Crit: A mixed methods examination of epistemic cognition during source evaluations

Courtney Denton Hurlbut

Ph.D. in Educational Psychology

Department of Education and Counselling Psychology

McGill University

Montréal, QC, Canada

March 2021

A thesis submitted to McGill University in partial fulfillment of the requirements of the Degree of Doctor of Philosophy in Educational Psychology

Abstract

Individuals have more access than ever to view, assess and share information. Yet, they continue to struggle during source evaluations and, as a result, accept inaccurate information as truth. Researchers have documented individuals' limited source evaluation behaviours, including their reliance on content-based rather than epistemic criteria to examine reliability. Despite the profound impact of poor source evaluation skills on individuals and society, educational researchers have not widely examined individuals' behaviours in authentic online environments. One promising avenue to better understand individuals' source evaluations is by exploring their epistemic cognition, or their thinking about the epistemic properties of specific knowledge claims and sources. Epistemic cognition researchers have used broad conceptualizations of this construct and controlled settings to investigate individuals' behaviours. As such, it is unclear whether researchers' findings translate to individuals' thinking during source evaluations on the internet. To address this gap in the literature, three studies were conducted to investigate individuals' behaviours using authentic tasks. The first manuscript reports college students' epistemic metacognition related to source evaluations from focus groups interviews and their epistemic cognition during two evaluation tasks. The second manuscript reports the efficacy of an intervention developed to improve individuals' source evaluations on the internet. Based on these empirical studies, theoretical contributions, implications, limitations, and future directions will be discussed.

Keywords: epistemic cognition, source evaluation, online tutorial

Résumé

Les individus ont plus que jamais accès pour consulter, évaluer et partager des informations. Pourtant, ils continuent de se débattre lors des évaluations des sources et, par conséquent, acceptent des informations inexactes comme étant la vérité absolve. Les chercheurs ont documenté les comportements limités d'évaluation des sources des individus, y compris leur dépendance à des critères basés sur le contenu plutôt que sur des critères épistémiques pour examiner la fiabilité. Malgré l'impact profond des faibles compétences d'évaluation des sources sur les individus et la société, les chercheurs en éducation n'ont pas largement examiné les comportements des individus dans des environnements en ligne authentiques. Une avenue prometteuse pour mieux comprendre les évaluations des sources des individus consiste à explorer leur cognition épistémique ou leur réflexion sur les propriétés épistémiques des revendications et des sources de connaissances spécifiques. Les chercheurs en cognition épistémique ont utilisé de vastes conceptualisations de ce concept et des paramètres contrôlés pour étudier les comportements des individus. En tant que tel, il est difficile de savoir si les conclusions des chercheurs se traduisent par la réflexion des individus lors des évaluations de sources sur l'internet. Pour combler cette lacune dans la littérature, trois études ont été menées pour étudier les comportements des individus dans des environnements authentiques et contrôlés. Le premier manuscrit rapporte la métacognition épistémique des étudiants du CÉGEP lié aux évaluations de la source des entretiens de groupes de discussion et leur cognition épistémique au cours de deux tâches d'évaluation. Le deuxième manuscrit se rend compte de l'efficacité d'une intervention développée pour améliorer les évaluations des sources individuelles sur l'internet. Sur la base de ces études empiriques, les contributions théoriques, les implications, les limites et les orientations futures seront discutées.

Table of Contents

List of Tables

List of Figures

List of Appendices

Dedication

For Toby and Kendrick, who grew through triumphs as much as trials.

Acknowledgements

I am grateful for the myriad of support I have received along this journey. First and foremost, I wish to express my sincere appreciation to my supervisor, Dr. Krista R. Muis, who provided many learning opportunities to gain valuable research experience over the last four years. Thank you for challenging me to become a better researcher, writer, and mentor. I appreciate the time you have invested in my success.

I would also like to extend my deepest gratitude to Dr. Adam Dubé, who provided instrumental feedback to improve the quality of my research and writing. Thank you for extending opportunities to grow and apply my knowledge about source evaluations.

I am also grateful to the institutions that recognized the importance of my research and the educators that welcomed me into their classrooms. Without your generous support and guidance, this project would not have been possible.

A very special thanks to Cara Singh, for your help with data collection and practical suggestions, and Skylar Armstrong, for your unparalleled support during data analysis. Many thanks to my lab mates who walked this path with me. Thank you to Marianne Chevrier, Ivana Di Leo, Nesma Etoubashi, Kelsey Losenno, Brendan Munzar, and Maria Rizk who taught me the ropes in the lab. Thanks also to my academic friends, Stephanie Beck, Chiung-Fang Chang, Hafiz Hashim, Emily Jonas, Gulshah Kacmaz, So Yeon Lee, Shan Li, and Jenny Zheng. I am grateful for every deep conversation over a cup of tea and each special celebration with you all.

To my partner, Yen Cam, who brightens each day by saying “epistemology,” thank you. Your love has been one of the greatest gifts along this arduous road.

I would also like to acknowledge the financial support during my doctoral studies from the Fonds de Recherche du Québec – Culture et Société (FRQSC).

Preface and Contribution of Authors

I am the primary author on all chapters included in this thesis and I am responsible for their content. I independently wrote the chapters and Drs. Krista R. Muis and Adam Dubé provided feedback. Skylar Armstrong assisted with the search procedure described in Chapter 2. I received feedback from co-authors on the empirical manuscripts (Chapters 3 and 4). Their specific contributions are summarized below using Brand and colleagues' (2015) Contributor Roles Taxonomy.[1]

Author Contributions for Chapters 3 & 4

CAD: Conceptualization, Methodology, Project administration, Investigation, Data curation, Supervision, Formal analysis, Writing – Original Draft

KRM: Funding acquisition, Supervision, Conceptualization, Methodology, Resources, Writing – Review & Editing

SA: Formal analysis, Writing – Review & Editing

AD: Supervision, Conceptualization, Methodology, Writing – Review & Editing

[1]Brand, A., Allen, L., Altman, M., Hlava, M., & Scott, J. (2015). Beyond authorship: attribution, contribution, collaboration, and credit. *Learned Publishing*, *28*(2), 151-155.

Chapter 1

Introduction

The World Health Organization's declaration of an "infodemic" in March 2020 reinvigorated discussions about the quality of information found online. Faster than ever before, individuals can access information from around the globe. Unfortunately, much of the available information is unfiltered and uncontrolled, which makes source evaluations difficult. In an era where dependence on the traditional gatekeepers of knowledge is no longer sufficient (e.g., professional editors, Jenkins, 2009), everyone is responsible for determining the reliability of the information they encounter (Strømsø & Kammerer, 2016). This task can quickly become challenging when source information is unavailable (Greene, 2016). All ages struggle with source evaluations if they have not received focused training (Braasch et al., 2013; Hämäläinen et al., 2020; Wiley et al., 2009; Zhang & Duke, 2011). As a result, misinformation spreads at alarming rates about numerous topics, including health-related topics such as COVID-19 (Kouzy et al., 2020).

To address this concerning trend, many places have adopted educational standards aimed at improving students' 21st-century skills, including their ability to evaluate sources. For example, Australian policymakers have developed an extensive description of digital skills for each discipline and education level, which includes authentic skill-building activities (e.g., evaluate the reliability and usefulness of a source as evidence of an historical event), achievement standards (e.g., an "above satisfactory" student can analyze the value of primary and secondary sources in an argument), and illustrative examples of student performance for each skill (e.g., portfolio with source analysis, Australian Curriculum, 2020). In Canada, Québec policymakers have developed a general digital competency framework that includes desired skills (e.g., using rigorous criteria to determine content-based and source-based reliability), sample themes without description (e.g., media literacy), and broad achievement standards (e.g.,

an "advanced" student can use appropriate criteria to determine reliability, MEES, 2018). Chinn and colleagues (2020) argued that the lack of specificity, as seen in Québec's plan and many others', undermines the efficacy of their efforts to train critical information consumers. Additionally, rapid changes in technology suggest that individuals' evaluation skills require ongoing updates (Ala-Mutka, 2011). To support skill development, Chinn and colleagues (2020) recommended that educators design authentic learning environments to prepare students for a broad range of epistemic activities, including evaluating sources, claims, evidence, and arguments. That is, preparing individuals for epistemically unfriendly environments, such as the internet, by developing their epistemic cognition (Greene & Yu, 2016; Sinatra & Chinn, 2012). Epistemic cognition (EC) is used to describe thinking about the acquisition, justification, and use of knowledge (Hofer, 2016).

Educators and EC researchers alike have ignored the importance of authentic practice in online environments—favouring curated environments that simplify the complexities of source evaluations online instead (Barzilai & Chinn, 2018; Chinn et al., 2020). Consequently, researchers' findings about students' source evaluations have been somewhat incorrectly generalized to their behaviours in authentic settings. For example, Mason and colleagues (2010) acknowledged that individuals require new skills to navigate internet source evaluations, yet the researchers continued to design offline environments to assess these behaviours afterward (Mason et al., 2011, 2014, 2018). Despite similarities between online and offline source evaluations, the latter does not account for the unfiltered environment found on the internet. Further, researchers in diverse fields have documented varying degrees of EC during source evaluations. For example, Greene and colleagues (2014, 2018) found that individuals employ limited epistemic evaluation criteria to assess source reliability, whereas Kąkol and colleagues

(2017) observed high rates of individuals' epistemic criteria use. Given the vastly different methods and results used as well as the successes of source evaluation trainings in controlled settings (e.g., Pretorius, 2018), further investigation about source evaluations in authentic online environments is required.

