thus the data is at a system level only, and does not single out individual students.

Second, in addition to the ORCA test score data itself, the ORCA was accompanied by TIUS, SIUS surveys, as well as demographic data, which is an important first step in identifying issues of fairness. For example, disaggregated performance data could be used to identify construct-irrelevant variance between groups. For example, does the video games ORCA scenario favour males over females (as females are less likely to play video games and have background knowledge about them)? This disaggregated performance data could also be used to explore barriers and facilitators to achievement. As Davis, Sumara, & Luce-Kapler (2015) have

argued, “those who own the latest tools and who have the broadest access to digital services have huge advantages over those who do not” (p. 128). These inequities are “inevitable,” argued Davis et al. (2015), and “it takes effort to monitor and modulate them” (p. 128). The ORCA and its accompanying survey instruments could be an important tool in this effort. For example, data from the ORCA have already been used to justify one-to-one laptop programs in schools as research has shown that access to networked technologies mitigates the digital divide (Kennedy et al., 2016). On balance, while there is a room for improvement with regards to the design of the ORCA, the ORCA is an innovative assessment that provides much needed data with regards to students’ ability to conduct and communicate online research, in addition to the systemic and demographic data which may mitigate these abilities. Furthermore, caution should be exercised if the ORCA were to change from its current use as a formative assessment to being used for summative assessment purposes.

Limitations

Validation work is a complex and ongoing process. As such, there are a number of limitations to this study. For one, I relied on data from my previous novice (Grade 7 – 9 students) and expert (knowledge workers) study (see Chapter IV) to investigate response processes. Obviously, this elicited a wide range of response processes as these groups vary with regards to age, expertise, and their ability to articulate their response processes. However, I feel that the benefit of having included adult expert participants to obtain validity evidence for an assessment intended for Grade 7 students is that I was able to articulate a much broader range of response processes. In this way, were adaptations made to the scoring tool (i.e., the use of an analytic rubric), more sensitive and nuanced achievement indicators could be articulated.

Further, the adults in the study were much more able to articulate the responses that they used, which led to a deeper understanding of them.

Additionally, the sample size was relatively small ($n = 20$). In order to better understand response processes used during the ORCA, it would be fruitful for future research to examine response processes with students solely from Grade 7, as this is the age group for which the ORCA is intended. A large, stratified random sample of high and low performing students would make for more generalizable results. Further, in this study I only investigated response processes of the cosmetic contact lenses version of the ORCA, which focussed specifically on investigating conflicting claims. Future research could examine the learn more about ORCA versions to see if these present different response processes. Another short coming of this study is that I have not researched how the ORCA data was actually used and interpreted by administrators and teachers in their schools and classrooms. While I have discussed how the ORCA data should be used and interpreted, it would be useful to know what is actually happening with this data in order to see whether it is being used appropriately or not.

Conclusion

This study sought to address the following research question: How should test score data from the ORCA be used and interpreted? The ORCA represents a novel and important advancement in designing Internet-based, large-scale literacy assessments. While other writing and literacy assessments have included digital components, none until the ORCA has required students to actually use the Internet to perform an inquiry-based research task. Embedded into the design of the ORCA were a number of features—including a search engine, websites, social media platform, instant messaging, email, wiki, and even avatars—which participants cited as giving the ORCA an authentic feel that mirrored their everyday ORW practices. Thus, ORCA

test score data should be used by stakeholders to better understand students' ORW competency, something that is often not captured by other assessments, particularly large-scale ones.

This study used the IDAF (Slomp, 2016) to appraise the design of the ORCA in order to better understand the ways in which its assessment results should be used and interpreted. This model was used to appraise the ORCA in terms of its aims, foundational elements, assessment program, scoring system, assessment results, and assessment consequences. During an earlier phase of this study, I identified the target domain of ORW. Then, in this phase of the study, I juxtaposed the target domain with the sample of observations from the ORCA. What I found was that a breadth—though not necessarily a depth—of skills were represented on the ORCA across the domains of Locate, Evaluate, Synthesize, and Communicate. Perhaps the most prominent issue of construct underrepresentation was seen in the Communicate domain. The issue was not with the Communicate task (i.e., wiki entry) itself, but rather with regards to the superficial scoring criteria compounded by the use of dichotomous scoring. An interrelated issue was posed by the timed, impromptu nature of the assessment. A number of participants reported feeling rushed during the ORCA, leading to issues of construct-irrelevant variance.

Concomitantly, the timed nature of the ORCA led to issues of construct underrepresentation as the writing and researching processes were truncated, leaving no time for an iterative, process-oriented approach. While many participants noted that the ORCA tasks and communications tools had an authentic feel, participants—particularly experts—also noted that they would never feel comfortable in their professional contexts publishing a wiki entry without extensive fact checking and revision. While ORW in real-life contexts often does have time constraints, the one imposed on the ORCA were much more restrictive. While a professional might have two days to write a wiki entry or blog post, participants on the ORCA took less than an hour. There

is also the danger that students might assume that it is appropriate to truncate the writing process without appropriate fact-checking and revision, such as was expected on the ORCA.

Just as with any other assessment, using ORCA data carries with it a series of consequences, both negative and positive, intended and unintended. One unintended consequence of designing the ORCA with high reliability measures, such as the use of dichotomous scoring, was that validity was sacrificed to an extent. Where this is most evident is with regards to the superficial scoring criteria used to assess the wiki entry. While the wiki entry writing task is in and of itself a robust task, the scoring criteria do not account for the complex, situated nature of ORW. A second unintended consequence was that the timed, impromptu nature of the ORCA led to participants truncating the researching and writing processes in order to meet the restrictive time constraints. Emphasizing timed, impromptu writing on standardized assessment could lead to a washback effect of teachers and students focussing on product over process, and resorting to superficial, formulaic responses, as some research has begun to indicate (Lau, 2013).

Aside from the negative, unintended consequences, there were many positive ones as well. Considering that class and school level reports—not individual reports—were provided to teachers and administrators, ORCA data was meant to inform program- and system-level decisions. Particularly, the ORCA was meant to assist teachers and administrators by providing an “early look at some of the types of skills that are likely to appear on the CCSS assessments” (Leu & New Literacies Research Lab, 2012). Because the ORCA is aligned with CCSS, the ORCA provides formative assessment feedback across a number of reading, writing, and science CCSS strands. Caution should be used, however, in interpreting ORCA results as an indication of writing proficiency, or literacy proficiency more generally. The ORCA features a particular

genre (argumentative writing), in a particular subject (science), using particular communication technologies (wiki, among others), and with a particular constraint of time (timed, versus sustained writing tasks). Thus, it would be inappropriate to generalize ORCA results beyond their stated context. Used appropriately, the ORCA should be used in conjunction with multiple and varied assessments over time.

Implications

This study has implications on a number of fronts, most notably with regards to informing how ORCA data might be used and interpreted, as well as how the ORCA might be redesigned in future iterations. Further, it is not a matter of if, but when, large-scale online literacy assessments become the standard. Thus, the ORCA and its accompanying validation work have the potential to share future writing assessments.

Considering the results of this study, I have the following recommendations:

- Use the ORCA as a formative assessment. Numerous studies have shown the deleterious effects of summative, large-scale writing assessments (Behizadeh, 2014; Hillocks, 2002; Slomp et al., 2014). While those who administer these tests claim to do so in the name of improving student learning, there is little to no evidence to support this claim. The ORCA is currently used as a formative assessment, and in my opinion, it should be kept this way.
- Interpret the ORCA results, particularly the Communicate scores, with caution. The dichotomous scoring technique used might suggest that students have no understanding of the communicate construct, when in fact they have at least a partial understanding. Further, because digital literacies knowledge and writing knowledge are so intertwined on the ORCA, students might score zero on the communicate domain because of a lack of

digital literacies knowledge (e.g., not knowing to click the pencil icon to edit the wiki), which should not suggest a lack of writing ability. Thus, the Communicate score should not be used to generalize about students' writing ability.

- Redesign the ORCA to eliminate dichotomous scoring for more complex tasks. For the task of writing the wiki entry, for example, an analytic rubric would provide more accurate and more nuanced feedback when compared to a dichotomous score.
- Interpret the results of the ORCA alongside multiple and varied forms of assessment over a prolonged period of time. This helps to mitigate the dangers of a single story (Behizadeh, 2014), the idea that one assessment result should ever be used to suggest a student's writing ability. Writing is complex, varied, and socially-situated. No one assessment can possibly measure writing ability. Teachers and stakeholders need to acknowledge this and recognize that standardized assessment results—from the ORCA or any other test for that matter—are merely partial representations of the nature and degree of student learning. In order to get a holistic picture of students' writing ability, the ORCA results are best used as one data point among many. For example, combining ORCA results with writing portfolio assessment (Behizadeh, 2014; Hicks et al., 2007; Pullman, 2002) might be a promising means to track student growth across a range of genres and contexts over time.
- Redesign the ORCA to include UDA features such as text-to-speech and speech-to-text capability; text descriptions of all graphics; adjustable font size; electronic dictionary, thesaurus, and translation tools; graphic organizers for organizing and planning writing; and extremely generous time limits. Owing to the fact that implementing UDA features

in digital assessments is relatively straightforward, this seems like an obvious and easy approach to increasing fairness for all students.

- Use the on-demand tasks of the ORCA to complement classroom writing, including parallel, but protracted online research writing tasks. Teachers can help students understand the differences between writing for an assessment genre as opposed to the iterative, process-oriented approaches required for protracted ORW tasks.

While the above uses, interpretations, and suggestions for redesign are specific to the ORCA, in many cases they can be applied to large-scale writing assessments more generally.

Because constructs such as writing are constantly in flux, it is important to note Messick's (1992) observation that “validity is an evolving property and validation is a continuing process” (p. 1487). This sentiment echoes Beaufort (2007), who cautions that there is no one, correct way of assessing writing: “[o]thers have repeatedly documented over and over the context-specific expectations about what counts as *good writing*,” adding that “writing standards are largely cultural and socially specific” (p. 11). Therefore, it was not the purpose of this validation study to argue whether or not the ORCA is valid. Rather, the goal was to present a comprehensive argument—using the IDAF—demonstrating the ways in which ORCA test scores should be used and interpreted, and to inform future design iterations. In conclusion, the ORCA is an innovative large-scale formative assessment that provides data to inform decisions at a program- and system-level and to inform research with regards to the assessment of New Literacies. Such data are needed as we strive to understand how students are adapting to writing in our global information age.

EXECUTIVE SUMMARY

Eighty-two 7th grade students at _____ Middle School completed two online research tasks in science, measuring a number of skills essential to the new Common Core State Standards (CCSS) in Reading and Writing. The assessments required students to conduct research in science on the Internet and communicate a short report of their findings online.

Performance in General Skill Areas

Your students performed best in the following general area(s) of online research and reading comprehension:

- *Synthesizing* (Mean = 6.44 out of 8, SD = 1.55, % correct = 80.5%)
- *Locating* (Mean = 4.56 out of 8, SD = 2.22, % correct = 57%)

Your students performed less well in the following area(s) of online research and reading comprehension:

- *Evaluating* (Mean = 4.06 out of 8, SD = 2.26, % correct = 50.8%)
- *Communicating* (Mean = 3.75 out of 8, SD = 2.35, % correct = 46.9%)

Performance in Specific Skill Areas

How did your students perform in specific skill areas related to online reading and research? Sixteen specific skill areas were assessed with two opportunities for each skill in the four areas: locate, evaluate, synthesize, and communicate. Thus, the maximum possible score was 2.0 for each skill.

The four skills most in need of improvement appear to be:

1. *Can students craft an explicit, unambiguous, response to the research question?* (Mean = 0.56, SD = 0.63)
2. *Can students communicate in a way that demonstrates awareness of audience, by including the correct address (email) or using a descriptive voice (wiki)?* (Mean = 0.63, SD = 0.81)
3. *Were students able to provide a logical explanation of the website's reliability?* (Mean = 0.63, SD = 0.72)
4. *Can students identify the author's point of view and provide one piece of evidence to support the answer they provide?* (Mean = 0.88, SD = 0.72)

The four strongest skills appear to be:

1. *Can the student summarize the information from the first two websites?* (Mean = 1.69, SD = 0.48)
2. *Can students identify the author/creator of the website?* (Mean = 1.63, SD = 0.62)
3. *Can the student summarize the information from the second two websites?* (Mean = 1.63, SD = 0.62)
4. *Can the student summarize information, related to the question, from a website?* (Mean = 1.56, SD = 0.51)

OVERVIEW OF THE PROJECT

During the 2011-12 school year, a sample of 7th grade students at your Middle School completed two assessments of their ability to conduct research online, read for comprehension, and construct a report of their work. The assessments took place online with laptop computers. The assessments were developed around an online reading and writing task that will define students' learning in the 21st century – the ability to conduct research on the Internet and communicate a short report of their findings online. This requires successful performance on a number of Common Core State Standards (CCSS) in both reading and writing:

Reading (Anchor Standards):

1. Read closely to determine what the text says explicitly and to make logical inferences from it; cite specific textual evidence when writing or speaking to support conclusions drawn from the text.
6. Assess how point of view or purpose shapes the content and style of a text.
7. Integrate and evaluate content presented in diverse media and formats, including visually and quantitatively, as well as in words.
8. Delineate and evaluate the argument and specific claims in a text, including the validity of the reasoning as well as the relevance and sufficiency of the evidence.
9. Analyze how two or more texts address similar themes or topics in order to build knowledge or to compare the approaches the authors take.

Writing (Anchor Standards):

1. Write arguments to support claims in an analysis of substantive topics or texts, using valid reasoning and relevant and sufficient evidence.
2. Write informative/explanatory texts to examine and convey complex ideas and information clearly and accurately through the effective selection, organization, and analysis of content.
6. Use technology, including the Internet, to produce and publish writing and to interact and collaborate with others.
7. Conduct short as well as more sustained research projects based on focused questions, demonstrating understanding of the subject under investigation.
8. Gather relevant information from multiple print and digital sources, assess the credibility and accuracy of each source, and integrate the information while avoiding plagiarism.