The purpose of this thesis is to investigate individuals' EC during source evaluations and empirically evaluate a training aimed at supporting their evaluation behaviours. To identify the role of EC in source evaluations, Chapter 2 presents a literature review of personal epistemology theory (Part 1) and empirical work on source evaluations from learning sciences (Part 2). Chapters 3 and 4 present a multiphase mixed methods research program that examines college students' epistemic metacognition about source evaluations and their EC during evaluation tasks (Chapter 3) to develop a training for adults that supports their ability to evaluate information on the internet (Chapter 4). The chapters within this thesis address the following questions:

Chapter 2

1. What is epistemic cognition and how has it been conceptualized?
2. How can epistemic cognition research inform future research about source evaluations?

Chapter 3

3. What characterizes college students' epistemic metacognitive knowledge about source evaluations on the internet?
4. How do college students' epistemic ideals contribute to their overall source evaluations?

Chapter 4

5. Do adults in an intervention group that received source evaluation training evaluate sources more accurately than those in a comparison group?
6. How do adults' epistemic ideals contribute to their overall source evaluations?
7. How do adults' epistemic ideals contribute to the complexity of their written argumentation?

Chapter 5

8. What are promising avenues for future epistemic cognition research on authentic source evaluations?

By addressing these questions, this thesis contributes new knowledge to the understanding of EC during source evaluations on the internet and further contributes to understanding how source evaluation trainings can be designed to support these important skills. As a result, this thesis provides guidance for researchers and educators to better develop and leverage EC to promote critical evaluation of unfiltered information on the internet.

Overview of the Chapters

This thesis captures the needs assessment, program development and program evaluation testing of a multiphase research program aimed at improving source evaluation skills. In Chapter 2, relevant theoretical and empirical literature are synthesized to provide a foundation for developing the online training. Previous researchers' theoretical and methodological choices are also highlighted and critiqued to identify gaps in the literature. The review explored theorists' conflicting conceptualizations of EC and researchers' limited use of online environments to explore EC. The limitations of previous work guided methodological choices throughout the research program.

Chapter 3 presents two empirical studies that explored college students' epistemic metacognition about source evaluations and their EC during two different evaluation tasks. In Study 1, participants discussed the epistemic ideals and the reliable epistemic processes they engage in to evaluate information on the internet during focus group interviews. Participants also assessed the reliability of an online news article. In Study 2, participants were asked to rank-order two news articles and justify their rank-ordering. Consistent with the literature, participants described and used a variety of criteria to assess reliability, including important epistemic ideals such as source expertise, evidence quality, and corroboration. These insights were used to develop an online training for adults.

Chapter 4 describes the training developed to improve adults' source evaluation skills. The training was based on a widely endorsed evaluation tool, the CRAAP test, which includes questions that help the user assess a source's currency, relevance, authority, accuracy, and purpose. To examine the efficacy of the training, adults across Canada either received source evaluation training before completing an evaluation task or proceeded directly to the task without training. Like Study 2, adults' EC was captured through their rank-order justifications and, additionally, through their essay responses. Results from the study support critiques of the CRAAP test as a tool to evaluate sources on the internet. Future research should examine epistemic ideals in relation to epistemic aims and processes during online source evaluations to better understand the importance of emphasising ideals or processes in trainings.

To conclude, Chapter 5 provides a summary of the contributions and limitations of the work described in this thesis. Using a pragmatist approach to investigate source evaluations on the internet, the strengths of the methods and analyses used in education research were used in

more authentic contexts and tasks. Additionally, the present thesis extends education research on to a novel population whose source evaluations have been underexplored.

References

Ala-Mutka, K. (2011). *Mapping digital competence: Towards a conceptual understanding* [PDF document]. Institute for Prospective Technological Studies. ftp://s-jrcsvqpx102p.jrc.es/pub/EURdoc/JRC67075_TN.pdf

Australian Curriculum. (2020). *Digital technologies*. Australian Curriculum. https://www.australiancurriculum.edu.au/f-10-curriculum/technologies/digital-technologies/?year=12983&year=12984&year=12985&year=12986&year=12987&strand=Digital+Technologies+Knowledge+and+Understanding&strand=Digital+Technologies+Processes+and+Production+Skills&capability=ignore&capability=Literacy&capability=Numeracy&capability=Information+and+Communication+Technology+%28ICT%29+Capability&capability=Critical+and+Creative+Thinking&capability=Personal+and+Social+Capability&capability=Ethical+Understanding&capability=Intercultural+Understanding&priority=ignore&priority=Aboriginal+and+Torres+Strait+Islander+Histories+and+Cultures&priority=Asia+and+Australia's+Engagement+with+Asia&priority=Sustainability&elaborations=true&elaborations=false&scotterms=false&isFirstPageLoad=false

Barzilai, S., & Chinn, C. A. (2018). On the goals of epistemic education: Promoting apt epistemic performance. *Journal of the Learning Sciences*, *27*(3), 353–389. https://doi.org/10.1080/10508406.2017.1392968

Braasch, J. L. G., Bråten, I., Strømsø, H. I., Anmarkrud, Ø., & Ferguson, L. E. (2013). Promoting secondary school students' evaluation of source features of multiple documents. *Contemporary Educational Psychology*, *38*(3), 180–195. https://doi.org/10.1016/j.cedpsych.2013.03.003

Chinn, C. A., Barzilai, S., & Duncan, R. G. (2020). Education for a "post-truth" world: New directions for research and practice. *Educational Researcher*, 0013189X2094068. https://doi.org/10.3102/0013189X20940683

Greene, J. A. (2016). Interacting epistemic systems within and beyond the classroom. In J. A. Greene, W. A. Sandoval, & I. Bråten (Eds.), *Handbook of epistemic cognition* (pp. 265–278). Routledge.

Greene, J. A., Copeland, D. Z., Deekens, V. M., & Seung, B. Y. (2018). Beyond knowledge: Examining digital literacy's role in the acquisition of understanding in science. *Computers & Education*, *117*, 141–159. https://doi.org/10.1016/j.compedu.2017.10.003

Greene, J. A., & Yu, S. B. (2016). Educating critical thinkers: The role of epistemic cognition. *Policy Insights from the Behavioral and Brain Sciences*, *3*(1), 45–53. https://doi.org/10.1177/2372732215622223

Greene, J. A., Yu, S. B., & Copeland, D. Z. (2014). Measuring critical components of digital literacy and their relationships with learning. *Computers & Education*, *76*, 55–69. https://doi.org/10.1016/j.compedu.2014.03.008

Hämäläinen, E. K., Kiili, C., Marttunen, M., Räikkönen, E., González-Ibáñez, R., & Leppänen, P. H. T. (2020). Promoting sixth graders' credibility evaluation of Web pages: An intervention study. *Computers in Human Behavior*, *110*, 106372. https://doi.org/10.1016/j.chb.2020.106372

Hofer, B. K. (2016). Epistemic cognition as a psychological construct: Advancements and challenges. In J. A. Greene, W. A. Sandoval, & I. Bråten (Eds.), *Handbook of epistemic cognition* (pp. 31–50). Routledge.

Jenkins, H. (2009). *Confronting the challenges of participatory culture: Media education for the 21st century*. MIT Press.

Kąkol, M., Nielek, R., & Wierzbicki, A. (2017). Understanding and predicting Web content credibility using the Content Credibility Corpus. *Information Processing & Management*, *53*(5), 1043–1061. https://doi.org/10.1016/j.ipm.2017.04.003

Kouzy, R., Abi Jaoude, J., Kraitem, A., El Alam, M. B., Karam, B., Adib, E., Zarka, J., Traboulsi, C., Akl, E., & Baddour, K. (2020). Coronavirus goes viral: Quantifying the COVID-19 misinformation epidemic on Twitter. *Cureus*. https://doi.org/10.7759/cureus.7255

Mason, L., Ariasi, N., & Boldrin, A. (2011). Epistemic beliefs in action: Spontaneous reflections about knowledge and knowing during online information searching and their influence on learning. *Learning and Instruction*, *21*(1), 137–151. https://doi.org/10.1016/j.learninstruc.2010.01.001

Mason, L., Boldrin, A., & Ariasi, N. (2010). Epistemic metacognition in context: Evaluating and learning online information. *Metacognition and Learning*, *5*(1), 67–90. https://doi.org/10.1007/s11409-009-9048-2

Mason, L., Junyent, A. A., & Tornatora, M. C. (2014). Epistemic evaluation and comprehension of web-source information on controversial science-related topics: Effects of a short-term instructional intervention. *Computers & Education*, *76*, 143–157. https://doi.org/10.1016/j.compedu.2014.03.016

Mason, L., Scrimin, S., Tornatora, M. C., Suitner, C., & Moè, A. (2018). Internet source evaluation: The role of implicit associations and psychophysiological self-regulation. *Computers & Education*, *119*, 59–75. https://doi.org/10.1016/j.compedu.2017.12.009

Ministère de l'Enseignement supérieur (MEES). (2018). *Digital action plan for education and higher education* [PDF document]. http://www.education.gouv.qc.ca/fileadmin/site_web/documents/ministere/PAN_Plan_action_VA.pdf

Pretorius, L. (2018). Experiential and self-discovery learning in digital literacy: Developing the discernment to evaluate source reliability. *College & Undergraduate Libraries, 25*(4), 388–405. https://doi.org/10.1080/10691316.2018.1530626

Sinatra, G. M., & Chinn, C. A. (2012). Thinking and reasoning in science: Promoting epistemic conceptual change. In K. R. Harris, S. Graham, T. Urdan, A. G. Bus, S. Major, & H. L. Swanson (Eds.), *APA educational psychology handbook: Application to learning and teaching* (Vol. 3, pp. 257–282). American Psychological Association.

Strømsø, H. I., & Kammerer, Y. (2016). Epistemic cognition and reading for understanding in the internet age. In J. A. Greene, W. A. Sandoval, & I. Bråten (Eds.), *Handbook of epistemic cognition* (pp. 230–246). Routledge.

Wiley, J., Goldman, S. R., Graesser, A. C., Sanchez, C. A., Ash, I. K., & Hemmerich, J. A. (2009). Source evaluation, comprehension, and learning in internet science inquiry tasks. *American Educational Research Journal, 46*(4), 1060–1106. https://doi.org/10.3102/0002831209333183

Zhang, S., & Duke, N. K. (2011). The impact of instruction in the WWWDOT framework on students' disposition and ability to evaluate web sites as sources of information. *The Elementary School Journal, 112*(1), 132–154. https://doi.org/10.1086/660687

Chapter 2

Literature Review

Many individuals worldwide spend hours on the internet each day (Hootsuite & We Are Social, 2020). Whether they are looking at news updates on COVID-19 or researching another topic, they are constantly evaluating which sources to trust. During a source evaluation, a source's credibility and the adequacy of its' content are assessed for the user's needs. Their evaluation processes may include thinking about the epistemic properties of the source and its' content (Barzilai & Zohar, 2014). Individuals' personal epistemology is one important aspect of understanding their source evaluation behaviours. Specifically, their EC may support or constrain their ability to accurately evaluate online sources. For example, individuals that employ reliable epistemic processes, such as knowledge-based validation or corroboration of evidence, tend to demonstrate more accurate source evaluations (Mason et al., 2014); however, individuals that engage in such processes may still struggle differentiating the quality of sources (Mason et al., 2018). Whereas educational researchers have documented reliance on content- and design-based features to establish trustworthiness (e.g., task relevance, Braasch et al., 2013; Gerjets et al., 2011; Kiili et al., 2008), the implications of these findings for online source evaluations are undermined by dissenting theoretical conceptualizations of EC and empirical investigations that lack ecological validity. To address these concerns, the purpose of this literature review is to examine how theoretical and empirical explorations of EC can inform the development and implementation of source evaluation trainings. Part 1 will review theoretical conceptualizations of personal epistemology and, specifically, EC. Part 2 will examine empirical evidence of individuals' EC during source evaluations.

Part 1. What is epistemic cognition and how has it been conceptualized?

Learning and, more recently, online learning has served as an important context for examining personal epistemology. Broadly, an individual's personal epistemology refers to their

thinking about knowledge and how they come to know (Hofer, 2001; Hofer & Pintrich, 1997). Theorists have delineated the multifaceted nature of personal epistemology, identifying EC, epistemic metacognition, epistemic motivation, and epistemic emotions as the primary areas of investigation in the field (Barzilai & Chinn, 2018; Barzilai & Zohar, 2014; Muis & Singh, 2018). Despite clear distinctions between these areas, the field continues to be in a state of "theoretical competition" (Sandoval, 2009, p. 153). Whereas some theorists argue to differentiate the aspects of personal epistemology (Muis & Singh, 2018), others have moved to unify the field's diverse body of work under the term EC, rather than personal epistemology (e.g., Chinn et al., 2011, 2014; Greene et al., 2008, 2018). As is evidenced by the categories of epistemic "somethings" in the *Handbook of Epistemic Cognition*, these theorists and researchers have come to include metacognitive and cognitive processes (e.g., epistemic ideals), motivation (e.g., epistemic values), affect (e.g. epistemic emotions), tasks and contexts (e.g., epistemic climate) and much more under the term EC (see Table 30.1, Greene et al., 2016).

To add further confusion, theorists have employed assorted labels to describe the cognitive and metacognitive elements of personal epistemology examined in the field, including reflective judgement (King & Kitchener, 1994), epistemic (Muis, 2007) or epistemological beliefs (Schommer, 1990), epistemological reflection (Baxter Magolda, 2004), EC (Chinn et al., 2011, 2014, 2016), epistemic thinking (Barzilai & Chinn, 2018; Barzilai & Zohar, 2014, 2016), practical epistemologies (Sandoval, 2005), and epistemological resources (Elby & Hammer, 2010; Hammer & Elby, 2002; Louca et al., 2004). Regardless of the term used, theorists have largely explored the nature of knowledge, or what knowledge is, and the nature of knowing, or how knowledge is constructed and evaluated (Hofer & Pintrich, 1997). Some theorists have also included factors related to the nature of learning (e.g., Baxter Magolda, 2004; Elby & Hammer,

2002, 2010; Schommer, 1990), which includes beliefs about how learning occurs (Hofer & Pintrich, 1997). Due to the variability within the field, concerns persist about differing conceptualizations of key constructs, including EC itself (Alexander, 2016).

So, what does *epistemic* mean? The term epistemic refers to being of knowledge or relating to knowledge (Kitchener, 2011). Thus, individuals' EC is a specific type of their cognition (e.g., reasoning, decision-making, Hofer, 2016; Sinatra, 2016) that focuses on the acquisition, justification, and use of knowledge (Greene et al., 2018). The purpose of Part 1 is to distinguish elements of EC from other epistemic constructs, which may have cognitive components (e.g., affect, Muis et al., 2018). Following Muis and Singh (2018), *personal epistemology* will be used to globally refer to the field of investigation. The field's component parts, such as epistemic metacognition, will also be used for classification. Within the global definition of personal epistemology, *epistemic beliefs* will be used to refer to an individual's tacit beliefs about the nature of knowledge or knowing that influence other facets of EC (Muis, 2007). The term *source evaluation* will be used to describe individuals' cognitive processes as they assess various features to determine reliability, and *source integration* will be used to describe individuals' reconciliation of multiple sources (Barzilai & Zohar, 2014). Further, use of the terms *online* and *offline* during discussions of source evaluations will describe a context with active, open internet access and with a disconnected digital or paper-based environment, respectively. To delineate the nature, structure, and function of EC, theoretical frameworks of personal epistemology will be reviewed and parsed, with an emphasis on cognitive and metacognitive frameworks. See Figure 1 for a classification of key constructs explored in personal epistemology.

Figure 1

Classification of Key Constructs from Selected Theorists

<table>
<tr><th>Epistemic Cognition</th><th>Epistemic Metacognition</th><th>Epistemic Motivation</th><th>Epistemic Emotions</th></tr>
<tr><td colspan="4">Epistemic aims, ideals, and reliable epistemic processes[1,2]</td></tr>
<tr><td colspan="2">Epistemic thinking[1,3]</td><td rowspan="4"></td><td>Epistemic experiences[1,3]</td></tr>
<tr><td colspan="2">Epistemological resources[4]</td><td>Epistemic emotions[8]</td></tr>
<tr><td>Epistemic strategies[1,3]</td><td>Epistemic metacognitive knowledge and skills[1,3]</td><td rowspan="3"></td></tr>
<tr><td>Epistemic artifacts or products</td><td>Epistemic and epistemological beliefs[5,6]</td></tr>
<tr><td></td><td>Epistemological views[7]</td></tr>
</table>

[1] Barzilai and Chinn (2018)
[2] Chinn and colleagues (2011, 2014)
[3] Barzilai and Zohar (2014, 2016)
[4] Elby and Hammer (2002, 2010)
[5] Hofer and Pintrich (1997)
[6] Schommer (1990)
[7] Kuhn (1991)
[8] Muis and colleagues (2018)

Frameworks of Personal Epistemology

Theorists have situated aspects of personal epistemology in existing mechanisms of cognitive and metacognitive processing (e.g., Barzilai & Zohar, 2014; Hofer, 2004; Kitchner, 1983; Muis, 2007; Muis et al., 2018). To illustrate the relationship between epistemic and non-epistemic processes during source evaluations, consider this simplified scenario:

> Toby, the average post-secondary student, is concerned about the efficacy of treatments for COVID-19 and decides to consult several online sources. After entering their keywords into the search bar, Toby selects the first webpage on the results page. Toby begins to read that source (cognition) and monitors their reading progress (metacognition) while comparing the source's content to their aims (epistemic metacognition). Toby considers the adequacy of evidence presented (epistemic metacognition) by noting any inconsistencies in the information presented (epistemic cognition). Ultimately, Toby decides to see if other sources use similar evidence to support their efficacy claims (epistemic metacognition). After returning to the results page, Toby assesses the relevance of the other search results (cognition) and selects the fourth webpage because it appears to be published by a reliable venue (epistemic cognition).

Toby may evaluate and integrate information from many sources before they decide to stop their search. Following Kitchener (1983), Barzilai and Zohar (2014) described individuals' epistemic processes as monitoring their non-epistemic processes. To examine how these processes have been incorporated into theory, Greene and colleagues' (2018) classification of theoretical work will be used to discuss the contributions and limitations of developmental, multidimensional, philosophical, and disciplinary perspectives.

Developmental Perspectives

Developmental theorists (e.g., King & Kitchener, 1994; Kuhn & Weinstock, 2002; Perry, 1970) have posited that epistemic development occurs over time. From this perspective, an individual may progress through distinct qualitative levels of thinking about the nature of knowledge and knowing that range from concrete to contextual (Hofer & Pintrich, 1997). Stemming from Perry's seminal research (1970), theorists have explored epistemic development

through women's ways of knowing (Belenky et al., 1986), relativistic thinking (Chandler et al., 1990), argumentative reasoning (Kuhn, 1991), reflective judgment (King & Kitchener, 1994), and epistemological reflection (Baxter Magolda, 2004). These theorists introduced relationships between personal epistemology and other constructs (e.g., critical thinking, King & Kitchener, 1994; Kuhn, 1991, 1999) and skills (e.g., argumentation, Kuhn, 1991). Following Muis and Singh (2018), Kuhn's framework of argumentative reasoning will be used to describe individuals' general progression through developmental stages (Kuhn, 1991; Kuhn et al., 2000; Kuhn & Weinstock, 2002).