THE ONLINE RESEARCH AND COMPREHENSION ASSESSMENT (ORCA)

Students were presented with two research questions relating to human body systems, a curriculum topic in seventh grade science, and were asked to conduct research online and write two short reports. One question asked students to communicate their report in an email message; the other asked them to communicate their report in a wiki. The skills that were evaluated in this assessment include:

1. Locating information online with a search engine;
2. Critically evaluating information that appears at a web page;
3. Synthesizing, or summarizing, information from multiple web sites; and
4. Communicating a short research report in an email message or on a wiki.

In the online assessment system, students were guided by an avatar who presented text messages within a social network. Students were guided through each aspect of online research and reading by this avatar within the online assessment system. Each research activity contained 16 possible score points. There were 32 total score points possible for both assessments.

The skills evaluated on this assessment are likely to be included in the upcoming Common Core State Standards (CCSS) assessments, scheduled to be used in this state in 2014. It is quite possible that the CCSS

assessment will include a task requiring students to conduct research online and write a short report, similar to the ones your students completed.

Your School Sample

The results reported below come from the sample of 82 students at your school who received parental permission to participate and who completed two assessment activities. These results are likely to provide a reasonable representation of student performance for the entire seventh grade cohort at your school. They may be used as you plan for the upcoming Common Core State Standard Assessments in 2014.

OVERALL PERFORMANCE

Table 1 (below) indicates that, on average, seventh graders at your school were able to successfully perform 58.8% of the items (mean = 18.81 out of 32 score points).

Table 1. Online Research and Reading Comprehension Scores For Two Assessments
Total Possible Score: 16 + 16 = 32

	n	Mean	SD	% Correct
Total Score	82	18.81	6.41	58.8%

PERFORMANCE WITH LOCATING, EVALUATING, SYNTHESIZING, AND COMMUNICATING ONLINE INFORMATION

How did students perform in the four general areas important to online research? Your students performed better in the following area(s) of online research:

- Synthesizing (mean = 6.44 out of 8, SD = 1.55, % correct = 80.5%)
- Locating (mean = 4.56 out of 8, SD = 2.22, % correct = 57%)

Your students performed less well in the following skill area(s) of online research:

- Evaluating (mean = 4.06 out of 8, SD = 2.26, % correct = 50.8%)
- Communicating (mean = 3.75 out of 8, SD = 2.35, % correct = 46.9%)

These results appear in Table 2.

Table 2. Online Research and Reading Comprehension Scores for The Four Major Skill Areas.
Total for Each Area: 4.0 + 4.0 = 8.0

Variable	n	Mean	SD	% Correct
Locate Score	82	4.56	2.22	57%
Evaluate Score	82	4.06	2.26	50.8%
Synthesize Score	82	6.44	1.55	80.5%
Communicate Score	82	3.75	2.35	46.9%
Total Score	82	18.81	6.41	58.8%

Locate Score

The locate score tells you how well your students were able to locate information on the Internet as they completed their research projects. This includes skills such as using keywords with a search engine to locate the best information to solve a problem, being able to locate correct links in a set of search results,

and communicating useful sites to others. The total possible locate score for each student was 8. If your students had a mean score of 4.0, this indicates that your students were able to successfully perform half of the locate tasks. The Common Core anchor standards where locate skills are required include:

Reading (Anchor Standards):

1. Read closely to determine what the text says explicitly and to make logical inferences from it; cite specific textual evidence when writing or speaking to support conclusions drawn from the text.
7. Integrate and evaluate content presented in diverse media and formats, including visually and quantitatively, as well as in words.

Writing (Anchor Standards):

6. Use technology, including the Internet, to produce and publish writing and to interact and collaborate with others.
7. Conduct short as well as more sustained research projects based on focused questions, demonstrating understanding of the subject under investigation.
8. Gather relevant information from multiple print and digital sources, assess the credibility and accuracy of each source, and integrate the information while avoiding plagiarism.

Evaluate Score

The evaluate score tells you how well your students were able to evaluate information on the Internet as they completed their research projects. This includes skills such as determining the reliability of information they encounter. The total possible evaluate score for each student was 8. If your students had a mean score of 4.0, this indicates that your students were able to successfully perform half of the evaluate tasks.

Reading (Anchor Standards):

1. Read closely to determine what the text says explicitly and to make logical inferences from it; cite specific textual evidence when writing or speaking to support conclusions drawn from the text.
6. Assess how point of view or purpose shapes the content and style of a text.
7. Integrate and evaluate content presented in diverse media and formats, including visually and quantitatively, as well as in words.
8. Delineate and evaluate the argument and specific claims in a text, including the validity of the reasoning as well as the relevance and sufficiency of the evidence.

Writing (Anchor Standards):

2. Write informative/explanatory texts to examine and convey complex ideas and information clearly and accurately through the effective selection, organization, and analysis of content.
6. Use technology, including the Internet, to produce and publish writing and to interact and collaborate with others.
8. Gather relevant information from multiple print and digital sources, assess the credibility and accuracy of each source, and integrate the information while avoiding plagiarism.

Synthesize Score

The synthesize score tells you how well your students were able to synthesize different information on the Internet as they completed their research projects. This includes skills such as summarizing the information from one or several web sites. The total possible synthesis score for each student was 8. If your students had a mean score of 4.0, this indicates that your students were able to successfully perform half of the synthesis tasks.

Reading (Anchor Standards):

1. Read closely to determine what the text says explicitly and to make logical inferences from it; cite specific textual evidence when writing or speaking to support conclusions

drawn from the text.

7. Integrate and evaluate content presented in diverse media and formats, including visually and quantitatively, as well as in words.
9. Analyze how two or more texts address similar themes or topics in order to build knowledge or to compare the approaches the authors take.

Writing (Anchor Standards):

1. Write arguments to support claims in an analysis of substantive topics or texts, using valid reasoning and relevant and sufficient evidence.
2. Write informative/explanatory texts to examine and convey complex ideas and information clearly and accurately through the effective selection, organization, and analysis of content.
6. Use technology, including the Internet, to produce and publish writing and to interact and collaborate with others.
7. Conduct short as well as more sustained research projects based on focused questions, demonstrating understanding of the subject under investigation.
8. Gather relevant information from multiple print and digital sources, assess the credibility and accuracy of each source, and integrate the information while avoiding plagiarism.

Communicate Score

The communicate score tells you how well your students were able to communicate information on the Internet as they completed their research projects. This includes skills such as composing a clear email message or wiki post. If your students had a mean score of 4.0, this indicates that your students were able to successfully perform half of the communicate tasks.

Reading (Anchor Standards):

5. Read closely to determine what the text says explicitly and to make logical inferences from it; cite specific textual evidence when writing or speaking to support conclusions drawn from the text.
7. Integrate and evaluate content presented in diverse media and formats, including visually and quantitatively, as well as in words.
9. Analyze how two or more texts address similar themes or topics in order to build knowledge or to compare the approaches the authors take.

Writing (Anchor Standards):

1. Write arguments to support claims in an analysis of substantive topics or texts, using valid reasoning and relevant and sufficient evidence.
1. Write informative/explanatory texts to examine and convey complex ideas and information clearly and accurately through the effective selection, organization, and analysis of content.
6. Use technology, including the Internet, to produce and publish writing and to interact and collaborate with others.
7. Conduct short as well as more sustained research projects based on focused questions, demonstrating understanding of the subject under investigation.
8. Gather relevant information from multiple print and digital sources, assess the credibility and accuracy of each source, and integrate the information while avoiding plagiarism.

PERFORMANCE IN SPECIFIC SKILL AREAS

How did your students perform in specific skill areas related to online reading and research? Sixteen specific skill areas were assessed with two opportunities for each skill in the four areas: locate, evaluate, synthesize, and communicate. Thus, the maximum possible score was 2.0 for each skill.

The four skills most in need of improvement appear to be:

1. Can students craft an explicit, unambiguous, response to the research question?
(Mean = 0.56, SD = 0.63)

2. Can students communicate in a way that demonstrates awareness of audience, by including the correct address (email) or using a descriptive voice (wiki)? (Mean = 0.63, SD = 0.81)
3. Were students able to provide a logical explanation of the website's reliability? (Mean = 0.63, SD = 0.72)
4. Can students identify the author's point of view and provide one piece of evidence to support the answer they provide? (Mean = 0.88, SD = 0.72)

The four strongest skills appear to be:

1. Can the student summarize the information from the first two websites? (Mean = 1.69, SD = 0.48)
2. Can students identify the author/creator of the website? (Mean = 1.63, SD = 0.62)
3. Can the student summarize the information from the second two websites? (Mean = 1.63, SD = 0.62)
4. Can the student summarize information, related to the question, from a website? (Mean = 1.56, SD = 0.51)

Table 3, below, shows the mean scores for each skill at your school.

Table 3. Online Research and Reading Comprehension Scores for The 16 Skills Evaluated in This Assessment. Total for Each Skill: 1.0 + 1.0 = 2.0

	Skill	Mean*	SD
<i>Reading to Locate Online Information</i>			
1.	Can students use appropriate key words for a search engine, entering both topic and claim as terms?	0.94	0.85
2.	Within a set of search results for the first search task, can students locate a link that will take them to information that will answer the question on the first click?	1.31	0.60
3.	Within a set of search results for the second search task, can students locate a link that will take them to information that will answer the question on the first click?	1.31	0.70
4.	Can students locate and provide two URLs for websites containing information that will answer the research question?	1.00	0.82
<i>Reading to Evaluate Online Information</i>			
5.	Can students identify the author/creator of the website?	1.63	0.62
6.	Can students provide an accurate detail about the author's level of expertise?	0.94	0.93
7.	Can students identify the author's point of view and provide one piece of evidence to support the answer they provide?	0.88	0.72
8.	Were students able to provide a logical explanation of the website's reliability?	0.63	0.72
<i>Reading to Synthesize Online Information</i>			
9.	Can the student summarize information, related to the question, from a website?	1.56	0.51
10.	Can the student summarize the information from the first two websites?	1.69	0.48
11.	Can the student summarize the information from the second two websites?	1.63	0.62
12.	Can the student summarize the information related to the question from four different websites?	1.56	0.81
<i>Writing to Communicate Online Information</i>			

13.	Can students navigate to the correct communication area (the correct email or the correct section of the wiki) for the communication task on the first click?	1.56	0.73
14.	Can students use the communication tool to post or send a message with an appropriate heading (wiki) or subject line (email)?	1.00	0.89
15.	Can students communicate in a way that demonstrates awareness of audience, by including the correct address (email) or using a descriptive voice (wiki)?	0.63	0.81
16.	Can students craft an explicit, unambiguous, response to the research question?	0.56	0.63

* Max = 2.0

CONCLUSION

This report has provided an early look at some of the types of skills that are likely to appear on the CCSS assessments. I hope these data will be useful in your planning. If you have any questions, please let me know.

Respectfully Submitted,

Donald J. Leu, Ph.D.

John and Maria Neag Endowed Chair in Literacy and Technology

Director, The New Literacies Research Lab

<http://neag.uconn.edu/donald.leu/>

Neag School of Education

University of Connecticut

249 Glenbrook Road

Storrs, CT 06269-2033

Appendix I – Profile of Participants

Table1. Expert Participants: Overview of Demographic Characteristics and Experiences with Online Research

Pseudonym	Occupation	Age	Gender	Educational Attainment / Grade Level	Frequency of online research	Preferred Communication Media	Preferred Research Media
Trina	Technical writer / Business analyst	51	Female	Bachelor's degree	Daily or more	Social networking sites; email	Twitter; Google alerts; news articles; research papers
Catherine	Community Outreach Lead, education programs and partnerships	30	Female	Master's degree	Weekly or more	Social networking sites; email; blogs/wikis	Twitter; professional / peer-reviewed journal articles
Robert	Nuclear safety analyst	41	Male	Bachelor's degree	Weekly or more	Email; reports; websites	Web sites; professional publications and newsletters; reports
Lisa	Registered Massage Therapist	35	Female	College diploma	Monthly or more	Social networking sites; social bookmarking sites; blogs/wikis	News and magazine articles; meetings and/or presentations; video (e.g., television news, vodcasts) and/or Audio (e.g., radio news, podcasts)
Margaret	English Teacher, Curriculum Coach	55	Female	Bachelor's degree	Daily or more	Social networking sites; microblogs; blogs/wikis	Social networking sites; professional publications/newsletters; meetings and/or presentations
Robin	Financial Management Analyst	42	Female	Bachelor's degree	Daily or more	Email; reports; meetings/presentations	Reports; websites; meetings and/or presentations
Laura	Director and Senior Counsel (lawyer)	52	Female	Bachelor's degree	Daily or more	Email; meetings/presentation; websites	Professional publications/newsletters; websites; dictionaries and language guides
Michelle	Professor of Educational Technologies in a Faculty of Education	45	Female	Doctorate	Weekly or more	Blogs /Wikis; email; social networking sites	Professional / peer-reviewed journal articles; websites; social networking sites

Julie	Registered Dietician	34	Female	Bachelor's degree	Monthly or more	Meetings/presentations; email; pamphlets/brochures	Listserves; professional / peer-reviewed journal articles; electronic medical records
Colin	Program Engineer, Regulatory Administration	37	Male	Bachelor's degree	Weekly or more	Email	Email; reports; websites

Table 2. Novice Participants: Overview of Demographic Characteristics and Experiences with Online Research