Kuhn and colleagues' work (Kuhn, 1991; Kuhn et al., 2000; Kuhn & Weinstock, 2002) described aspects of metacognitive processes related to knowledge and knowing through four stages: realist, absolutist, multiplist, and evaluativist. According to Kuhn, young children begin at the realist stage, where they view knowledge as a direct copy of reality. Since they believe knowledge is certain and transmitted from an external source (e.g., an authority, such as a parent), children do not engage in critical thinking. Researchers have found that children believe most information found on the internet is true (Hirsh, 1999; Schacter et al., 1998). At the absolutist stage, individuals continue to view knowledge as certain, but their understanding of knowledge has also taken on the quality of objectivity, wherein facts are definitively right or wrong. Like the realist stage, individuals believe that knowledge reflects reality and is transmitted from an external source. However, they engage in critical thinking to evaluate the veracity of information. Following Kuhn, Barzilai and Zohar (2012) classified students that believed only one website could be right as absolutists. These students reported using relevance and website familiarity as criteria to choose websites.

Kuhn (1991) asserted that individuals transition to the multiplist stage through access to conflicting paradigms. At this stage, individuals view knowledge as uncertain and constructed. Multiplists believe that different or conflicting perspectives are equally valid. As a result, they do not engage in critical thinking about information because they view this act as irrelevant. Barzilai and Zohar (2012) found that multiplists believed that more than one website could provide the truth, but that each site was equally valid in doing so. At the highest stage, evaluativist, individuals continue to consider knowledge uncertain and constructed. They believe multiple perspectives require evaluation of argument quality and supporting evidence. Further, evaluativists believe that knowledge and knowledge claims that can be verified can be used to generate tentative conclusions. As such, individuals engage in critical thinking to support their evaluations and valid judgments. In accordance with Kuhn's description, Barzilai and Zohar (2012) found that evaluativists described corroboration strategies more frequently than their absolutist counterparts.

The developmental perspective, as illustrated by Kuhn and colleagues' framework, describes a unidimensional progression of epistemic development. Higher stages of epistemic development emerge during late adolescence or adulthood (Hofer & Pintrich, 1997), with advancement occurring as a result of cognitive disequilibrium (Sandoval et al., 2016). According to Kienheus and colleagues (2008), cognitive dissonance occurs when individuals are exposed to new experiences that conflict with their existing beliefs, as seen in the critical transition between absolutist and multiplist. As such, various learning environments have served as important contexts to explore personal epistemology, including inquiries about socio-scientific topics (e.g., Muis et al., 2020; Trevors et al., 2017) and about historical events (e.g., Barzilai & Eshet-Alkalai, 2015; Barzilai & Weinstock, 2015). Due to differing findings in various domains,

theorists have highlighted problematic assumptions within the developmental perspective, including the domain-general approach to personal epistemology (e.g., Chandler et al., 1990; Chinn & Sandoval, 2018; Muis et al., 2006), contradictions between theory and expert practices (e.g., Bromme et al., 2010; Chinn et al., 2011; Chinn & Rinehart, 2016), and higher levels of development emerging in adolescence (Greene et al., 2018; Cartiff et al., 2020). Such challenges have led to domain-dependent developmental theories, including Kuhn and colleagues' later work (2000) and Chandler and colleagues' work (1990), as well as disciplinary perspectives (e.g., Elby & Hammer, 2010; Hammer & Elby, 2002; Louca et al., 2004; Sandoval, 2005). Instead of a single cohesive shift to higher levels of thinking (e.g., multiplist to evaluativist), domain-dependent theories described differences in progression for distinct domains. For example, Toby may transition to a higher stage of epistemic development in fashion opinions (i.e., taste) before transitioning in online research about COVID-19. Further, several developmental theories accounted for individuals that appeared to be in more than one stage at a time (e.g., Chandler et al, 1990; King & Kitchener, 1994; Perry, 1970).

Adding to the understanding of personal epistemology as domain-specific, Chandler and colleagues (Chandler et al., 1990; Chandler et al., 2002) explained that individuals transition to higher levels of thinking in areas that have little epistemic content prior to transitioning in areas with significant epistemic content. They described the developmental progression from an objective view of knowledge, like Kuhn's absolutist, to a stage where individuals have faith in their own judgement of the adequacy of knowledge claims, like Kuhn's evaluativist. The second level of Chandler and colleagues' (1990; 2002) framework differentiated their conceptualization from other developmental theorists (e.g., Kuhn, Perry). At the second level, individuals view knowledge as a human construction, where they can travel one of two paths: the individual can

either become a (a) dogmatist that believes knowledge is constructed and only comes from an authority figure, or (b) a skepticist that believes knowledge is constructed and we cannot know knowledge or the truth. In Chandler and colleagues' (1990; 2002) dichotomy, the dogmatist equates to Kuhn's multiplist, whereas the skepticist describes a new concept for the field. Chandler and colleagues' (1990) inclusion of the skepticist accounts for the fact that not all adults reach the highest stage of epistemic development. If Toby became a skepticist, they would view themself as a source of knowledge, but also acknowledge that the truth cannot be attained. Toby's beliefs may manifest similarly to Kuhn's absolutist, where the first (or most convenient) source is deemed acceptable. That is, Toby would view source evaluation and integration as irrelevant.

Despite Chandler and colleagues' (1990) contribution, Murphy and colleagues (2007) argued that the concept of skepticism does not support theorists' and researchers' efforts to understand personal epistemology. Specifically, to understand personal epistemology during source evaluations, Barzilai and Chinn (2020) identified four dynamics that have led to the post-truth era—rather than describing skeptics, they highlighted deficits in knowledge and skills, commitments to truth, and disagreements about knowing. Undoubtedly, developmental theorists' work, including Chandler and colleagues' (1990; 2002), stimulated further exploration of personal epistemology in other perspectives. The unidimensional portrayal of epistemic development in the developmental approach captured the nature of knowledge factors that underlie the nature of knowing factors featured in later multidimensional frameworks (Barzilai & Zohar, 2014). Next, frameworks developed in response to developmental theorists' work are presented.

Multidimensional Perspectives

Multidimensional theorists challenged previous conceptualizations by proposing dimensions of epistemological (Schommer, 1990) and epistemic beliefs (Hofer & Pintrich, 1997). Originally, Schommer (1990) hypothesized that beliefs were independent (Hofer & Pintrich, 1997), but Hofer (2004) later amended this hypothesis, stating that individuals' beliefs integrate in an overarching, predictable pattern. Like the developmental perspective that captured individuals' overall epistemic progression, multidimensional theorists argued that individuals' beliefs function on continua that range from concrete to contextual. Multidimensional theorists and researchers have examined epistemic beliefs in terms of the domain-general (e.g., Hofer & Pintrich, 1997), domain-specific (e.g., Barzilai & Weinstock, 2015; Muis et al., 2006), and topic-specific properties (e.g., Merk et al., 2018; Merk, Rosman, et al., 2017; Merk, Schneider, et al., 2017). Whereas some similarities between paradigms exist, Schommer's seminal research (1990) initiated empirical examination of their theorized dimensions. Chinn and colleagues noted that "many subsequent publications have explicitly drawn on this influential conceptualization" (2011, p. 141).

Schommer's work (1990; Schommer-Aikins, 2004) proposed five dimensions that could be evaluated quantitatively. The dimensions were fixed ability, quick learning, simple knowledge, certain knowledge, and a hypothesized fifth dimension, source of knowledge. Like Schommer, Hofer and Pintrich (1997) described four dimensions: certainty of knowledge, simplicity of knowledge, source of knowledge, and justification for knowing. Hofer and Pintrich's (1997) framework, which does not include the nature of learning factors Schommer (1990) did, will be used to elaborate on the hypothesized dimensions. The first two dimensions, certainty and simplicity of knowledge, are considered the nature of knowledge factors because

they describe the limits of knowledge. The certainty of knowledge dimension ranges from the belief that knowledge is absolute and static to the belief that knowledge is uncertain and evolving. This conceptualization is consistent with developmental (e.g., Kuhn, 1991; King & Kitchener, 1994) and other multidimensional (Schommer, 1990) frameworks, where the high end of the continuum reflects being open to knowledge revision because the individual understands that knowledge is tentative. The second dimension, simplicity of knowledge, ranges from the belief that information is isolated and disconnected to the belief that concepts are highly connected and interrelated. Using Toby's experience as an example, their belief that knowledge is uncertain and complex during their online navigation may manifest in the use of multiple sources to corroborate evidence.

The other two dimensions in Hofer and Pintrich's (1997) framework, source of knowledge and justification for knowing, are considered the nature of knowing factors because they describe criteria for knowing. The source of knowledge dimension ranges from the belief that knowledge is transmitted from authorities (i.e., from outside the self) to the belief that it is constructed by the individual as they interact with the environment (i.e., constructivist). Like developmental frameworks, the critical transition on the source of knowledge dimension occurs when an individual views themselves as a knower that can construct knowledge through interaction. Lastly, the justification for knowing dimension ranges from justification through direct observation or transmission from authority to use of rules of inquiry to evaluate and integrate multiple sources. This dimension describes how an individual evaluates knowledge claims, including their evaluation of experts and use of evidence (Hofer & Pintrich, 1997). The nature of knowing factors are highly related as beliefs about the source of knowledge may influence the types of justifications deemed adequate to support a knowledge claim. In the

example, Toby's belief that knowledge is a human construction may allow more diverse evidence to be accepted as sufficient, such as evidence from memory, reasoning, or perception.