Pseudonym	Grade	Age	Gender	Frequency of Internet Use	Preferred Online Activities	Other Details
Helen	7	12	Female	20 minutes/day at school; 20 minutes/day at home	Email; online research for school projects; blogging	Fraternal twin of Jake; loves to draw and write fictional stories; uses Google Classrooms and Chromebooks at school
Jake	7	12	Male	20 minutes/day at school; 20 minutes/day at home	Web searching to learn more about something; research for school projects	Fraternal twin of Helen; won a national essay writing contest
Sue	6	11	Female	20 minutes/day at school 60 minutes/day at home;	Research for school projects; listening to music; watching TV shows; computer games involving coding; Googling funny things about cats; Finding recipes	Uses iPads and Chromebooks at school; mom is a teacher; goes to a French immersion school
Jill	8	13	Female	20 minutes/day at school 20 minutes/day at home;	Research for school projects; messaging friends; occasional use of Instagram	Goes to French school; swims competitively
Cecilia	8	13	Female	60 minutes/day	Research for school projects; email	Homeschooled; sibling of Darryl; dad is a computer programmer
Darryl	9	14	Male	90 minutes/day	Email; social media (<u>Instagram</u>); researching school projects; gaming (<u>MineCraft</u> ; <u>Mario Kart</u>)	Homeschooled; speed reader; gamer; sibling of Cecilia
Natalie	7	12	Female	60 minutes/day at school; 60 minutes/day at home;	Playing games on iPod; YouTube; doing research for homework; email	Each class has a set of Chromebooks; goes to French school
Fiona	8	13	Female	30 minutes/day at school (twice a week); 10 minutes/day at home	Online language translators; research for school projects; watching movies; email	Identical twin of Anna; goes to French immersion school; mom is a teacher; Dad is VP product development at a high tech firm
Anna	8	13	Female	50 minutes/day at school (a couple of times a week);	Research for school projects; email	Identical twin of Fiona; goes to French immersion school; mom is a teacher; Dad is VP product development at a high tech firm; class uses Chromebooks

				30 minutes/day at home		
Cody	8	13	Male	“most of the day”	Online games (e.g., <u>Papa’s Pizzeria</u> and <u>Fire Ball</u>); research for school projects	Won a spoken word contest where he talked about his experience living with Asperger’s

CHAPTER 6 - CONCLUSION

Conclusion

This study concerns an appraisal of a large-scale assessment of New Literacies known as the Online Research Comprehension Assessment (ORCA). The ORCA is ground-breaking in the sense that it is the first large-scale literacy assessment to evaluate students not only in a digital context, but in a fully online context. During the ORCA, students are presented with inquiry-based research tasks and are then challenged to conduct and communicate research using the Internet. While completing the assessment, students are prompted to locate, evaluate, and synthesize online information. At the end of the assessment, they are prompted to communicate their findings via a classroom wiki or email.

The purpose of this meta-study was to investigate how ORCA test score data should be used and interpreted. Previous validation work done regarding the ORCA included cognitive labs completed over two years with approximately 300 students; pilot testing with 1600 students; and, a panel review of the ORCA by measurement and online research and comprehension experts. Further, previous reliability estimates by measurement experts determined that the KR-20 for the full scale ranged from .86 to .90 for the ORCA-Closed format, depending on the version of the ORCA. The KR-20 was .80 for the cosmetic contact lenses version used in this study (Leu et al., 2014). For each of the separate component values, the KR-20 values were low to moderate, which would be expected with only four items per component. These KR-20 values were .56 for Locate, .43 for Evaluate, .67 for Synthesize, and .51 for Communicate (Forzani, 2016).

My study expands on this initial work to consider a more holistic and interconnected appraisal of the ORCA by examining its validity, reliability, and fairness using the Integrated Design and Appraisal Framework (IDAF; Slomp, 2016). In doing so, my research demonstrated

how ORCA test score data should be used and interpreted, in what context, and what the consequences (intended and unintended, positive and negative) can be anticipated.

IDAF and other contemporary validation frameworks necessitate the development of robust constructs. In the case of writing, previous construct and theoretical work failed to take into account the profound ways in which the Internet has revolutionized the practice of writing. Finding no suitable construct in the literature, I decided to re-conceptualize the writing construct. Thus, RQ 1 of this dissertation asked, based on a systematic review of the literature, what is the construct underlying the ORCA? This led to my first study found in Chapter III entitled “Articulating an Integrated Contemporary Construct for the Digital Age.” For this study, I co-conducted a systematic review literature to integrate 50 years worth of writing research in order to articulate a contemporary writing construct (Corrigan & Slomp, 2017). After reviewing 101 texts, our analysis revealed the following domains germane to the construct of writing: metacognitive, critical discourse, discourse, rhetorical, genre, writing process, subject matter, and digital literacies knowledge. Further analysis revealed that the domain of digital literacies knowledge was integral to all other domains, thus instead of making it a domain of its own, we integrated it throughout the construct. Additionally, our analysis revealed that some domains were superordinate to other domains, as determined through our modelling in Nvivo. For example, our analysis demonstrated that metacognitive knowledge had a hierarchical relationship to all other domains. In other words, one could have metacognitive knowledge about writing process knowledge, or one could have writing process knowledge without a metacognitive understanding of it. Further, our findings led us to conclude that writing is a complex construct that embodies multiple knowledge domains—including metacognitive, critical discourse, discourse, rhetorical, genre, writing process, and subject matter knowledge—all of which are

interconnected with digital literacies knowledge. We brought these domains together to form a construct we call the integrated writing construct, which describes the knowledge domains necessary to develop expertise in writing in a digital age. We concluded by situating this construct within a bioecological model that elucidates the implications this construct has for writing curriculum, instruction, and assessment. I then used these constructs to inform my data analyses in subsequent studies in this dissertation.

The second study, Chapter IV, was entitled “Exploring the Differences between Student Writers and Knowledge Workers in Online Research Writing.” The research question posed here was, based on an empirical investigation of response processes, what is the construct underlying the ORCA? This led to my examination of the differences in responses processes between experts and novices when completing the ORCA, and in online research writing (ORW) more generally. Participants for this study included 10 novices between sixth and ninth grade, as well as 10 experts, all of whom were employed full-time in the knowledge economy and did online research and writing frequently for a living, according to their questionnaire responses. Using a novice-expert approach enabled me to observe and analyze a continuum of processes and how these differed between the two groups. All participants were invited to a university research site to complete an ORCA. Following this, I used cued retrospective reporting (van Gog et al., 2005) to elicit response processes from my participants, in other words, the cognitive and metacognitive processes elicited by the ORCA. Then, I asked participants to complete a Venn diagram demonstrating how their ORW practices were both similar and different to the ones they used at school or work. My interviews opened with a discussion of these Venn diagrams, followed by the questions posed in the interview protocol (found on p. 202). From here I asked participants to share with me an ORW artefact from school or work as we moved

into a semi-structured interview wherein we further explored their ORW practices. Data sources for this study included ORCA test scores, but also cued retrospective reporting and interview transcripts, questionnaires, and writing artefacts from school and work. The results suggest that there are complex and sophisticated cognitive and metacognitive processes underlying online research writing and, that some of these processes are unique to online contexts. Secondly, both quantitative and qualitative results suggest significant differences between novice and expert groups.

While novices and experts engaged in many common processes when it came to online research writing—e.g., they used keyword searches, skimmed texts to find the main idea, checked the author’s credentials, and summarized ideas from across multiple sources to communicate their findings—in general, the data suggested that experts possessed broader and more sophisticated cognitive and metacognitive processes. It is also possible that experts were simply better at articulating these response processes. These more complex processes were evident across the domains of Locate, Evaluate, Synthesize, and Communicate. For example, when it came to locating online information, experts were more aware of their filter bubble—that search engines use algorithms to retrieve information based on their prior search behaviours—than were novices. A prime example of this was when Judy, a registered dietician, sought out online information from a variety of perspectives, not only from the perspective of her fellow dieticians. She did this in order to better understand and communicate with multiple stakeholders, from family physicians to her patients. Experts also sought out information from multiple sources including search engines, online databases, Twitter, social networking, and a variety of alternative sources (e.g., blogs). By contrast, novices used predominantly search

engines and tended to trust the top result on the search engine results page, as has been shown in previous studies (Pan et al., 2007).

In terms of evaluating online information, novices portrayed a Wikipedia instinct (Corrigan, 2017), that is, the tendency to trust the credibility of secondary sources. Experts, on the other hand, expressed cynicism toward secondary sources such as Wikipedia, opting instead to go back to the primary source of information, or what is known as backward reference searching or even second-level backward reference searching. However, even expert participants did not engage in backward reference searching or other reliability checks arbitrarily. Rather, they engaged in triaging (Corrigan, 2017), deciding which tasks were most important to themselves and other stakeholders, and then allocating their time appropriately. If, for example, they were engaged in high-stakes ORW—such as writing a professional blog or peer-reviewed article—they conducted rigorous fact checking and engaged in numerous strategies to do so. These strategies included background checking of authors and publications, backward reference searching, second-level backward reference searching, triangulating evidence from multiple sources, and engaging with numerous and at times competing perspectives. This breadth of strategies was complemented with a more nuanced perspective with regards to credibility: whereas novices viewed the credibility of online sources dichotomously (i.e., credible or not), experts viewed it on a continuum (Corrigan, 2017).

In terms of synthesizing online information, novices tended to summarize information by closely paraphrasing the original text. By contrast, experts were better able to synthesize information, considering competing perspectives to form their own judgement (Corrigan, 2017). Experts would often combine their professional experience and expertise with information they

gathered during their research to generate new knowledge, which is a form of what Deschryver (2014) called generative synthesis.

When communicating online information, experts were much more aware of the discourse communities for whom they were writing, choosing genre and rhetorical features that supported their purpose for writing and the audience to whom they were writing. Novices, by contrast, were so preoccupied with what they were saying (knowledge telling) that they, generally speaking, did not consider their audience or purpose for writing (or what is known as knowledge transforming; Bereiter & Scardamalia, 1987). Another interesting finding of this study was that, suggesting once again that the digital native myth (Selwyn, 2009) is inaccurate, experts utilized a much broader range of online communications technologies for ORW than did novices. Novices were seen primarily using Google Docs, Slides and email. While novices might use social networking or other more alternative platforms in their personal lives, this did not transfer to their using these technologies to communicate online research, whether at home or school. Meanwhile, experts spoke of using more traditional online platforms (e.g., email and online reports) in addition to Twitter, websites, blogs, and social networking platforms to communicate their research findings. Lastly, experts tended to develop much more sophisticated organizational and time management skills as they completed writing tasks that extended over protracted time periods. These more complex and extended tasks also necessitated greater collaboration among colleagues with complementary areas of expertise. To accomplish these more protracted, complex, and collaborative tasks, experts were seen using a variety of metacognitive strategies, oftentimes offloading these tasks to web-based applications (e.g., productivity, project management, and reference managements applications).

What then do these findings suggest about ways in which ORCA test score data should be used and interpreted? The third study, Chapter V, “Using the Integrated Design and Appraisal Framework to Appraise a Large-Scale Assessment of New Literacies” examined just that. The research question explored in this study was, using the IDAF, how should ORCA test scores be used and interpreted, considering the results of the preceding literature review and empirical study, and additional research? Thus, in this study I used the results of the previous two studies to determine the target domain (i.e., the full range of possible observations associated with writing, and online research writing more specifically) in order to juxtapose that with the sample of observations elicited by the ORCA. In doing so, I examined each domain of the construct (i.e., Locate, Evaluate, Synthesize, and Communicate) describing which aspects of each domain are represented, underrepresented, or ignored on the ORCA. Examples of construct underrepresentation included students being given websites to search for instead of coming up with their own. Also, while students are prompted to evaluate the credibility of the authors who wrote the website content, they are not asked to critically evaluate the arguments made by these authors. As a final example, while writing a wiki entry to an inquiry-based task is an authentic and complex task, the writing skills and competencies needed to write such an entry are not reflected in the assessment criterion. Despite these examples of construct underrepresentation—no one assessment can measure all aspect of a construct after all—the ORCA does indeed align with numerous seventh-grade CCSS strands such as Reading Standards for Informational Texts, Reading Standards for Literacy in Science, and Writing Standards (see [Common Core Standards and the ORCA Project](#) for more). Additionally, the ORCA is measuring aspects of the writing construct that have been traditionally ignored on both classroom and traditional large-scale

writing assessments, which focus more on literacy (Lankshear & Knobel, 2007) than a contemporized construct (Corrigan & Slomp, 2017).

This study also examined issues of construct-irrelevant variance with the ORCA. One of the major concerns was with the issue of time, with most participants reporting that they felt rushed completing the ORCA. The lack of time also compounded construct underrepresentation because participants were unable to engage in process-oriented writing, which led to more formulaic and superficial ORW. Meanwhile, participants noted that were they to complete this same sort of writing task (e.g., a professional wiki entry) in the real world, they would engage in rigorous fact checking and multiple rounds of revision. Otherwise, they said, their writing would be unprofessional and perhaps even unethical, were they to disseminate inaccurate or incomplete information to their stakeholders.

A further issue of construct-irrelevant variance was created by participants' lack of digital literacies knowledge. For example, many novice participants did not realize that they could click on the hyperlink of the author's name in order to evaluate his or her level of expertise. When participants were provided this opportunity in the interview following the ORCA, they oftentimes gave correct answers whereas they did not on the ORCA. Further, in terms of the communicate score, one novice participant did not know how to click on the pencil icon to edit the classroom wiki. That participant would have received a score of 0/4 for the Communicate domain were she to have taken the ORCA in an official capacity. These examples demonstrate the complexity of assessing students on researching and writing challenges that are so intertwined with digitality. This further illustrates that results should be interpreted with caution: receiving a score of 0/4 on the communicate domain might suggest students' lack of digital

literacies knowledge (e.g., how to click on the pencil icon to edit the wiki page) and little about students' actual writing ability.