The four hypothesized dimensions described by Hofer and Pintrich (1997) can be tied back to developmental descriptions of personal epistemology. As such, some theorists have developed integrative models to address limitations of purely developmental or multidimensional frameworks (e.g., Barzilai & Weinstock, 2015; Bendixen & Rule, 2004; Greene et al., 2008; Rule & Bendixen, 2010). In integrative models, personal epistemology is described as both developmental and multidimensional. According to Merk and colleagues (2018), these models assume that several dimensions (e.g., nature of knowledge factors, Rule & Bendixen, 2010) develop independently over time. Drawing from developmental (Baxter Magolda, 2004) and multidimensional theorists (Schommer, 1990; Hofer & Pintrich, 1997), Bendixen and Rule's (2004; Rule & Bendixen, 2010) integrative model presents a clearer path for epistemic development than previous approaches. Set within a learning context, their mechanism of change described the requirements which must be met for an individual to experience epistemic development: epistemic doubt, epistemic volition, and resolution strategies. Epistemic doubt refers to a specific type of cognitive disequilibrium, where the individual examines their beliefs about knowledge or knowing. Bendixen (2002) explained that an individual's independence and their exposure to beliefs that conflict with currently held beliefs are important for epistemic doubt to occur (i.e., cognitive dissonance). This condition is similar to Kuhn's (1991) description of the transition from absolutist to multiplist. In the example, Toby might question their beliefs about the source of knowledge when examining knowledge claims from "experts' with varying degrees of reliability. Bendixen and Rule (2004; Rule & Bendixen, 2010) also elaborated on epistemic volition, which describes an individuals' directed effort to change beliefs about

knowledge and knowing. Toby might enact a procedure to examine source reliability to advance their epistemic beliefs. Lastly, Bendixen and Rule (2004; Rule & Bendixen, 2010) described resolution strategies, which are the tactics an individual uses to achieve belief change, such as reflection and interaction with others. Like developmental theorists, Bendixen and Rule (2004; Rule & Bendixen, 2010) explained that an individuals' progression through these three conditions were not guaranteed. They asserted that even if all requirements were met for change to occur, an individuals' beliefs about knowledge and knowing could regress. As such, according to this framework, Toby's beliefs could revert to their original beliefs about knowledge and knowing at any point.

Despite contributions, early multidimensional work (e.g., Hofer & Pintrich, 1997; Kuhn, 1999) was also criticized for the domain-general approach taken toward personal epistemology (e.g., Hofer, 2000; Muis et al., 2006; Sandoval et al., 2016). Such challenges led to Muis and colleagues' (2006) development of the Theory of Integrated Domains in Epistemology (TIDE) framework and Merk and colleagues' (2018) extended TIDE framework. Muis and colleagues' (2006) TIDE framework described the progression of belief development from general epistemic beliefs to domain-specific epistemic beliefs, as an individual deepens their understanding of different domains. This progression indicated that an individual could exhibit a range of epistemic beliefs in different areas and on various points of each continuum. The shift in approach to domain-specific and topic-specific has resulted in the development of many self-report instruments to measure epistemic beliefs, such as the Internet-Specific Epistemological Questionnaire (ISEQ, Bråten et al., 2005), which was based on Hofer and Pintrich's (1997) four dimensions of personal epistemology.

Within the multidimensional perspective, heavy use of self-reported measures of epistemic beliefs have generated a large body of work. However, these instruments have received criticism for psychometric issues (Sandoval et al., 2016), which undermines the accuracy of the results produced. For example, several researchers that have used the ISEQ have documented associations between self-reported beliefs and behaviours (e.g., Kammerer et al., 2013, 2015; Strømsø & Bråten, 2010). Kammerer and colleagues (2013) found that individuals who believed the internet was a reliable source were less likely to examine source reliability on search engine results pages and were more certain of their search decisions. Further, Kammerer and colleagues (2013) found that individuals who doubted the need to evaluate sources were more likely to have one-sided arguments about the topic. Researchers have had difficulty validating the full structure of the ISEQ, reporting two factors rather than four (justification and general epistemology, Bråten et al., 2005; Knight et al., 2017). The instability of factor structures associated with the multidimensional perspective may be due to operationalizations tacit epistemic beliefs (DeBacker et al., 2008; Sandoval et al., 2016). Issues with operationalizations may, in part, be related to the vague descriptions of beliefs that fall between the concrete and contextual ends of each dimension's spectrum. Greene and colleagues (2008) questioned the appropriateness of measuring complex beliefs on single spectrum, suggesting that the source of knowledge and justification for knowing dimensions are underdeveloped. Psychometric concerns about self-report instruments have led researchers to explore personal epistemology via think-aloud protocols (Greene et al., 2014, 2018) and practical reasoning tasks (Mason et al., 2011, 2014, 2018). Next, philosophically-informed frameworks that can elaborate on the nature of knowledge factors are presented.

Philosophical Perspectives

Stemming from the work of philosophical epistemologists, the philosophical perspective has primarily examined the features of knowledge, its limits and justifications, and the sources of knowledge (Murphy et al., 2012). The nature and limits of knowledge, according to Greene and colleagues (2008), refer to the types of knowledge claims that can be justified as knowledge. Justification can occur through a variety of sources, including perception, memory, introspection, inference, and testimony (Chinn et al., 2011). In contrast to previous perspectives, philosophically-informed theorists have emphasized the role of justification, as opposed to knowledge (Greene et al., 2008). They have also delineated areas that were underdeveloped in previous frameworks, specifically the nature of knowing factors.

Like the multidimensional perspective, Royce (1959) proposed two belief dimensions about knowing: beliefs about how knowledge is derived and about how knowledge is justified. Royce (1959) further explained that there were distinct sub-dimensions about knowing (e.g., knowledge is derived and justified through inference). These sub-dimensions were later transformed into Royce's (1978) three ways of knowing: rationalism, empiricism, and metaphorism. Each way of knowing involved distinct criteria and was associated with a cognitive process. Returning to the example, if Toby were primarily classified as a rationalist by Royce, they would prioritize the use of logical consistency as a criterion to determine whether to accept or reject the knowledge claims encountered online. If Toby were primarily classified as an empiricist, they would prioritize the use of quantitative evidence for credentialled experts about treatments for COVID-19 to make their information evaluations. Lastly, if Toby were classified as a metaphorist, who believed that knowledge is derived through intuition and justified through universality, they would prioritize constructing their own understanding of the issue by

integrating knowledge claims from multiple sources. Given these different beliefs about knowing, and the processes associated with each way of knowing, Royce (1978) suggested that an individual may prioritize one way of knowing above the others but may use all three ways of knowing. Other theorists drawing from philosophical epistemology (e.g., Barzilai & Chinn, 2018) may classify this hierarchical prioritization structure as adapted to the situation.

Barzilai and Chinn (2018) introduced the Apt-AIR framework, an integrated framework that described five aspects of epistemic thinking in terms of individuals' epistemic aims, ideals, and reliable processes. Drawing from their previous multidimensional (Barzilai & Zohar, 2014, 2016) and philosophical work (Chinn et al., 2011, 2014), Barzilai and Chinn (2018) aimed to promote apt epistemic performance, or individuals' ability to competently and reliably seek knowledge (Sosa, 2015), by elaborating on their epistemic thinking. Epistemic thinking includes both EC and epistemic metacognition (Barzilai & Zohar, 2014). Aspects 1-3 of their framework will be emphasized as these relate directly to epistemic thinking, whereas Aspects 4 and 5 relate to individuals' motivational-affective dispositions and participation in social settings, respectively.

In Aspect 1, Barzilai and Chinn (2018) described the cognitive components of epistemic aims, ideals, and processes (i.e., EC). They explained that an individual's epistemic aims are the objectives and importance they set for their cognition or action (e.g., understanding, Chinn et al., 2014). An individual's aims can influence how they process information (Greene et al., 2018). For example, Toby's epistemic aim may be to determine if information from a health website is reliable, which will influence what source and information they select. An individual evaluates whether their epistemic aim has been met through comparison with their epistemic ideals, which describe their domain-specific standards or criteria (Barzilai & Chinn, 2018). Toby may evaluate

health websites based on their level of authority and accuracy because trustworthiness is one of their epistemic ideals. According to Barzilai and Chinn (2018), epistemic ideals influence what reliable epistemic processes are engaged in, including evaluative, creative and communicative processes. Reliable epistemic processes refer to the strategies and skills an individual uses to consistently achieve their objectives and create epistemic products (e.g., an essay). To establish trustworthiness, Toby may examine content-and source-based features, such as the websites' readability and author's credentials, respectively. In Aspect 1, the role of domain-general and domain-specific processes are accounted for, and the foundation is set to describe the importance of adapting epistemic thinking to perform aptly in diverse environments (Barzilai & Chinn, 2018).

In Aspect 2, Barzilai and Chinn (2018) described how contexts, tasks and domains require individuals to engage in distinct types of EC. These situations may influence the types of epistemic aims an individual sets, their epistemic ideals and the processes they engage in. A situated view of EC was elaborated on by Chinn and Sandoval (2018) who argued that epistemic processes may appear similar, but, under scrutiny, differ substantially across domains (e.g., individuals' source evaluations are important in science and history). For example, Toby's epistemic aim to determine the reliability of a health website would be useful for a task in a science communication course and in a history course assignment about pandemics. Yet, they may select similar epistemic ideals and engage in distinct epistemic processes. For both tasks, Toby may select an epistemic ideal of trustworthiness because the tasks require source evaluation. To complete the science communication task, Toby may establish reliability by solely examining the evidence within the website. In contrast, for the history task, Toby may use lateral reading to corroborate and contextualize the website's information. To elaborate on

Aspects 1 and 2, Barzilai and Chinn (2018) described individuals' epistemic metacognition in Aspect 3.

In Aspect 3, Barzilai and Chinn (2018) emphasized the role of individuals' epistemic metacognitive skills and knowledge in regulating their competent epistemic performance. Following Flavell and colleagues (Flavell, 1979; Flavell et al., 2002), Barzilai and Zohar (2014, 2016) defined epistemic metacognitive skills as regulatory processes of their EC that involve planning, monitoring and evaluating knowledge and epistemic processes. Due to the variability of situations described in Aspect 2, individuals' epistemic metacognitive skills support their ability to identify, plan for, control and monitor their progress toward their epistemic aims, their use of appropriate epistemic ideals, and their enactment of reliable processes. For example, Toby's metacognitive skills during the history task may lead them to seek corroboration to determine the reliability of the website because this is an important standard in the domain. Barzilai and Zohar (2014, 2016) also elaborated on epistemic metacognitive knowledge, defining this construct as knowledge, beliefs, ideas, and theories about the nature of knowledge and knowing, which includes knowledge about persons, strategies and tasks. Individuals' epistemic metacognitive knowledge supports their understanding of what is required to achieve specific epistemic aims, of when and why epistemic ideals are important to apply, and of how and why to evaluate the reliability of epistemic processes. Toby's epistemic metacognitive knowledge during the history task may support their understanding of what processes are required to determine if the website is reliable (e.g., corroboration) and how to carry out those processes. Together, Toby's metacognitive skills and knowledge will influence every aspect of their source evaluation—from the generation of their epistemic aims to the reliable processes they enact. The

Apt-AIR framework has addressed important shortcomings that have been presented in previous frameworks, including the distinction between EC and epistemic metacognition.

Whereas Royce's (1979) work maintained a domain-general approach like the developmental and multidimensional perspectives before, Barzilai and Chinn (2018) incorporated a situated perspective of personal epistemology that has been called for by disciplinary theorists (Sandoval, 2012). Unlike previous theorists, philosophically-informed theorists also elaborated on individuals' epistemic processes, including their evaluation of sources and justification of knowledge. This emphasis is aligned with the shift to empirical validation of theory using think-aloud protocols (e.g., Barzilai & Zohar, 2012; Ferguson et al., 2012; Mason et al., 2010) rather than self-reports. This perspective is useful for understanding individuals' EC during source evaluations on the internet and designing empirical investigations. Specifically, Barzilai and Chinn's (2018) framework details the impact of different tasks and contexts on EC, suggesting that findings about individuals' source evaluation behaviours must be situated in the environment findings are intended to understand. Next, the disciplinary perspective which has further examined the situated nature of personal epistemology is presented.

Disciplinary Perspectives

In contrast to the previous frameworks described, the disciplinary perspective emphasized the application of theory to practice and instruction. These theorists have also examined personal epistemology in terms of its domain-general (e.g., Elby & Hammer, 2010), domain-specific (e.g., science, Sandoval, 2005, 2014), and context-specific properties (e.g., multiple text comprehension, Ferguson et al., 2012). Sandoval and colleagues (Sandoval, 2014, 2017; Sandoval & Çam, 2011) classified such investigations as situated, where an individual

employs cognitive resources during an activity. Under the situated perspective, individuals' EC occurs in interaction with the material resources or other individuals involved in a specific situation (Greeno, 2015). Greene (2016) differentiated the relevant situations in epistemic cognition research: individual, individual in interaction, and the system level. The first two levels will be emphasized here. Sandoval (2014) argued that situated experiences inform individuals' development of further cognitive resources to use in subsequent situations. Within the disciplinary perspective, researchers' findings are described as contextualized rather than inconsistent (Sandoval & Redman, 2015). As an example of contextualization, Toby may state that evidence about health information is most reliable from a government or clinical website because they have used these types of websites in previous experiences with similar tasks. However, Toby may provide a lower reliability rating for a government website during the present task because the evidence and justification of the site's viewpoint are inferior to the justifications provided by another source type. Theorists (Chinn & Sandoval, 2018; Elby & Hammer, 2010) and researchers (e.g., Kienhues et al., 2016) have noted that epistemic practices can vary within the same topic as a result of task presentation. The two influential disciplinary frameworks described below focused on tacit and localized beliefs or resources that are activated conditionally (Sandoval, 2014).

To support tailored instruction, Elby and Hammer (2010) argued that epistemological resources were context-specific cognitive resources that an individual may use to understand and reflect on their epistemic knowledge, activities, forms, and stances. Rather than a set of stable beliefs, Hammer and Elby (2002) described individuals' epistemological resources as finer grained than concepts or beliefs, where multiple resources may be activated simultaneously in a local network. As such, epistemic development is demonstrated through a change in activation of

these resources, rather than a progression through stages as seen in developmental frameworks (Hammer & Elby, 2002). Through activation, individuals' epistemological resources were theorized to influence their learning. Hammer and Elby (2010) posited four categories of resources: resources to understand (a) the general nature of knowledge and how it is derived, such as knowledge as propagated stuff, where knowledge comes from an authority figure, (b) familiar activities, such as accumulation, which describes the retrieval of information, (c) the form of an activity, such as an essay, where Toby may call upon specific resources to construct a five-paragraph essay, and (d) stances toward knowledge, such as doubt, which describes when Toby has neither accepted or rejected a piece of information (Hammer & Elby, 2002). Returning to the example, during Toby's online search, their epistemological resources such as accumulation, facts, sorting, propagated stuff, and fabricated stuff may be activated as they determine source reliability. Whereas Hammer and Elby (2010) theorized that cognitive structures could not be characterized due to the context-specific nature of epistemological resources, Sandoval (2005) agreed that practical epistemologies were situated, but that this cognitive structure could also be described.

Sandoval (2005, 2014) asserted that situated theories of epistemic development better account for the array of findings about individuals' expressed ideas about knowledge and knowing and their practices during knowledge construction (i.e., empirical paradox). Due to the situated nature of cognition, Sandoval (2012) argued that EC occurs during joint activity, the construction of epistemic artifacts (e.g., scientific argument), or reflection on these activities. Sandoval (2005) defined practical epistemologies as epistemological beliefs that guide knowledge production. Like philosophically-informed theorists, Sandoval and colleagues have focused on the nature of knowing rather than the nature of knowledge in their empirical

investigations of science argumentation (e.g., Ryu & Sandoval, 2012; Sandoval & Çam, 2011). For example, Ryu and Sandoval (2012) identified four epistemic criteria that they deemed central to understanding scientific argumentation: causal structure, causal coherence, citation of evidence, and evidentiary justification. Like epistemic ideals (Chinn et al., 2014), they described causal structure as the understanding that a scientific argument should present causal claims. Casual coherence refers to how well the claims presented fit together, where a good scientific argument has multiple coherences related to claims and the argument is easy to follow. Citation of evidence refers to the support that is provided to justify claims, where a good scientific argument includes relevant evidence for all claims and this evidence is fully explained in context. Lastly, Ryu and Sandoval (2012) described evidentiary justification, which is the asserted relationship between causal claims and the evidence presented. In Toby's case, these epistemic criteria would be important for an argument about COVID-19 treatments. Consistent with the shift to examine processes in the field, the disciplinary perspective attempts to bridge the gap between theory and practice.

Disciplinary theorists elaborated on the situated nature of individuals' EC which was included, and at times glossed over, in other perspectives. Like criticisms of developmental and multidimensional perspectives, Elby and Hammer's (2010) vague description of epistemological resources has made validation difficult. To compare the four perspectives presented in this literature review, Table 1 highlights the main assumptions and concepts featured in the selected work.

Table 1

Comparison of Theoretical Perspectives

	Developmental		**Multidimensional**		**Philosophical**		**Disciplinary**	
	Kuhn (1991)	Chandler and colleagues (1990)	Hofer and Pintrich (1997)	Bendixen and Rule (2002, 2010)	Royce (1969, 1978)	Barzilai and Chinn (2018)	Elby and Hammer (2002, 2010)	Sandoval (2005, 2014)
Metacognitive, Cognitive or Both	Metacognitive	Metacognitive	Metacognitive	Metacognitive	Both	Both	Both	Both
Main Assumption(s)	Domain-general Stable beliefs	Domain-specific Intraindividual differences	Domain-general Stable beliefs	Unstable beliefs over time	Domain-specific Prioritize one way of knowing	Situated	Situated Unstable beliefs over time	Situated
Key Constructs	1 Realist 2 Absolutist 3 Multiplist 4 Evaluativist	1 Realist 2 Dogmatist – Skepticist 3 Rationalist	<u>Nature of knowledge</u> • Certainty • Simplicity <u>Nature of knowing</u> • Source • Justification	<u>Epistemic change</u> 1 Epistemic doubt 2 Epistemic volition 3 Resolution strategies	A. Rationalist B. Empiricist C. Metaphorist	• Epistemic aims, ideals, processes • Epistemic metacognitive skills • Epistemic metacognitive knowledge	<u>Epistemological resources</u> • Propagated stuff • Familiar activities • Activity form • Stances	Practical epistemologies

Note. In the "Key Constructs" section, numbers are used to represent processes. Letters are used to represent profiles. Bullet points are used to represent components.

Theoretical Concerns

Broadly, concerns remain regarding key constructs, mechanisms of epistemic change, and disparate empirical findings in the field.

Key Constructs

The field of personal epistemology consists of four distinct areas of focus: EC, epistemic metacognition, epistemic motivation, and epistemic emotions (Muis & Singh, 2018). Some theorists have combined several areas (e.g., Chinn et al., 2011, 2014; Chinn & Rinehart, 2016) or included non-epistemic constructs (e.g., Hammer & Elby, 2002). These broader conceptualizations of personal epistemology have been opposed due to conflation concerns. For example, based on Muis and Singh's (2018) classification, Barzilai and Chinn's Apt-AIR framework would more appropriately fall under personal epistemology rather than epistemic thinking because they included elements of epistemic thinking, epistemic motivation (i.e., Aspect 4) and epistemic emotions (i.e., Aspect 5). Given the decentralized state of the field, it is essential to distinguish EC from other epistemic constructs. Similarly, separating epistemic constructs from other cognitive and metacognitive phenomena is important for construct validity and meaningful generation of theory.