A final issue of construct-irrelevant variance was the lack of Universally Designed Assessment (UDA) features such as text-to-speech and speech-to-text capability; text descriptions of all graphics; adjustable font size; electronic dictionary, thesaurus, and translation tools; graphic organizers for organizing and planning writing; and extremely generous time limits (National Center on Educational Outcomes, 2002). Although this was not a major issue in my study, one novice participant who went to a French school and speaks English and French at home did note that there was some vocabulary used on the various websites on the ORCA that she did not understand. In this case, a translation tool embedded into the ORCA would have been beneficial. Future iterations of the ORCA should consider UDA features as a means of addressing issues of fairness.

Also with regards to fairness, it is not known at this time whether group variance on ORCA performance is construct relevant or irrelevant. Preliminary research has identified group variance in terms of gender (Corrigan, 2014a) and socio-economic status (Leu et al., 2015). In order to determine whether or not the variance among groups is construct relevant, further studies investigating response processes would be beneficial (Goodwin & Leech, 2003).

What then are the consequences—positive and negative, intended and unintended—of using the ORCA? There were two negative, unintended consequences of note. The first was that, due to dichotomous scoring procedures, the feedback available to stakeholders was less nuanced and less accurate. Secondly, the timed, impromptu nature of the researching and writing tasks on the ORCA promoted more superficial and formulaic researching and writing practices. Because research has shown the effects of washback—that classroom instructional and assessment

practices are strongly influenced by large-scale assessments—there is a concern that teachers will prioritize the writing product over the process. Further, students may conflate the superficial, product-oriented writing used on large-scale assessments with the type of writing required for more extended writing tasks beyond the test. Teachers need to help students recognize that even when extensive fact-checking and recursive writing processes are not used in testing situations, they are absolutely necessary beyond the assessment genre.

On the whole, however, I feel that the affordances outweigh the limitations because the ORCA is more formative than summative in nature—particularly because the assessment is not used in any way to make a judgement about the nature and degree of an individual student’s learning. With the ORCA, there is no danger of creating a single story (Behizadeh, 2014), such as might be the case with other large-scale writing assessments where one score is—inappropriately—used to make a generalization about a student’s overall writing ability. In fact, with the ORCA, there are no reports of individual student achievement generated. Rather, ORCA data is combined with demographic and survey data (i.e., the Teacher Internet Use Survey and the Student Internet Use Survey) to better understand systemic and classroom barriers to acquiring New Literacies. In fact, data from the ORCA have already been used to justify one-to-one laptop programs in schools as research has shown that access to networked technologies mitigates the digital divide (C. Kennedy et al., 2016). Finally, the ORCA is the only large-scale assessment to date to assess New Literacies. Even classroom assessment—which traditionally has better construct validity--tends to focus on literacy and seldomly on New Literacies (Lankshear & Knobel, 2007). Thus, the ORCA provides a valuable form of feedback to stakeholders on an aspect of the construct typically ignored on other assessments.

Implications

This study has implications on a number of fronts, particularly with regards to the ORCA's use, interpretation, and redesign. Here I wish to reiterate, although more succinctly, the recommendations I made in Chapter V:

- Use the ORCA as a formative assessment.
- Interpret the ORCA results, particularly the Communicate scores, with caution.
- Interpret the results of the ORCA alongside multiple and varied forms of assessment over a prolonged period of time.
- Redesign the ORCA to eliminate dichotomous scoring for more complex tasks.
- Redesign the ORCA to include Universal Design for Assessment features.
- Use the on-demand tasks of the ORCA to complement classroom writing, including parallel, but protracted online research writing tasks.

In addition to the above recommendations, this study has also contributed to the field of writing by reconceptualizing the construct of writing in light of the tremendous changes the practice has undergone since the advent of the Internet. Furthermore, this study has elucidated many cognitive and metacognitive processes required to write in online contexts not previously discussed in the literature—most notably, triaging and the Wikipedia instinct. It is my hope that these findings will help inform curricular, instructional, and assessment design.

Future Directions for Research

We are living in an era of accountability (Schoen & Fusarelli, 2008). If large-scale assessment is here to stay, can we do it better? What if large-scale assessment could actually be used to improve student learning, rather than being used solely for accountability and gate-keeping purposes (Nagy, 2000)? What would such an assessment look like? While large-scale

assessments are much maligned (often with good reason), they arguably serve a valuable function in our educational systems. For one, their standardized conditions facilitate comparisons among other students, schools, and regions in order to identify areas of strength and areas for improvement (Gareis & Grant, 2015). Used in conjunction with classroom assessment, they provide a more holistic and reliable judgement of student learning: “neither standardized tests nor classroom assessments alone are sufficient to meet the needs of teachers to accurately and dependably gauge and thereby support continued student learning” (Gareis & Grant, 2015, p. 13). Data from large-scale assessments such as the ORCA can help to identify and address school-and system-level barriers and gaps in learning. Yet, the negative consequences accruing from many large-scale assessments are very real. The concept of formative, large-scale assessment is relatively new. Would changing the nature of these large-scale assessments from being summative to formative in nature lessen the negative consequences and amplify the positive ones? More research is needed into large-scale formative assessments such as the ORCA to show whether or not they can be used to improve student learning.

Further, how can technologically mediated large-scale writing assessments level the playing field? How do UDA features help English language learners, those with physical and/or learning disabilities, and students in general? Also, how might computer adaptive testing help all learners feel that they can succeed? In non-adaptive testing environments, students become discouraged by only being able to answer a few questions. What if the assessments adapted to each student based on his or her interests and/or level of readiness? Computer-mediated assessments show promise when it comes to designing more inclusive and fair tests.

Finally, how do we design assessments—including large-scale ones—that align with 21st century competencies? There is a growing impetus to shift from teaching skills to teaching competencies (Ontario Ministry of Education, 2016). As the OECD (2003) explains,

[a] competency is more than just knowledge or skills. It involves the ability to meet complex demands, by drawing on and mobilising psychosocial resources (including skills and attitudes) in a particular context. For example, the ability to communicate effectively is a competence that may draw on an individual's knowledge of language, practical IT skills and attitudes towards those with whom he or she is communicating. (p. 4)

In Chapter IV of this dissertation, participants (experts in particular) discussed how they were often challenged with writing tasks that were beyond their personal expertise which required them to collaborate with other experts, sometimes locally or sometimes around the world—in other words, complex writing tasks oftentimes require distributed intelligence (Pea, 1993). Thus, it would seem fitting that we would include collaborative writing in the curriculum.

Collaborative writing in particular, however, poses challenges for assessment because it is difficult for teachers to know who wrote what and assess students according to their individual contributions.

While teaching 21st century competencies—including critical thinking, communication, collaboration—is challenging in its own right, assessment will require fundamentally different paradigms and approaches. This is particularly true in large-scale contexts. In the case of a collaboratively written text, the whole is more than the sum of its parts. While we can determine who wrote what in a Google Document, for example, in a collaboratively written text, writers are continuously building off of one another, leveraging one another's unique skills and experiences, and expanding upon them as the text metamorphoses. One student might contribute more to the

content, another to the style, and still another to the organization, and so forth. In the same vein, we will require new paradigms and approaches as we face the challenge of assessing multimodal composition, hypertextual writing, and writing in multiple languages/dialects and modes for multiple purposes.

A contemporary, integrated writing construct (Corrigan & Slomp, 2017) will pose new challenges for instruction and assessment, but also new possibilities. We are living in times of deixis, where “[o]ur graduates are entering a world that is more competitive, globally connected, and technologically engaged than any other period in history” (Ontario Ministry of Education, 2017). Indeed, these are exciting times for writing researchers as we prepare students with the skills, strategies, dispositions, and social practices that will avail them to endless opportunities in our global village.

CHAPTER 7 - REFERENCES

References

- Abbott, R. D., & Berninger, V. W. (1993). Structural equation modeling of relationships among developmental skills and writing skills in primary- and intermediate-grade writers. *Journal of Educational Psychology, 85*, 478–508. <http://doi.org/10.1037/0022-0663.85.3.478>
- Aikens, N. L., & Barbarin, O. (2008). Socioeconomic differences in reading trajectories: The contribution of family, neighborhood, and school contexts. *Journal of Educational Psychology, 100*, 235–251. <http://doi.org/10.1037/0022-0663.100.2.235>
- Alamargot, D., & Fayol, M. (2009). Modelling the development of written composition. In R. Beard, D. Myhill, J. Riley, & M. Nystrand (Eds.), *Handbook of Writing Development* (Kindle, pp. 23–47). London, England: Sage Publications Inc.
<http://doi.org/10.4135/9780857021069.n3>
- Ambrose, S. B., Bridges, M. W., DiPietro, M., Lovett, M. C., & Norman, M. K. (2010). *How learning works*. San Francisco, CA: Jossey-Bass.
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
Retrieved from <https://ebooks.aera.net/user>
- Andrade, H. G., & Boulay, B. A. (2003). Role of rubric-referenced self-assessment in learning to write. *The Journal of Educational Research, 97*, 21–34. <http://doi.org/10.2307/27542460>
- Australian Curriculum, Assessment, and R. A. (n.d.). Critical and creative thinking. Retrieved October 24, 2017, from <https://www.australiancurriculum.edu.au/f-10-curriculum/general-capabilities/critical-and-creative-thinking/>
- Bakhtin, M. (1981). *The dialogic imagination: Four essays*. (M. Holquist, Ed., C. Emerson & M.

Holquist, Trans.). Austin: University of Texas Press.

Bastos, M. T., Travitzki, R., & Raimundo, R. (2012). Tweeting political dissent: Retweets as pamphlets in #FreeIran, #FreeVenezuela, #Jan25, #SpanishRevolution and #OccupyWallSt. In *The Internet, Policy, and Politics Conference* (pp. 1–22). Oxford, England: Oxford Internet Institute. Retrieved from http://ipp.oii.ox.ac.uk/sites/ipp.oii.ox.ac.uk/files/documents/BASTOS-TRAVITZKI-RAIMUNDO_-_Tweeting_Political_Dissent.pdf

Battiste, M. (1998). Enabling the autumn seed: Toward a decolonized approach to aboriginal knowledge, language, and education. *Canadian Journal of Native Education*, 22, 16–27.

Bazerman, C. (1994). *Constructing experience*. Carbondale, IL: Southern Illinois University Press. Retrieved from <https://books.google.ca/books?id=db9GhG1qOpUC&pg=PR4&lpg=PR4&dq=bazerman+1994+constructing+experience&source=bl&ots=AjKJdFHLHf&sig=oAuI4xGg1GqicIEba7-3d3jYTIw&hl=en&sa=X&ved=0ahUKEwj0KO4soHSAhVB2oMKHeV8DTsQ6AEIPDAF#v=snippet&q=writing+expertise&f=false>

Bazerman, C. (2016). What do sociocultural studies of writing tell us about learning to write? In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (2nd ed., pp. 11–23). New York: Guilford Press.

Beard, R., Myhill, D., Riley, J., & Nystrand, M. (2009). General introduction. In R. Beard, D. Myhill, J. Riley, & M. Nystrand (Eds.), *The SAGE handbook of writing development* (Kindle, pp. 1–5). London, England: Sage Publications, Inc.

Beaufort, A. (2007). *College writing and beyond: A new framework for university writing instruction*. Logan, UT: Utah State University Press.

- Beck, S. W., & Jeffery, J. V. (2007). Genres of high-stakes writing assessments and the construct of writing competence. *Assessing Writing, 12*, 60–79.
<http://doi.org/10.1016/j.asw.2007.05.001>
- Behizadeh, N. (2014). Mitigating the dangers of a single story: Creating large-scale writing assessments aligned with sociocultural theory. *Educational Researcher, 43*(3), 125–136.
<http://doi.org/10.3102/0013189X14529604>
- Behizadeh, N., & Engelhard, G. (2011). Historical view of the influences of measurement and writing theories on the practice of writing assessment in the United States. *Assessing Writing, 16*, 189–211. <http://doi.org/10.1016/j.asw.2011.03.001>
- Behizadeh, N., & Engelhard, G. (2015). Valid writing assessment from the perspectives of the writing and measurement communities. *Pensamiento Educativo: Revista de Investigación Educativa Latinoamericana, 52*(2), 34–54. <http://doi.org/10.7764/PEL.52.2.2015.3>
- Bereiter, C., Burtis, P. J., & Scardamalia, M. (1988). Cognitive operation in constructing main points in written composition. *Journal of Memory and Language, 27*(3), 261–278.
- Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition* (Kindle). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.
- Berlin, J. A. (1992). Poststructuralism, cultural studies, and the composition classroom: Postmodern theory in practice. *Rhetoric Review, 11*, 16–33.
<http://doi.org/10.1080/07350199209388984>
- Berlin, J. A., & Inkster, R. P. (1980). Current-traditional rhetoric: Paradigm and practice. *Freshman English News, 8*, 1–14.
- Biesta, G. (2010). Pragmatism and the philosophical foundations of mixed methods research. In A. Tashakkori & C. Teddlie (Eds.), *SAGE handbook of mixed methods in social and*

behavioral research (2nd., pp. 95–117). Thousand Oaks, CA: Sage.

Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2004). Working inside the black box: Assessment for Learning in the classroom. *Phi Delta Kappan*, *86*, 8–21.

<http://doi.org/10.1177/003172170408600105>

Black, P., & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, *80*(2), 139–144, 146–148.

Blom, J. N., & Hansen, K. R. (2015). Click bait: Forward-reference as lure in online news headlines. *Journal of Pragmatics*, *76*, 87–100. <http://doi.org/10.1016/j.pragma.2014.11.010>

Bohrnstedt, G. W., & Stancavage, F. B. (2015). *Maintaining the validity of the National Assessment of Educational Progress in a Common Core based environment*. San Mateo, CA. Retrieved from <http://www.air.org/sites/default/files/Validity-NAEP-Common-Core-Environment-March-2015.pdf>

Brady, L. A. (1993). A contextual theory for business writing. *Journal of Business and Technical Communication*, 452–471. <http://doi.org/0803973233>

Brandt, D. (2003). Changing literacy. *Teachers College Record*, *105*(2), 245–260.

<http://doi.org/10.1111/1467-9620.00238>

Brandt, D. (2005). Writing for a living: Literacy and the knowledge economy. *Written Communication*, *22*, 166–197. <http://doi.org/10.1177/0741088305275218>

<http://doi.org/10.1177/0741088305275218>

Breetvelt, I., van den Bergh, H., & Rjlaarsdam, G. (1994). Relations between writing processes and text quality: when and how? *Cognition and Instruction*, *12*, 103–123.

http://doi.org/10.1207/s1532690xci1202_2

Brent, D. (2011). Transfer, transformation, and rhetorical knowledge: Insights from transfer theory. *Journal of Business and Technical Communication*, *25*, 396–420.

<http://doi.org/10.1177/1050651911410951>

- Bronfenbrenner, U. (1994). Ecological models of human development. In *International Encyclopedia of Education* (2nd Editio, pp. 37–43). Elsevier.
- Bronfenbrenner, U., & Morris, P. A. (2006). The bioecological model of human development. In R. M. Lerner & W. Damon (Eds.), *Handbook of child psychology* (6th Editio, pp. 793–828). New York, NY: Wiley. <http://doi.org/10.1017/CBO9781107415324.004>
- Brookhart, S. M. (2009). Assessment, gender and in/equity. In C. Wyatt-Smith & J. J. Cumming (Eds.), *Educational assessment in the 21st century: Connecting theory and practice* (pp. 119–136). New York, NY: Springer. <http://doi.org/10.1007/978-1-4020-9964-9>
- Bruning, R. H., & Kauffman, D. F. (2016). Self-efficacy beliefs and motivation in writing development. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (2nd ed., pp. 160–193). New York: Guilford Press.
- Bryman, A. (1984). The debate about quantitative and qualitative research: A question of method or epistemology? *The British Journal of Sociology*, 35(1), 75–92.
- Bryson, M., Bereiter, C., Scardamalia, M., & Joram, E. (1991). Going beyond the problem as given: Problem solving in expert and novice writers. In R. J. Sternberg & P. A. Frensch (Eds.), *Complex problem solving: Principles and mechanisms* (pp. 61–84). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Bulger, M. E., Mayer, R. E., & Metzger, M. J. (2014). Knowledge and processes that predict proficiency in digital literacy. *Reading and Writing*, 27, 1567–1583.
<http://doi.org/10.1007/s11145-014-9507-2>
- Burdick, H., Swartz, C. W., Stenner, J. A., Fitzgerald, J., Burdick, D., & Hanlon, S. T. (2013). Measuring students' writing ability on a computer-analytic developmental scale: An

exploratory validity study. *Literacy Research and Instruction*, 52, 255–280.

<http://doi.org/10.1080/19388071.2013.812165>

- Camp, R. (2009). Changing the model for the direct assessment of writing. In B. Huot & P. O’Neill (Eds.), *Assessing writing: A critical sourcebook* (pp. 102–130). Urbana, Illinois: National Council of Teachers of English & Bedford/St. Martin’s.
- Canagarajah, A. S. (1996). “Nondiscursive” requirement in academic publishing, material resources of periphery scholars, and the politics of knowledg production. *Written Communication*, 13. <http://doi.org/10.1177/0741088396013004001>
- Chappuis, S., & Stiggins, R. J. (2002). Classroom assessment for learning. *Educational Leadership*, 60(1), 40–43.
- Chenoweth, N. A., & Hayes, J. R. (2001). Fluency in writing: Generating text in L1 and L2. *Written Communication*, 18, 80–98. <http://doi.org/10.1177/0741088301018001004>
- Christ, T. W. (2010). Teaching mixed methods and action research: Pedagogical, practical, and evaluative considerations. In *SAGE handbook of mixed methods in social & behavioral research* (2nd ed.). Los Angeles: Sage Publications.
- Cohen, L., Manion, L., & Morrison, K. (2013). *Research methods in education* (7th ed.). New York, NY: Routledge/Taylor & Francis Group.
- Coiro, J., Knobel, M., Lankshear, C., & Leu, D. J. (Eds.). (2008). *Handbook of research on New Literacies*. New York, NY: Routledge/Taylor & Francis Group.
- Coiro, J. L. (2015). Purposeful, critical, and flexible: Vital dimensions of online reading and learning. In R. J. Spiro, M. DeSchryver, M. Schira Hagerman, P. M. Morsink, & P. Thompson (Eds.), *Reading at a crossroads? Disjunctures and continuities in current conceptions and practices* (pp. 53–64). New York, NY: Routledge.

- Coiro, J. L., & Dobler, E. (2007). Exploring the online reading comprehension strategies used by sixth-grade skilled readers to search for and locate information on the Internet. *Reading Research Quarterly, 42*, 214–257. <http://doi.org/10.1598/RRQ.42.2.2>
- Coiro, J. L., Knobel, M., Lankshear, C., & Leu, D. J. (2008). Chapter 1: Central issues in new literacies and new literacies research. In *Handbook of Research on New Literacies* (pp. 1–21). Lawrence Erlbaum Associates, Taylor & Francis Group. Retrieved from <http://eprints.jcu.edu.au/7369/>
- Cole, N. (1997). *The ETS gender study: How females and males perform in educational settings*. Princeton, NJ: Educational Testing Service. Retrieved from https://archive.org/details/ERIC_ED424337
- Collins, A., Brown, J. S., & Holum, A. (1991). Cognitive apprenticeship: Making thinking visible. *American Educator, 15*(3), 6–11, 38–46.
- Collins, K. M. T., Onwuegbuzie, A. J., & Johnson, R. B. (2012). Securing a place at the table: A review and extension of legitimation criteria for the conduct of mixed research. *American Behavioral Scientist, 56*, 849–865. <http://doi.org/10.1177/0002764211433799>
- Common Core State Standards Initiative. (2010). *Common Core State Standards for English Language Arts*.
- Cope, B., & Kalantzis, M. (2009). “Multiliteracies”: New literacies, new learning. *Pedagogies: An International Journal, 4*, 164–195.
- Coppola, N. W., & Elliot, N. (2013). Conceptualizing the technical communication body of knowledge: Context, metaphor, and direction. *Technical Communication, 60*, 267–278. Retrieved from <http://www.scopus.com/inward/record.url?eid=2-s2.0-84891883479&partnerID=tZOtx3y1>

- Corrigan, J. A. (2014a). Writing assessment in a digital age: A two-state, large-scale study of multimodal composition. In *Writing Research Across Borders, International Society for the Advancement of Writing Research*. Paris, France: Université Paris-Ouest Nanterre La Défense.
- Corrigan, J. A. (2014b). Writing assessment in a digital age: A two-state, large-scale study of multimodal composition. In *Writing Research Across Borders*. Paris, France: International Society for the Advancement of Writing Research. Retrieved from http://conference2014.fr/wrab_resu_en.php
- Corrigan, J. A. (2017). Exploring the differences between student writers and knowledge workers in online research writing.
- Corrigan, J. A., & Slomp, D. H. (2017). New writing: Articulating a bioecological writing construct for the digital age.
- Creswell, J. W. (2007). *Qualitative inquiry & research design* (2nd.). Thousand Oaks, CA: Sage.
- Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches* (3rd ed.). Thousand Oaks, CA: Sage .
- Creswell, J. W. (2011). Controversies in mixed methods research. In N. K. Denzin & Y. S. Lincoln (Eds.), *The SAGE handbook of qualitative research* (4th, Kindl ed.). Thousand Oaks, CA: Sage.
- Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and conducting mixed methods research* (Kindle 2nd). Thousand Oaks, CA: Sage.
- Cronbach, L. J. (1989). Construct validation after thirty years. In R. L. Linn (Ed.), *Intelligence: Measurement, theory, and public policy* (pp. 147–171). Urbana, Illinois: University of Illinois Press.

- Cureton, E. E. (1951). Validity. In E. F. Lindquist (Ed.), *Educational Measurement* (pp. 621–694). Washington, DC: American Council on Education.
- Davenport, T. H. (2005). *Thinking for a living: How to get better performances and results from knowledge workers*. Boston, MA: Harvard Business School Press.
- Davis, B., Sumara, D., & Luce-Kapler, R. (2015). *Engaging minds - Cultures of education and practices of teaching* (3rd ed.). New York, NY: Routledge.
- Davis, D. S., Huang, B., & Yi, T. (2016). Making sense of science texts: A mixed-methods examination of predictors and processes of multiple-text comprehension. *Reading Research Quarterly*, 52, 227–252. <http://doi.org/10.1002/rrq.162>
- De La Paz, S., & Graham, S. (1997). Strategy instruction in planning: Effects on the writing performance and behavior of students with learning disabilities. *Exceptional Children*, 63, 167–181.
- de Milliano, I., van Gerlderen, A., & Slegers, P. (2012). Patterns of cognitive self-regulation of adolescent struggling writers. *Written Communication*, 29, 303–325. <http://doi.org/10.1177/0741088312450275>
- Deane, P. (2015). Developing and validating assessments focusing on the writing process (The CBAL key practice, writing, reviewing and publishing documents). Princeton, NJ: Educational Testing Service.
- Deleuze, G., & Guattari, F. (1983). *Anti-Oedipus: Capitalism and schizophrenia*. (M. Seem & H. R. Lane, Eds.) *Humanities*. Minneapolis: University of Minnesota Press. Retrieved from <http://www.amazon.com/Anti-Oedipus-Capitalism-Schizophrenia-Gilles-Deleuze/dp/0816612250>
- Denzin, N. K. (2008). The new paradigm dialogues and qualitative inquiry. *International Journal*

of Qualitative Studies in Education, 21, 315–325.

Denzin, N. K., & Lincoln, Y. S. (2011). The discipline and practice of qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *The SAGE handbook of qualitative research* (4th, Kindl ed.). Thousand Oaks, CA: Sage Publications Inc.

Derrida, J., & Ronell, A. (2009). The law of genre. *Critical Inquiry*, 7, 55–81.

<http://doi.org/10.1086/448088>

Deschryver, M. (2014). Higher order thinking in an online world: Toward a theory of web-mediated knowledge synthesis. *Teachers College Record*, 116(12), 1–44. Retrieved from <http://www.tcrecord.org/Content.asp?ContentId=17692>

Dewey, J., & Bentley, A. F. (2008). Knowing and the known. In J. A. Boydston (Ed.), *John Dewey, The Later Works, Volume 16, 1925-1953*. Carbondale: Southern Illinois UP.

Donkin, K. (2013, January 11). Social media helps drive Idle No More movement. *The Star*.

Toronto. Retrieved from

http://www.thestar.com/news/canada/2013/01/11/social_media_helps_drive_idle_no_more_movement.html

Downs, D., & Wardle, E. (2007). Teaching about writing, righting misconceptions: (Re)envisioning “First-Year Composition” as “Introduction to Writing Studies.” *College Composition and Communication*, 58, 552–584. <http://doi.org/10.2307/20456966>

Drijbooms, E., Groen, M. A., & Verhoeven, L. (2015). The contribution of executive functions to narrative writing in fourth grade children. *Reading and Writing*, 28, 989–1011.

<http://doi.org/10.1007/s11145-015-9558-z>

Earl, L. (2003). *Assessment as learning: Using classroom assessment to maximize student learning*. Thousand Oaks, CA: Corwin Press.

- Education Quality and Accountability Office. (2007). *Framework - Ontario Secondary School Literacy Test (EQAO)*. Toronto, Canada. Retrieved from <http://www.eqao.com/en/assessments/OSSLT/assessment-docs/framework-osslt.pdf>
- Education Quality and Accountability Office. (2016). *EQAO's technical report*. Toronto, Canada. Retrieved from <http://www.eqao.com/en/assessments/DMA-docs/technical-report-2015-2016.pdf>
- Ellingson, L. L. (2011). Analysis and representation across the continuum. In N. K. Denzin & Y. S. Lincoln (Eds.), *The SAGE handbook of qualitative research* (4th, Kindl ed.). Thousand Oaks, CA: Sage Publications Inc.
- Ellis, A. K., Denton, D. W., & Bond, J. B. (2014). An analysis of research on metacognitive teaching strategies. *Procedia - Social and Behavioral Sciences*, *116*, 4015–4024. <http://doi.org/10.1016/j.sbspro.2014.01.883>
- Emig, J. (1971). *The composing processes of twelfth graders*. Urbana, Illinois.
- Ercikan, K., & Roth, W.-M. (2006). What good is polarizing research into qualitative and quantitative? *Educational Researcher*, *35*(5), 14–23. <http://doi.org/10.3102/0013189X035005014>
- Flavell, J. H. (1979). Metacognition and cognitive monitoring. *American Psychologist*, *34*, 906–911. <http://doi.org/10.1037/0003-066X.34.10.906>
- Flower, L. (1989). Cognition, context, and theory building. *College Composition and Communication*, *40*, 282–311.
- Flower, L., & Hayes, J. R. (1980). The cognition of discovery: Defining a rhetorical problem. *College Composition and Communication*, *31*, 21–32.
- Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. *College Composition*

and Communication, 32, 365–387. <http://doi.org/10.2307/356600>

Forzani, E. (2016). *Individual differences in evaluating the credibility of online information in science: Contributions of prior knowledge, gender, socioeconomic status, and offline reading ability*. University of Connecticut. Retrieved from

<http://digitalcommons.uconn.edu/cgi/viewcontent.cgi?article=7457&context=dissertations>

Gareis, C. R., & Grant, L. W. (2015). *Teacher-made assessments: How to connect curriculum, instruction, and student learning* (2nd ed.). New York, NY: Routledge.

Gee, J. P. (1989). Literacy, discourse, and linguistics: Introduction. *Journal of Education*, 171, 5–176.

Gee, J. P. (1990a). *Social linguistics and literacies: Ideology in Discourses, critical perspectives on literacy and education*. Bristol, PA: Falmer Press.

Gee, J. P. (1990b). *Social linguistics and literacies: ideology in discourses*. New York: Falmer Press.