Sandoval criticized the epistemological resources framework, stating that "framing learning strategies like 'memorize' as epistemological 'resources' propagates the conflation of beliefs about knowledge with beliefs about or strategies for learning" (2005, p. 649). Learning is related to personal epistemology, but it is not the same construct (Sandoval, 2014). Despite this criticism, Elby and Hammer (2010) maintain that views of learning and views of knowledge and knowing are closely tied. Researchers also continue to include the nature of learning factors in their investigations (e.g., Ricco et al., 2010). Greene and colleagues' (2018) recent meta-analysis

of the relationship between personal epistemology and academic achievement found that inclusion of nature of learning factors has not skewed the empirical literature. However, clearly defining epistemic and non-epistemic constructs during empirical examination of EC is essential for understanding how individuals' EC specifically supports their source evaluation behaviours.

The nature of knowing factors have been the least elaborated on aspects of personal epistemology. Hofer and Pintrich (1997) identified that lower developmental levels were clearly related to aspects of personal epistemology. However Hofer and Pintrich (1997) explained that elements of personal epistemology were more difficult to trace at the more advanced stages,, as seen in Perry's (1970) scheme. This inconsistency has presented issues for operationalization of key constructs. Multidimensional frameworks have also received criticism for their narrow focus on epistemic beliefs, without integration of theory from philosophical epistemology. Their narrow conceptualizations may account for limited descriptions of the source of knowledge and justification for knowing dimensions. Barzilai and Chinn's (2018) Apt-AIR framework has attempted to fill in the blanks about the nature of knowing factors. These factors are the most important for determining the role EC plays in source evaluations and difficulties operationalizing constructs limit the empirical validation of theory. Alongside vague descriptions of these factors, theorists' descriptions of the mechanism of change has also been scant.

Mechanisms of Change

The developmental and multidimensional perspectives have offered limited description of the stimulus for change between levels of thinking or points on a continuum (i.e., epistemic change, Bendixen, 2002; Hofer, 2004; Kuhn & Weinstock, 2002). Schommer (1990) criticized the developmental assumption that a simultaneous shift in beliefs about the nature of knowledge and knowing occurs when an individual transitions from one stage to the next. Disciplinary

theorists have also called into question the coherence assumption that ties many perspectives together (Hammer & Elby, 2002; Sandoval & Redman, 2015). When a perspective is based upon stable beliefs, Hammer and Elby (2002) asserted that consistency is presumed. Disciplinary theorists suggest that variation in empirical findings discount the coherence assumption, classifying this widespread aspect of theorists' work as unwarranted (Sandoval & Redman, 2015). Despite such assertions, disciplinary theorists have not adequately described how epistemological resources or practical epistemologies can be developed to aid instruction either.

As described in Bendixen and Rule's (2004, 2010) integrative model, change requires epistemic doubt, where the individual gains more independence and is exposed to beliefs that conflict with those that are currently held. Epistemic change also requires directed effort and resolution strategies to achieve the desired change. Research has demonstrated changes in individuals' epistemic processes through instructional interventions (Muis & Duffy, 2013) lasting several months after initial assessment. These researchers induced epistemic doubt by asking students to engage in activities where they compare different approaches to problem-solving or examine multiple viewpoints. Beyond activities as catalysts for cognitive dissonance, Muis and colleagues (Muis et al., 2016; Muis & Duffy, 2013) have identified the environment, social interactions, and discourse as crucial aspects of epistemic change. Since personal epistemology has primarily focused on the individual's experience, it is understandable why these critical social elements have not been factored into many frameworks. Further development of Sandoval's (2005, 2014) situated view of personal epistemology and Barzilai and Chinn's (2018) Apt-AIR framework are promising avenues to better understand EC during interactions with others and materials. Overall, concerns about key constructs and how they relate to one another may account for the disparate findings within the field about the nature of knowledge

and knowing. Given the frameworks presented in this section, it is important to distinguish aspects of EC from other constructs.

Distinguishing Epistemic Cognition from Other Constructs

From the first appearance of the term EC (Kitchner, 1983), theorists have embedded personal epistemology into existing cognitive structures. At times, they have done so inappropriately. According to Hofer (2016), EC describes the mental processes involved in the acquisition, justification, and use of knowledge. Muis and Singh (2018) further differentiated these mental processes by setting a clear boundary between EC and epistemic metacognition. Individuals' EC is limited to setting epistemic aims, enacting epistemic strategies, engaging in reliable processes to attain epistemic achievements, and creating artifacts (e.g., written argument). Their EC encompasses conscious activity, such as thinking, reasoning, or attending to information. For example, during Toby's knowledge acquisition, they set epistemic aims and select epistemic ideals and reliable processes based on the epistemic aim choosen. During justification and use of knowledge, Toby enacts selected ideals and processes to reach their goals and create an artifact. Here, Toby's EC deals with the evaluation of others' explanations when deciding whether to accept a knowledge claim or the use of knowledge to justify their own claim in an argument. Through empirical examination, researchers have found that individuals' EC can be developed to improve source evaluations (e.g., Brante & Strømsø, 2018).

Whereas EC consists of explicit thinking, epistemic metacognition includes knowledge, skills, and experiences, such as epistemic beliefs, epistemological resources, and epistemic ideals. These subcomponents account for an individual's beliefs about the features, limits, explanations, and sources of knowledge, which will influence their EC. Epistemological resources also provide metacognitive knowledge about knowledge, activities, forms, and stances.

Epistemic metacognitive knowledge of persons or strategies and epistemological resources remain metacognitive, rather than cognitive, until the individual activates that knowledge for use as an epistemic process. As a result of this distinction, theoretical and empirical work examining epistemic metacognition should not fall under the umbrella of EC, but rather personal epistemology. There is a growing body of appropriately labelled research on the role of epistemic metacognitive knowledge and skills in source evaluation and knowledge construction (e.g., Barzilai & Ka'adan, 2017; Barzilai & Zohar, 2012). That is, how epistemic metacognition influences EC. Further research is required to ensure measurements of individuals' EC in real-time are effective and do not fall back on manifestations of epistemic beliefs.

Epistemic motivation and epistemic emotions are evoked by cognitive processes during knowledge construction and justification; however, these epistemic constructs are not cognition and should not be classified as such under the term EC (Muis et al., 2018). The desire to unify a fragmented field under the term EC has not improved the field of investigation's consistency or coherence as Alexander (2016) called for. To resolve some of the conceptual concerns discussed in this section, empirical validation of theory is required to better understand the nature, structure, and function of EC.

Conclusions

This section has reviewed prominent theorists' conceptualizations of personal epistemology and identified contributions and limitations for each perspective. Drawing from theory, the relationship between personal epistemology and learning has also been elaborated upon. Learning serves as a useful context to examine EC because individuals are often required to evaluate and integrate information from divergent sources. Given the affordances of internet access as a place to discover and distribute knowledge and information, the ability to effectively

sort fact from fiction is a crucial aspect of life in the digital age. Epistemic cognition has the potential to support this process. The next section will examine the role of EC in source evaluations.

Part 2. What role does epistemic cognition play in source evaluations?

Supporting the ability to assess the quality of information has become an important educational objective. Although this ability is not novel to the digital age, the interactivity between users and the abundance of available information distinguishes today's high epistemic stakes from previous periods. Eshet (2012) argued that individuals' evaluation skills underlie nearly all decisions they make in digital environments (e.g., data queries, navigational decisions). On a grander scale, Barzilai and Chinn (2018) described the long reaching influences of these skills on society—effective evaluators base their beliefs and behaviours on evidence. Accordingly, Toby's ability to examine information about COVID-19 treatments will impact their own health decisions (e.g., choosing whether to receive a treatment based on beliefs) and their voting decisions (e.g., supporting candidates that do not promote evidence-based health policies). Undoubtedly, the potential impact of improving source evaluations has stimulated interventions among educators (e.g., The News Literacy Project, 2020; Stanford History Education Group, 2020) and researchers (e.g., Clark et al., 2017; Mason et al., 2014; Wiley et al., 2009; Zhang & Duke, 2011). Despite the high stakes of poor source evaluation skills on the internet, education researchers have primarily investigated EC in controlled environments (e.g., offline or multiple documents context). Thus, their findings may not accurately reflect thinking about knowledge claims, sources, and evidence found online. The purpose of the following section is to review empirical evaluations of EC during source evaluations in both controlled and authentic environments.

Evaluating Sources on the Internet

Based on Barzilai and Chinn's (2018) Apt-AIR Framework, EC is activated as individuals set epistemic aims, select ideals, and enact reliable processes to assess if ideals are met. As artifacts, they can also create arguments using reliable epistemic processes. Whether researchers have used the term EC or not, they have examined these three aspects of EC in hopes of better understanding how their participants evaluate information. Given that learning has been a useful context to investigate spontaneous and prompted EC, researchers in various fields have utilized search engine result pages and website evaluations as a context for their investigations. Researchers have designed numerous tasks to examine thinking, including tasks that ask participants to identify the source type or viewpoint (e.g., Barzilai & Eshet-Alkalai, 2015; Gui & Argentin, 2011), evaluate the reliability of information of multiple sources (e.g., Mason et al., 2014), or compare biased information (e.g., Alkali & Amichai-Hamburger, 2004). These contexts and tasks have exhibited varying degrees of authenticity with relevant investigations ranging from examining multiple documents (e.g., Ferguson et al., 2012) to examining authentic webpages on the internet (e.g., Greene et al., 2014, 2018).

Within personal epistemology, researchers have primarily ignored key aspects of digital contexts, such as the option to corroborate information with external sources or use context specific criteria to justify reliability ratings. Instead, these researchers have primarily drawn from multiple document research (e.g., Perfetti et al., 1999; Rouet et al., 2017) and reading research (e.g., Leu et al., 2017) to explore EC during source evaluations. In practice, these researchers have focused on the literacy aspect of source evaluations (i.e., vetting and integrating information). To better illustrate researchers' emphasis on literacy, in the *Handbook of Epistemic Cognition*, only one chapter describes the role of EC in online source evaluations. This chapter,

written by two multiple document researchers (Strømsø & Kammerer, 2016), was limited to a discussion of reading for understanding. As a result of this limited conceptualization, personal epistemology researchers have utilized elements found in digital environments (e.g., blogs, Barzilai et al., 2015; Barzilai & Eshet-Alkalai, 2015), but have primarily focused on comprehension-based outcomes when exploring individuals' EC. Whereas multiple documents theory and research is well supported for use in multiple documents contexts, key elements novel to online experiences may not be accounted for. To explore the impact of this limited conceptualization in personal epistemology research, empirical work from multiple fields will be examined.