Gee, J. P. (1999). *The New Literacy Studies and the “social turn.”*

Gee, J. P. (2014). Decontextualized language: A problem, not a solution. *International Multilingual Research Journal*, 8, 9–23. <http://doi.org/10.1080/19313152.2014.852424>

Geisler, C. (1994). Literacy and expertise in the academy. *Language and Learning across the Disciplines*, 1, 35–57.

Gilster, P. (1997). *Digital literacy*. New York: John Wiley & Sons Inc.

Goodwin, L. D., & Leech, N. L. (2003). The meaning of validity in the new Standards for Educational and Psychological Testing: Implications for measurement courses.

Measurement and Evaluation in Counseling and Development, 36, 181–191.

Graham, S., Gillespie, A., & McKeown, D. (2013). Writing: Importance, development, and

- instruction. *Reading and Writing*, 26, 1–15. <http://doi.org/10.1007/s11145-012-9395-2>
- Graham, S., & Harris, K. R. (2013). Designing an effective writing program. In S. Graham, C. A. MacArthur, & J. Fitzgerald (Eds.), *Best practices in writing instruction* (2nd ed.). New York, NY: Guilford Press.
- Graham, S., & Sandmel, K. (2011). The process writing approach: A meta-analysis. *The Journal of Educational Research*, 104, 396–407. <http://doi.org/10.1080/00220671.2010.488703>
- Graves, R., & Hyland, T. (Eds.). (2017). *Writing assignments across the undergraduate disciplines: An impetus for change*. Edmonton, Canada: Inkshed Publications.
- Greene, J. C. (2007). *Mixed methods in social inquiry*. San Francisco, CA: Jossey-Bass.
- Greene, J. C., & Hall, J. N. (2010). Dialectics and pragmatism: Being of consequence. In A. Tashakorri & C. Teddlie (Eds.), *SAGE handbook of mixed methods in social and behavioral research* (2nd., pp. 119–144). Thousand Oaks, CA: Sage.
- Guba, E. G. (1987). What have we learned about naturalistic evaluation? *Evaluation Practice*, 8(1), 23–43.
- Guba, E. G., & Lincoln, Y. S. (2009). Paradigmatic controversies, contradictions, and emerging confluences. In N. K. Denzin & Y. S. Lincoln (Eds.), *The Sage handbook of qualitative research* (3rd., pp. 191–215). Thousand Oaks, CA: Sage Publications Inc.
- Gutierrez, K. D. (2016). 2011 AERA presidential address: Designing resilient ecologies: Social design experiments and a new social imagination. *Educational Researcher*, 45, 187–196. <http://doi.org/10.3102/0013189X16645430>
- Gutiérrez, K. D., Baquedano-López, P., & Tejada, C. (1999). Rethinking diversity: Hybridity and hybrid language practices in the third space. *Mind, Culture, and Activity*, 6, 286–303. <http://doi.org/10.1080/10749039909524733>

- Hacker, D. J., Keener, M. C., & Kircher, J. C. (2009). Writing is applied metacognition. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of metacognition in education* (Kindle). New York: Taylor & Francis.
- Halliday, M. A. K. (2003). *On languages and linguistics*. (J. Webster, Ed.). London: Continuum.
- Hamp-Lyons, L. (1997). Washback, impact and validity: ethical concerns. *Language Testing*, 14, 295–303. <http://doi.org/10.1177/026553229701400306>
- Hanson, F. A. (2000). How tests create what they are intended to measure. In *Assessment: Social practice and social product* (pp. 67–81). London, UK: Routledge Falmer.
- Hayes, J. R. (1996). A new framework for understanding cognition and affect in writing. In C. M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences, and applications* (pp. 1–27). Mahwah, NJ: Lawrence Erlbaum Associates.
- Hayes, J. R. (2012). Modeling and remodeling writing. *Written Communication*, 29, 369–388. <http://doi.org/10.1177/0741088312451260>
- Hayes, J. R., & Flower, L. S. (1980). Identifying the organization of writing process. In L. W. Gregg & E. R. Steinberg (Eds.), *Cognitive processes in writing: An interdisciplinary approach* (pp. 3–30). Hillsdale, NJ: Erlbaum.
- Herrington, A., Hodgson, K., & Moran, C. (Eds.). (2009). *Teaching the new writing: Technology, change, and assessment in the 21st-century classroom*. New York: Teachers College Press and National Writing Project.
- Hesse-Biber, S., & Johnson, R. B. (2013). Coming at things differently: Future directions of possible engagement with mixed methods research. *Journal of Mixed Methods Research*, 7(2), 103–109. <http://doi.org/10.1177/1558689813483987>
- Hicks, T., Russo, A., Autrey, T., Gardner, R., Kabodian, A., & Edington, C. (2007). Rethinking

- the purposes and processes for designing digital portfolios. *Journal of Adolescent & Adult Literacy*, 50, 450–458. <http://doi.org/10.2307/40015497>
- Hillocks, G. J. (2002). *The testing trap: How state writing assessments control learning*. New York: Teachers College Press.
- Holmqvist, K., Nystrom, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures* (Kindle). New York, NY: Oxford University Press.
- Hopfenbeck, T. N. (2016). The power of PISA – limitations and possibilities for educational research. *Assessment in Education: Principles, Policy & Practice*, 23, 423–426. <http://doi.org/10.1080/0969594X.2016.1247518>
- Howe, K. R. (1988). Against the quantitative-qualitative incompatibility thesis or dogmas die hard. *Educational Researcher*, 17(8), 10. <http://doi.org/10.2307/1175845>
- Hudson, R. (2016). Grammar instruction. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (2nd ed., pp. 288–300). New York, NY: Guilford Press.
- Hull, G. A., & Katz, M.-L. (2006). Crafting an agentive self: Case studies of digital storytelling. *Research in the Teaching of English*, 41, 43–81.
- Hull, G. A., & Nelson, M. E. (2005). Locating the semiotic power of multimodality. *Written Communication*, 22, 224–261. <http://doi.org/10.1177/0741088304274170>
- Huot, B. (1990). The literature of direct writing assessment: Major concerns and prevailing trends. *Review of Educational Research*, 60, 237. <http://doi.org/10.2307/1170611>
- Inoue, A. B. (2009). The technology of writing assessment and racial validity. In C. Schreiner (Ed.), *Handbook of research on assessment technologies, methods, and applications in higher education*. Hershey, PA: Information Science Reference. <http://doi.org/10.4018/978->

1-60566-667-9.ch006

- Ivanic, R. (2004). Discourses of writing and learning to write. *Language and Education, 18*, 220–245. <http://doi.org/10.1080/09500780408666877>
- Jeffery, J. V. (2009). Constructs of writing proficiency in US state and national writing assessments: Exploring variability. *Assessing Writing, 14*, 3–24. <http://doi.org/10.1016/j.asw.2008.12.002>
- Jewitt, C. (2008). Multimodality and literacy in school classrooms. *Review of Research in Education, 32*(1), 241–267. <http://doi.org/10.3102/0091732X07310586>
- Johnson, R. B., & Gray, R. (2010). A history of philosophical and theoretical issues for mixed methods research. In A. Tashakkori & C. Teddlie (Eds.), *SAGE handbook of mixed methods in social and behavioral research* (Second, pp. 69–94). Thousand Oaks, CA: Sage.
- Johnson, R. B., Onwuegbuzie, A. J., & Turner, L. A. (2007). Toward a definition of mixed methods research. *Journal of Mixed Methods Research, 1*(2), 112–133. Retrieved from <http://mmr.sagepub.com/content/1/2/112.short>
- Jones, S., Myhill, D., Bailey, T., & Bailey, T. (2013). Grammar for writing? An investigation of the effects of contextualised grammar teaching on students' writing. *Reading and Writing, 26*, 1241–1263. <http://doi.org/10.1007/s11145-012-9416-1>
- Journal Citation Reports*®. (2015).
- Kalantzis, M., Cope, B., & Harvey, A. (2003). Assessing multiliteracies and the new basics. *Assessment in Education, 10*(1), 15–26. <http://doi.org/10.1080/0969594032000085721>
- Kamil, M. L. (2015). Past, present, and future conditions and practices of reading. In R. J. Spiro, M. DeSchryver, M. Schira Hagerman, P. M. Morsink, & P. Thompson (Eds.), *Reading at a crossroads? Disjunctures and continuities in current conceptions and practices* (pp. 139–

- 147). New York, NY: Routledge.
- Kane, M. T. (2006). Validation. In R. L. Brennan (Ed.), *Educational measurement* (4th ed., pp. 17–63). Westport, CT: ACE/Praeger Series on Higher Education.
- Kane, M. T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement, 50*, 1–73. <http://doi.org/10.1111/jedm.12000>
- Kellogg, R. T. (2008). Training writing skills: A cognitive developmental perspective. *Journal of Writing Research, 1*, 1–26. Retrieved from <http://dx.doi.org/10.17239/jowr-2008.01.01.1>
- Kennedy, C., Rhoads, C., & Leu, D. J. (2016). Online research and learning in science: A one-to-one laptop comparison in two states using performance based assessments. *Computers and Education, 100*, 141–161. <http://doi.org/10.1016/j.compedu.2016.05.003>
- Kennedy, M. L. (1985). The composing process of college students writing from sources. *Written Communication, 2*, 434–456. <http://doi.org/10.1177/0741088385002004006>
- Kent, S. C., & Wanzek, J. (2016a). The relationship between component skills and writing quality and production across developmental levels: A meta-analysis of the last 25 years. *Review of Educational Research, 86*, 570–601. <http://doi.org/10.3102/0034654315619491>
- Kent, S. C., & Wanzek, J. (2016b). The relationship between component skills and writing quality and production across developmental levels: A meta-analysis of the last 25 years. *Review of Educational Research, 86*, 570–601. <http://doi.org/10.3102/0034654315619491>
- Kiili, C., Laurinen, L., & Marttunen, M. (2008). Students evaluating Internet sources: From versatile evaluators to uncritical readers. *Journal of Educational Computing Research, 39*, 75–95. <http://doi.org/10.2190/EC.39.1.e>
- Kinloch, V. (2010). *Harlem on our minds: place, race, and the literacies of urban youth*. New York: Teachers College Press.

- Kirkpatrick, L. C., & Klein, P. D. (2016). High-achieving high school students' strategies for writing from internet-based sources of information. *Journal of Writing Research, 8*, 1–46. <http://doi.org/10.17239/jowr-2016.08.01.01>
- Kress, G. (2003). *Literacy in the new media age*. New York: Routledge/Taylor & Francis Group.
- Kress, G., & van Leeuwen, T. (1996). *Reading images: The grammar of visual design*. New York: Routledge/Taylor & Francis Group.
- Lankshear, C., & Knobel, M. (2007). Researching New Literacies: Web 2.0 practices and insider perspectives. *E-Learning, 4*(3), 224–240. <http://doi.org/10.2304/elea.2007.4.3.224>
- Lankshear, C., & Knobel, M. (2008). *Digital literacies: Concepts, policies and practices*. (C. Lankshear & M. Knobel, Eds.)*Names*. New York: Peter Lang Pub Inc.
- Lankshear, C., & Knobel, M. (2011). *New Literacies: Everyday practices and social learning* (3rd ed.). New York: Open University Press.
- Lau, A. (2013). Timed writing assessment as a measure of writing ability: A qualitative study. *Discussions, 9*(2), 1–6. Retrieved from <http://www.inquiriesjournal.com/articles/798/timed-writing-assessment-as-a-measure-of-writing-ability-a-qualitative-study>
- Law, J. (2004). *After method: Mess in social science research*. New York, NY: Routledge/Taylor & Francis Group.
- Lea, M. R., & Street, B. V. (2006). The “academic literacies” model: Theory and applications. *Theory Into Practice, 45*, 368–377. http://doi.org/10.1207/s15430421tip4504_11
- Leijten, M., Van Waes, L., Schriver, K., & Hayes, J. R. (2014). Writing in the workplace: Constructing documents using multiple digital sources. *Journal of Writing Research, 5*, 285–337. <http://doi.org/10.17239/jowr-2014.05.03.3>
- Leijten, M., & Waes, L. Van. (2000). Inputlog: A multimethod approach describing cognitive

writing processes using keystroke logging. Retrieved from

http://www.inputlog.net/docs/2011_manuscript_submitted.pdf

- Leu, D. J., Coiro, J. L., Kulikowich, J., & Cui, W. (2012). Using the psychometric characteristics Of Multiple-Choice, Open Internet, and Closed (simulated) Internet formats to refine the development of Online Research and Comprehension Assessments in science: Year three of the ORCA Project. In *Literacy Research Association Conference*. San Diego, CA, USA.
- Leu, D. J., Forzani, E., Burlingame, C., Kulikowich, J., Sedransk, N., Coiro, J. L., & Kennedy, C. (2013). The new literacies of online research and comprehension: Assessing and preparing students for the 21st century with Common Core State Standards. In S. B. Neuman & L. B. Gambrell (Eds.), *Quality reading instruction in the age of Common Core Standards* (pp. 219–236). Newark, DE: International Reading Association. Retrieved from http://massreading.org/wp-content/uploads/2014/03/16_496_CH16_IRA_Neuman-clean.pdf
- Leu, D. J., Forzani, E., Rhoads, C., Maykel, C., Kennedy, C., & Timbrell, N. (2015). The new literacies of online research and comprehension: Rethinking the reading achievement gap. *Reading Research Quarterly*, *50*(1), 37–59. <http://doi.org/10.1002/rrq.85>
- Leu, D. J., Kinzer, C. K., Coiro, J. L., & Cammack, D. W. (2004). Toward a theory of New Literacies emerging from the Internet and other Information and Communication Technologies. In *Theoretical models and processes of reading. (5th ed.)* (5th ed., pp. 1570–1613). International Reading Association.
- Leu, D. J., Kinzer, C. K., Coiro, J. L., Castek, J., & Henry, L. A. (2013). New Literacies: A dual-level theory of the changing nature of literacy, instruction , and assessment. In D. E. Alvermann, N. J. Unrau, & R. B. Ruddell (Eds.), *Theoretical models and processes of reading. (6th ed.)* (pp. 1150–1181). International Reading Association.

- Leu, D. J., Kulikowich, J. M., Sedransk, N., Coiro, J. L., Liu, C., Cui, W., ... Maykel, C. (2014). The ORCA Project: Designing technology-based assessments for online research, comprehension, and communication. In *American Educational Research Association* (pp. 1–60). Philadelphia, PA.
- Leu, D. J., Kulikowich, J., Sedransk, N., & Coiro, J. L. (n.d.). *Assessing online reading comprehension: The ORCA Project*. Washington, DC.
- Leu, D. J., & New Literacies Research Lab. (2012). *Sample report - Results: Seventh grade students' ability to conduct research online and write a short report, a Common Core Anchor Standard*. Storrs, CT.
- Leu, D. J., Slomp, D. H., Zawilinski, L., & Corrigan, J. A. (2016). Writing research through a New Literacies lens. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (2nd ed., pp. 41–53). New York: Guilford Press.
- Leu, D. J., Zawilinski, L., Castek, J., Banerjee, M., Housand, B., Liu, Y., & O'Neil, M. (2007). What is new about the new literacies of online reading comprehension. In A. Berger, L. Rush, & J. Eakle (Eds.), *Secondary school reading and writing: What research reveals for classroom practices*. (pp. 37–68). Chicago, IL: NCTE/NCRL. Retrieved from http://www.newliteracies.uconn.edu/pub_files/What_is_new_about_new_literacies_of_online_reading.pdf
- Levy, Y., & Ellis, T. J. (2006). A systems approach to conduct an effective literature review in support of Information Systems Research. *Informing Science Journal*, 9, 181–212.
- Lincoln, Y. S., Lynham, S. A., & Guba, E. G. (2011). Paradigmatic controversies, contradictions, and emerging confluences, revisited. In N. K. Denzin & Y. S. Lincoln (Eds.), *The SAGE handbook of qualitative research* (4th, Kindl ed.). Thousand Oaks, CA: Sage Publications

Inc.

- Luke, A. (2012). Critical literacy: Foundational notes. *Theory Into Practice, 51*, 4–11.
<http://doi.org/10.1080/00405841.2012.636324>
- Lyons, S. R. (2000). Rhetorical sovereignty: What do American Indians want from writing?
College Composition and Communication, 51, 447–468.
<http://doi.org/doi.org/10.2307/358744>
- MacArthur, C. A., & Graham, S. (2016). Writing research from a cognitive perspective. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (2nd ed., pp. 24–40). New York: Guilford Press.
- Madden, M., & Jones, S. (2008). *Wired workers: Who are they, what they're doing online*. Washington, DC. Retrieved from <http://www.pewinternet.org/2008/09/24/networked-workers/>
- Mailloux, S. (1989). *Rhetorical power*. Ithaca, NY: Cornell University Press.
- Marttunen, M., & Laurinen, L. (2012). Participant profiles during collaborative writing. *Journal of Writing Research, 4*, 53–79. <http://doi.org/10.17239/jowr-2012.04.01.3>
- Masny, D. (2011). Multiple Literacies Theory: Exploring futures. *Policy Futures in Education, 9*, 494–504. <http://doi.org/10.2304/pfie.2011.9.4.494>
- Mateos, M., Martín, E., Villalón, R., & Luna, M. (2008). Reading and writing to learn in secondary education: Online processing activity and written products in summarizing and synthesizing tasks. *Reading and Writing, 21*, 675–697. <http://doi.org/10.1007/s11145-007-9086-6>
- Maxwell, J. A., & Mittapalli, K. (2010). Realism as a stance for mixed methods research. In A. Tashakkori & C. Teddlie (Eds.), *SAGE handbook of mixed methods in social and*

- behavioral research* (2nd., pp. 45–168). Thousand Oaks, CA: Sage.
- McDonald, B., & Boud, D. (2003). The impact of self-assessment on achievement: The effects of self-assessment training on performance in external examinations. *Assessment in Education: Principles, Policy & Practice*, *10*, 209–220. <http://doi.org/10.1080/0969594032000121289>
- McKoski, M. M. (1995). A legacy of developmental writing. *Journal of Developmental Education*, *19*(2), 8–10.
- McLuhan, M. (1964). The medium is the message. In *Understanding media: The extensions of man* (pp. 3–21). New York: McGraw-Hill.
- Mertens, D. M. (2012). What comes first? The paradigm or the approach? *Journal of Mixed Methods Research*, *6*(4), 255–257. <http://doi.org/10.1177/1558689812461574>
- Messenger, W. E., de Bruyn, J., Brown, J., & Montagnes, R. (2004). *The Canadian writer's handbook* (4th ed.). Don Mills, ON: Oxford University Press.
- Messick, S. (1989). Meaning and values in test validation: The science and ethics of assessment. *Educational Researcher*, *18*(2), 5–11. <http://doi.org/10.3102/0013189X018002005>
- Messick, S. (1992). Validity of test interpretation and use. In M. C. Alkin (Ed.), *Encyclopedia of educational research* (6th ed., pp. 1487–1495). New York: MacMillan.
- Messick, S. (1996). Validity and washback in language testing. *Language Testing*, *13*, 241–256. <http://doi.org/10.1177/026553229601300302>
- Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). *Qualitative data analysis: A methods sourcebook* (3rd ed.). Thousand Oaks, CA: Sage Publications Inc.
- Miller, C. R. (1984). Genre as social action. *Quarterly Journal of Speech*, *70*, 151–167. <http://doi.org/10.1080/00335638409383686>
- Mills, K. A. (2010). A review of the “Digital Turn” in the New Literacy Studies. *Review of*

- Educational Research*, 80, 246–271. <http://doi.org/10.3102/0034654310364401>
- Mills, K. A. (2013). Multimodal and monomodal discourses of marketization in higher education: Power, ideology, and the absence of the image. In *Proceedings of the American Educational Research Association*. San Francisco, CA.
- Mitchell, R., & Taylor, M. (1979). The integrating perspective: An audience-response model for writing. *College English*, 41, 247–271. <http://doi.org/doi.org/10.2307/376441>
- Moore, P. (1997). Rhetorical vs . instrumental approaches to teaching technical communication. *Technical Communication*, 44, 163–173.
- Morgan, D. L. (2007). Paradigms lost and pragmatism regained: Methodological implications of combining qualitative and quantitative methods. *Journal of Mixed Methods Research*, 1(1), 48–76. <http://doi.org/10.1177/2345678906292462>
- Morrow, K. (1986). The evaluation of tests of communicative performance. In M. Portal (Ed.), *Innovations in language testing*. London, UK: NFER/Nelson.
- Moss, P. A. (1994). Can there be validity without reliability? *Educational Researcher*, 23(2), 5–12. <http://doi.org/10.3102/0013189X023002005>
- Myers, J., & Beach, R. (2001). Hypermedia authoring as critical literacy. *Journal of Adolescent & Adult Literacy*, 44(6), 233–246.
- Nagy, P. (2000). The three roles of assessment: Gatekeeping, accountability, and instructional diagnosis. *Canadian Journal of Education / Revue Canadienne de l'éducation*, 25(4), 262–279. <http://doi.org/10.2307/1585850>
- National Assessment Governing Board; U.S. Department of Education. (2010). *Writing framework for the 2011 National Assessment of Educational Progress*. Washington, DC.
- National Center on Educational Outcomes. (2002). *Universally designed assessments: Better*

- assessments for everyone! Policy Directions, number 14.* Minnesota. Retrieved from <https://nceo.umn.edu/docs/OnlinePubs/PolicyDirections14.pdf>
- National Writing Project. (2010). *Because digital writing matters: Improving student writing in online and multimedia environments.* (D. N. DeVoss, E. Eidman-Aadahl, & T. Hicks, Eds.). San Francisco, CA: Jossey-Bass.
- Neuman, S. B., & Celano, D. (2012). *Giving our children a fighting chance: poverty, literacy, and the development of information capital.* New York: Teachers College Press.
- New London Group. (1996). A pedagogy of multiliteracies: Designing social futures. *Harvard Educational Review, 66*, 60–92. <http://doi.org/10.17763/haer.66.1.17370n67v22j160u>
- Newkirk, T. (2000). Misreading masculinity: Speculations on the great gender gap in writing. *Language Arts, 77*(4), 294–300 CR–Copyright © 2000 National Council. <http://doi.org/10.2307/41483068>
- Niglas, K. (2010). The multidimensional model of research methodology: An integrated set of continua. In A. Tashakkori & C. Teddlie (Eds.), *SAGE handbook of mixed methods in social and behavioral research* (2nd., pp. 215–236). Thousand Oaks, CA: Sage.
- O’Hara, K. P., Taylor, A., Newman, W., & Sellen, A. J. (2002). Understanding the materiality of writing from multiple sources. *Int. J. Human-Computer Studies, 56*, 269–305. <http://doi.org/10.1006/ijhc.2001.0525>
- OECD. (2011). *PISA 2009 results : Students on line: Digital technologies and performance* (Vol. Vi). Retrieved from <http://dx.doi.org/10.1787/9789264112995-en>
- Olinghouse, N. G., Graham, S., & Gillespie, A. (2015). The relationship of discourse and topic knowledge to fifth graders’ writing performance. *Journal of Educational Psychology, 107*, 391–406. <http://doi.org/10.1037/a0037549>

- Olive, T., & Passerault, J.-M. (2012). The Visuospatial Dimension of Writing. *Written Communication, 29*(3), 326–344. <http://doi.org/10.1177/0741088312451111>
- Ontario Ministry of Education. (2007). *The Ontario curriculum, English (Grades 9 and 10)*. Toronto, Canada: Queen's Printer for Ontario.
- Ontario Ministry of Education. (2016). *21st century competencies: Foundation document for discussion. Phase 1. Towards defining 21st century competencies for Ontario*. Queen's Printer for Ontario.
- Ontario Ministry of Education. (2017). Innovation Learning Fund. Retrieved April 6, 2018, from http://edugains.ca/newsite/21stCenturyLearning/innovation_learning_funding.html
- Onwuegbuzie, A. J. (2003). Effect sizes in qualitative research: A prolegomenon. *Quality & Quantity, 37*, 393–409. <http://doi.org/10.1023/A:1027379223537>
- Onwuegbuzie, A. J. (2012). Putting the MIXED back into quantitative and qualitative research in educational research and beyond: Moving toward the radical middle. *International Journal of Multiple Research Approaches, 6*(3), 192–219. <http://doi.org/10.1007/s00267-012-9864-z>
- Onwuegbuzie, A. J., & Frels, R. K. (2013). Towards a new research philosophy for addressing social justice issues: critical dialectical pluralism 1.0. *International Journal of Multiple Research Approaches, 7*(1), 9–26.
- Onwuegbuzie, A. J., & Frels, R. K. (2016). *Seven steps to a comprehensive literature review: A multimodal and cultural approach* (Kindle). London, England: Sage.
- Onwuegbuzie, A. J., Johnson, R. B., & Collins, K. M. T. (2009). Call for mixed analysis: A philosophical framework for combining qualitative and quantitative approaches. *International Journal of Multiple Research Approaches, 3*(2), 114–139.
- Onwuegbuzie, A. J., & Leech, N. L. (2005). On becoming a pragmatic researcher: The

importance of combining quantitative and qualitative research methodologies. *International Journal of Social Research Methodology*, 8(5), 375–387.

<http://doi.org/10.1080/13645570500402447>

Organisation for Economic Co-operation and Development (OECD). (2003). *Key competencies for a successful life and well-functioning society. The definition and selection of key competencies: Executive summary*. Göttingen: Hogrefe and Huber Publishers.

Pajares, F., & Valiante, G. (2001). Gender differences in writing motivation and achievement of middle school students: A function of gender orientation? *Contemporary Educational Psychology*, 26(3), 366–381. <http://doi.org/10.1006/ceps.2000.1069>

Pan, B., Hembrooke, H., Joachims, T., Lorigo, L., Gay, G., & Granka, L. (2007). In Google we trust: Users' decisions on rank, position, and relevance. *Journal of Computer-Mediated Communication*, 12(3), 801–823. <http://doi.org/10.1111/j.1083-6101.2007.00351.x>

Pariser, E. (2011). *The filter bubble*. New York, NY: Penguin Press.

Parsons, S. A., & Gallagher, M. A. (2016). A content analysis of nine literacy journals, 2009–2014. *Journal of Literacy Research*, 48, 476–502.
<http://doi.org/10.1177/1086296X16680053>

Patel, V. L., & Groen, G. J. (1991). The general and specific nature of medical expertise: A critical look. In K. A. Ericsson & J. Smith (Eds.), *Toward a general theory of expertise: Prospects and limits* (pp. 93–125). Cambridge, UK: Cambridge University Press.

Pea, R. D. (1993). Practices of distributed intelligence and designs for education. In *Distributed cognitions: Psychological and educational considerations* (pp. 47–87). Cambridge, UK: Cambridge University Press.

Philips, D. C. (2011). Muddying the waters even more. In C. F. Conrad & R. C. Serlin (Eds.),

- The SAGE handbook for research in education* (2nd ed., pp. 13–28). Thousand Oaks, CA: Sage Publications Inc.
- Pirolli, P. (2007). *Information foraging theory: Adaptive interaction with information* (Vol. 20). NY, New York: Oxford University Press.
- Powell, W. W., & Snellman, K. (2004). The Knowledge Economy. *Annual Review of Sociology*, 30(1), 199–220. <http://doi.org/10.1146/annurev.soc.29.010202.100037>
- Prior, P. (2006). A sociocultural theory of writing. In C. A. MacArthur, S. Graham, & J. FitzGerald (Eds.), *Handbook of writing research* (Kindle, pp. 54–66). New York, NY: The Guilford Press.
- Proske, A., & Kapp, F. (2013). Fostering topic knowledge: Essential for academic writing. *Reading and Writing*, 26, 1337–1352. <http://doi.org/10.1007/s11145-012-9421-4>
- Pullman, G. (2002). Electronic portfolios revisited: The efolios project. *Computers and Composition*, 19(2), 151–169. [http://doi.org/10.1016/S8755-4615\(02\)00109-3](http://doi.org/10.1016/S8755-4615(02)00109-3)
- Rapp, C. (2010). Aristotle's Rhetoric. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Spring 201). Department of Philosophy, Stanford University. Retrieved from <http://plato.stanford.edu/archives/spr2010/entries/aristotle-rhetoric/>
- Reinhardt, W., Schmidt, B., Sloep, P., & Drachsler, H. (2011). Knowledge worker roles and actions — Results of two empirical studies. *Knowledge and Process Management*, 18(3), 150–174. <http://doi.org/10.1002/kpm>
- Reinking, D., & Colwell, J. (2015). A brief history of information sources in the late 20th and early 21st centuries (a simulation). In R. J. Spiro, M. DeSchryver, M. Schira Hagerman, P. M. Morsink, & P. Thompson (Eds.), *Reading at a crossroads? Disjunctures and continuities in current conceptions and practices* (pp. 3–21). New York, NY: Routledge.

- Rice, J. (2015). Para-expertise, tacit knowledge, and writing problems. *College English*, 78, 117–138.
- Rodriguez, V. (2012). The teaching brain and the end of the empty vessel. *Mind, Brain, and Education*, 6, 177–185. <http://doi.org/10.1111/j.1751-228X.2012.01155.x>
- Ronald, K., & Volkmer, J. (1989). Another competing theory of process: The student's. *Journal of Advanced Composition*, 9, 83–96. Retrieved from <http://www.jstor.org/stable/20865673>
- Rose, D. (2016). New developments in genre-based literacy pedagogy. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (2nd ed., pp. 227–242). New York: Guilford Press.
- Rosenblatt, L. M. (1988). *Writing and reading: The transactional theory. Technical Report No. 416*. Champaign, Illinois. Retrieved from https://www.ideals.illinois.edu/bitstream/handle/2142/18044/ctrstreadtechrepv01988i00416_opt.pdf?s..
- Ross, J. A. (2006). The reliability, validity, and utility of self-assessment. *Practical Assessment, Research & Evaluation*, 11(10), 1–13.
- Sainsbury, M. (2009). Developing writing in a high-stakes environment. In R. Beard, D. Myhill, J. Riley, & M. Nystrand (Eds.), *The SAGE handbook of writing development* (pp. 545–560). Los Angeles.
- Saldaña, J. (2013). *The coding manual for qualitative researchers* (2nd ed.). London, UK: Sage Publications Inc.
- Salibrici, M. M. (1999). Dissonance and rhetorical Inquiry: A Burkean model for critical reading and writing. *Journal of Adolescent & Adult Literacy*, 42, 628–637.
- Sandelowski, M. (2008). Theoretical saturation. In L. M. Given (Ed.), *The Sage encyclopedia of*

- qualitative methods* (Vol. 1, pp. 875–876). Sage.
- Sandelowski, M., Volis, C. I., & Knafl, G. (2009). On quantizing. *Journal of Mixed Methods Research, 3*, 208–222. <http://doi.org/10.1177/1558689809334210>
- Schlagal, B. (2013). Best practices in spelling and handwriting. In S. Graham, C. A. MacArthur, & J. Fitzgerald (Eds.), *Best practices in writing instruction* (pp. 257–283). New York, NY: The Guilford Press.
- Schoen, L., & Fusarelli, L. D. (2008). Innovation, NCLB, and the fear factor: The challenge of leading 21st-century schools in an era of accountability. *Educational Policy, 22*, 181–203. <http://doi.org/10.1177/0895904807311291>
- Schrivver, K. (2012). What we know about expertise in professional communication. In V. W. Berninger (Ed.), *Past, present, and future contributions of cognitive writing* (pp. 275–312). New York, NY: Psychology Press.
- Schultz, K., & Fecho, B. (2000). Society's child: Social context and writing development. *Educational Psychologist, 35*, 51–62. http://doi.org/10.1207/S15326985EP3501_6
- Schwartz, D. L., Bransford, J. D., & Sears, D. (2005). Efficiency and innovation in transfer. In J. P. Mestre (Ed.), *Transfer of learning from a modern multidisciplinary perspective* (pp. 1–51). Greenwich, CT: Information Age Publishing.
- Scopus®. (2016). *CiteScore Journal Metrics*.
- Selwyn, N. (2009). The digital native – myth and reality. *Aslib Proceedings: New Information Perspectives, 61*(4), 364–379. <http://doi.org/10.1108/00012530910973776>
- Shermis, M. D., Burstein, J., & Leacock, C. (2016). Applications of computers in assessment and analysis of writing. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (2nd ed., pp. 403–416). New York, NY: Guilford Press.

- Slomp, D. H. (2007). *Trapped between paradigms : Composition pedagogy in the context of a twelfth grade standardized writing assessment*. University of Alberta.
- Slomp, D. H. (2008). Harming not helping: The impact of a Canadian standardized writing assessment on curriculum and pedagogy. *Assessing Writing, 13*, 180–200.
<http://doi.org/10.1016/j.asw.2008.10.004>
- Slomp, D. H. (2012). Challenges in assessing the development of writing ability: Theories, constructs and methods. *Assessing Writing, 17*, 81–91.
<http://doi.org/10.1016/j.asw.2012.02.001>
- Slomp, D. H. (2016). An integrated design and appraisal framework for ethical writing assessment. *The Journal of Writing Assessment, 9*, 1–15.
- Slomp, D. H., Corrigan, J. A., & Sugimoto, T. (2014). A framework for using consequential validity evidence in evaluating large-scale writing assessments: A Canadian study. *Research in the Teaching of English, 48*, 276–302.
- Smith, J. K., & Heshusius, L. (1986). Closing down the conversation: The end of the quantitative-qualitative debate among educational inquirers. *Educational Researcher, 15*(1), 4–12.
- Sommers, N. (1980). Revision strategies of student writers and experienced adult writers. *College Composition and Communication, 31*, 378–388.
- Spivey, N., & King, J. R. (1989). Readers as writers composing from sources. *Reading Research Quarterly, 24*(1), 7–26.
- Stagg Peterson, S., McClay, J., & Main, K. (2011). An analysis of large-scale writing assessments in Canada (Grades 5-8). *Alberta Journal of Educational Research, 57*(4), 424–445. Retrieved from <http://ajer.synergiesprairies.ca/ajer/index.php/ajer/article/view/947>

- Statistics Canada. (2012). *2011 census of population: Linguistic characteristics of Canadians*. Ottawa, ON. Retrieved from <http://www12.statcan.gc.ca/census-recensement/2011/as-sa/98-314-x/98-314-x2011001-eng.pdf>
- Stenmark, J. K. (Ed.). (1991). *Mathematics assessment: Myths, models, good questions, and practical suggestions*. Reston, VA: National Council of Teachers of Mathematics.
- Stolarek, E. A. (1994). Prose modeling and metacognition: The effect of modeling on developing a metacognitive stance toward writing. *Research in the Teaching of English*, 28, 154–174. Retrieved from <http://www.jstor.org/stable/40171331> .
- Stotsky, S. (1990). On planning and writing plans. Or beware of borrowed theories! *College Composition and Communication*, 41, 37–57. <http://doi.org/10.2307/357882>
- Street, B. V. (1995). *Social literacies: Critical approaches to literacy in development, ethnography, and education*. London, UK, UK: Longman.
- Street, B. V. (2003). What's "new" in New Literacy Studies? Critical approaches to literacy in theory and practice. *Current Issues in Comparative Education*, 5, 77–91.
- Stroupe, C. (2000). Visualizing English: Recognizing the hybrid literacy of visual and verbal authorship on the Web. *College English*, 62, 607–632. <http://doi.org/10.2307/378964>
- Swarts, J. (2010). Recycled writing: Assembling actor networks from reusable content. *Journal of Business and Technical Communication*, 24, 127–163.
- Tashakkori, A., & Teddlie, C. (Eds.). (2010). *Handbook of mixed methods in social and behavioral research* (2nd.). Thousand Oaks, CA: Sage.
- van den Bergh, H., Rijlaarsdam, G., & van Steendam, E. (2016). Writing process theory: A functional dynamic approach. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (2nd ed., pp. 57–71). New York: Guilford Press.

- van Dijk, J. A. G. M. (2006). Digital divide research, achievements and shortcomings. *Poetics*, 34(4–5), 221–235. <http://doi.org/10.1016/j.poetic.2006.05.004>
- van Gog, T., Paas, F., van Merriënboer, J. J. G., & Witte, P. (2005). Uncovering the problem-solving process: Cued retrospective reporting versus concurrent and retrospective reporting. *Journal of Experimental Psychology. Applied*, 11(4), 237–44. <http://doi.org/10.1037/1076-898X.11.4.237>
- Van Heertum, J., & Share, J. (2006). A new direction for multiple literacy education. *McGill Journal of Education*, 41, 249–265.
- van Weijen, D. (2009). *Writing processes, text quality, and task effects. Empirical studies in first and second language writing*. Utrecht University. Retrieved from <https://dspace.library.uu.nl/handle/1874/33624>
- Vasudevan, L., Schultz, K., & Bateman, J. (2010). Rethinking composing in a digital age: Authoring literate identities through multimodal storytelling. *Written Communication*, 27, 442–468. <http://doi.org/10.1177/0741088310378217>
- Veenman, M. V. J., Hout-Wolters, B. H. A. M., & Afflerbach, P. (2006). Metacognition and learning: conceptual and methodological considerations. *Metacognition and Learning*, 1, 3–14. <http://doi.org/10.1007/s11409-006-6893-0>
- Voss, R. F. (1983). Janet Emig's The Composing Processes of Twelfth Graders: A Reassessment. *College Composition and Communication*, 34, 278–283.
- Vygotsky, L. S. (1997). *Problems of general psychology: The collected works of L. S. Vygotsky, Volume I*. (N. Minick, Ed.). New York: Plenum.
- Wakely, M. B., Hooper, S. R., de Kruif, R. E., & Swartz, C. (2006). Subtypes of written expression in elementary school children: A linguistic-based model. *Developmental*

- Neuropsychology*, 29, 125–159. http://doi.org/10.1207/s15326942dn2901_7
- Wang, M., Haertel, G. D., & Walberg, H. J. (1993). Toward a knowledge base for school learning. *Review of Educational Research*, 63, 249–294.
- Wardle, E., & Roozen, K. (2012). Addressing the complexity of writing development: Toward an ecological model of assessment. *Assessing Writing*, 17, 106–119.
<http://doi.org/10.1016/j.asw.2012.01.001>
- Warren, J. E. (2011). “Generic” and “specific” expertise in English: An expert/expert study in poetry interpretation and academic argument. *Cognition and Instruction*, 29, 349–374.
<http://doi.org/10.1080/07370008.2011.607929>
- White, E. M., Elliot, N., & Peckham, I. (2015). *Very like a whale: The assessment of writing programs*. Logan, UT: Utah State University Press.
- Wiggins, G. (1994). The constant danger of sacrificing validity to reliability: Making writing assessment serve writers. *Assessing Writing*, 1, 129–139. [http://doi.org/10.1016/1075-2935\(94\)90008-6](http://doi.org/10.1016/1075-2935(94)90008-6)
- Wiley, J., Goldman, S. R., Graesser, A. C., Sanchez, C. A., Ash, I. K., & Hemmerich, J. A. (2009). Evaluation, comprehension, and learning in Internet science inquiry tasks. *American Educational ...*, 46, 1060–1106. <http://doi.org/10.2307/40284747>
- Witte, S. P. (1992). Context, text, intertext: Toward a constructivist semiotic theory of writing. *Written Communication*, 9, 237–308. <http://doi.org/10.1177/0741088392009002003>
- Wood, J. W. (2002). Defining literacies. In *Annual Meeting of the American Educational Research Association*. New Orleans, LA.
- Yan, B. (2005). *Demand for skills in Canada: The role of foreign outsourcing and Information-Communication Technology*. Ottawa. Retrieved from

<http://publications.gc.ca/Collection/Statcan/11F0027M/11F0027MIE2005035.pdf>

- Yancey, K. B., Robertson, L., & Taczak, K. (2014). *Writing across contexts: transfer, composition, and sites of writing* (Kindle). Logan, UT: Utah State University Press.
- Yanchar, S. C., & Williams, D. D. (2006). Reconsidering the compatibility thesis and eclecticism: Five proposed guidelines for method use, *35*(9), 3–12.
- Yin, R. K. (2009). *Case study research: Design and methods* (Kindle, 4t). Thousand Oaks, CA: Sage.
- Zieky, M. J. (2016). Fairness in test design and development. In N. J. Dorans & L. L. Cook (Eds.), *Fairness in educational assessment and measurement* (pp. 9–31). New York, NY: Routledge: Taylor & Francis Group.

The Statement of Contribution of Co-Author

Chapter 3 of this dissertation, the article entitled *Articulating a Contemporary Integrated Writing Construct*, was co-authored with my thesis co-advisor Dr. David Slomp of the University of Lethbridge. The construct came about through multiple conversations I had with Dr. Slomp over the years that he supervised my Ph.D. The construct is an elaboration of the concept I articulated in one of my comprehensive examination papers. Dr. Slomp helped me to expand on these ideas. Situating the results within a bioecological framework was the product of his thinking. I devised the idea for a systematic, mixed methods review of the literature, and I both collected and analyzed this data. Then, I wrote up the results initially, which Dr. Slomp helped me to write, refine, and revise.