Empirical Investigations of Epistemic Cognition on the Internet

In September and October 2020, ten databases (PsycINFO, ERIC, Communication and Mass Media Complete, FRANCIS, JSTOR, SocINDEX, International Bibliographic of the Social Sciences, ABI/INFORM Collection, Web of Science Core Collection and Scopus) were searched using the keywords "web AND credibility" to conduct a broad survey of recent literature on source evaluations. Records were initially limited to peer reviewed journal articles published after 2000 in English. Thirty-five relevant papers from a similar search conducted in 2019 were carried over to the review. The present search process identified 2,135 publications. After pre-screening records for relevance based on their title and removing duplicates, 346 records remained. Abstracts were examined to remove studies that did not measure individuals' EC on the internet, resulting in 190 remaining articles. These full text articles were reviewed for measurement of EC during a webpage evaluation task, and articles were excluded if they (a) did not measure EC during source evaluation (n = 54), (b) used a social media or user-generated webpage task (n = 24), (c) analyzed the webpage contents (n = 15), (d) used another type of

learning environment (n = 9, Sabourin et al., 2013), (e) were not in English (n = 12), or were not empirical (n = 3). Seventy-three articles were selected for detailed review. Findings are discussed across four categories: individual differences, evaluation criteria, intervention studies, and source integration. Studies with post-secondary students were prioritized; however, studies with younger age groups were included when no equivalent was found in the age group of interest.

Individual Differences and Source Evaluations

Researchers have explored EC by mapping the spontaneous criteria participants use to assess the reliability or trustworthiness of a source and its claims. Barzilai and Zohar (2014) defined source evaluation as an epistemic strategy that individuals use to specifically examine the epistemic properties of a source. These evaluations include paying attention to the quality of the source, such as its surface-level features (e.g., author, venue, date of publication, Bråten et al., 2009), differences between sources (e.g., Flanagin & Metzger, 2007), or potential biases (e.g., Eshet-Alkalai & Chajut, 2009). As a result, individuals may use their source evaluation to interpret the evidence presented (Britt & Aglinskas, 2002; Wineburg, 1991). Researchers have specifically explored individuals source evaluations by observing how they evaluate and justify their evaluations of sources. Three studies with varying levels of authenticity are highlighted.

Drawing from previous multiple documents research, Barzilai, Thomm, and colleagues (2020) investigated university students' epistemic processes as they rated the trustworthiness of dissenting expert accounts. In 2 studies, the researchers asked participants to read two texts written by experts with distinct perspectives on the same topic, rate and justify their agreement with the text's claim, and rate and justify their ranking of the author's trustworthiness. Each participant was either assigned to read texts about a familiar topic or an unfamiliar topic. In Study 1, the researchers found that participants' source ratings were dependent on the source's

position and their level of familiarity with the topic. Unsurprisingly, participants rated sources that they agreed with as more trustworthy, primarily relying on knowledge validation strategies to evaluate the trustworthiness of the familiar topic and on discourse-based strategies to evaluate the unfamiliar topic. Many researchers have documented the tendency to identify belief-consistent information as more credible (e.g., Meppelink et al., 2019). In Study 2, Barzilai, Thomm and colleagues (2020) modified the author's credentials to make their expertise more explicit and the titles of the familiar topic texts. The researchers also added a third text that described reasons for expert disagreement. Despite these changes, they found that participants exhibited similar processes for each text pairing. However, in Study 2, the researchers observed that participants used fewer sourcing strategies to evaluate source claims in the familiar topic than in the unfamiliar topic, suggesting that individuals do not engage in evaluation processes when claims and sources present belief-consistent information. Irrespective of the environment source evaluations are made in, individuals hold beliefs about the topic which impacts their ability to evaluate source trustworthiness. Many researchers have documented the strong influence of individuals' beliefs in multiple documents contexts (e.g., Maier & Richter, 2013, 2016; Richter & Maier, 2017; Van Strien et al., 2014), and specifically, in online source evaluations (e.g., Meppelink et al., 2019). To mitigate the impact of beliefs on source evaluations during statistical analyses, some researchers have included measures of topic-specific beliefs (e.g., Taylor et al., 2015).

Jung and colleagues (Jung et al., 2016) examined the influence of source expertise and message accuracy on university students' perceptions of webpage credibility in an offline environment. Participants were assigned to read two texts with either high or low source expertise cues and with either high or low message accuracy. The researchers found that

participants rated sources with high message accuracy as more credible than inaccurate sources. Further, participants who had low prior knowledge about the topic rated accurate sources with no expertise cue and inaccurate sources with high expertise cues as more credible than the other respective texts, suggesting the importance of prior topic knowledge in source evaluations. Lucassen and Schraagen (2013) found that individuals with topic familiarity were less influenced by source cues, which may account for Barzilai, Thomm and colleagues' (2020) finding that individuals with prior topic knowledge relied on knowledge-based validation strategies to directly evaluate the quality of familiar information (Richter & Schmid, 2010a). To evaluate reliability, source cues may be more influential when looking for unfamiliar information on the internet; however, individuals with low prior knowledge may have more difficulty assessing the quality of sources, and subsequently accept inaccurate information. To better understand the role of prior knowledge in source evaluations, these constructs has been included in theory (Lucassen & Schraagen, 2011, 2013) and in research designs (e.g., as a covariate and outcome variable, Greene et al., 2014, 2018).

Cho and colleagues (2018) explored upper-secondary students' epistemic aims and processes in an authentic learning activity. Participants were prompted to think aloud as they made epistemic judgments about a variety of information sources on the internet. The purpose of participants' search was to pose a meaningful question about the topic and justify the importance of their question. The researchers found that participants engaged in distinct levels of EC to assess source authority, ranging from acritical processing to highly analytical assessments of the author's purpose and publishing type. They observed highly related source evaluation processes such that participants who noticed surface-level features also tended to assess source authority and those who made these types of assessments tended to critically analyze knowledge claims

and reconcile multiple perspectives. In addition to exploring EC, the researchers investigated participants' epistemic metacognitive knowledge and inferred their epistemic monitoring and regulation of the task. Unsurprisingly, they found substantial associations between participants' epistemic cognitive and metacognitive processes such that participants who made critical assessments espoused beliefs that knowledge is complex and contextual. To better understand the highest and lowest performers' epistemic processes, the researchers characterized each groups' behaviours and found that the highest performers sought to examine sources that differed in reliability. They also employed epistemic planning and monitoring to integrate multiple perspectives and examine the accuracy of knowledge claims and sources to establish reliability. Conversely, the lowest performers less critically evaluated sources and knowledge claims and struggled to plan and implement ways to fill in their knowledge gaps. The relationship between EC and epistemic metacognition has been extensively explored by Barzilai and colleagues in theoretical (Barzilai & Zohar 2014, 2016; Barzilai & Chinn, 2018) and relevant empirical work (Barzilai & Ka'adan, 2017; Barzilai & Zohar, 2012).

Whether occurring on- or off-line, source evaluations are profoundly influenced by beliefs, prior knowledge, and epistemic metacognition. Multiple documents researchers have suggested that individuals often do not consider source or justification features when evaluating the reliability of new or belief-inconsistent knowledge (e.g., Bråten et al., 2016; Wineburg, 1991), which may be accounted for by findings about the influence of prior knowledge and beliefs source evaluations. Although individual differences may similarly influence source evaluations in online and offline environments, according to situated theorists, cognitive processes will differ. As Sandoval (2012) explained, diverse empirical findings are contextual,

rather than inconsistent. Individual differences may help identify patterns of evaluations criteria use. Next, findings about evaluation criteria are reviewed.

Evaluation Criteria

Researchers have documented varied use of evaluation criteria and epistemic processes to assess reliability, an important epistemic ideal for online source evaluations. Perceptions of trustworthiness and source credibility have been widely used as indicators of reliability. Under these broad terms, researchers have observed individuals attend to content-based features, such as comprehensibility, as well as epistemic features, such as source expertise and evidence (e.g., Braasch et al., 2013). Typically, using controversial socio-scientific topics, researchers have measured evaluation criteria through think aloud protocols (e.g., Cho et al., 2018; Greene et al., 2014, 2018) and written responses (e.g., Halverson et al., 2010). To highlight the variety of criteria employed during source evaluations on the internet, three studies conducted online are reviewed.

Greene and colleagues (2014, 2018) investigated the role of self-regulated learning and epistemic processes in online learning by asking undergraduate and graduate students to think-aloud as they learned about a socio-scientific topic on the internet. Think aloud data were coded for indicators of their self-regulated learning and EC. Indicators of EC were reduced to epistemic aims, source evaluations, and methods of justification. Overall, EC was observed much less frequently than self-regulated learning processes, with source evaluation the most prominent of the epistemic processes. In this study, source evaluation was limited to investigations of the source and perceived author bias, which may be due to participants' verbalizations or the researchers' coding schemes. However, individuals' low source evaluation behaviours have been documented by other researchers (e.g., Barzilai et al., 2015; Mason et al., 2011). Greene and

colleagues (2018) conducted a follow up study altering the task focus from an epistemic aim of knowledge to an aim of understanding. With this change, the researchers observed students engage in EC more frequently than in their previous work. These findings, along with findings on task instructions (e.g., Kienhues et al., 2016), supports Barzilai and Chinn's (2018) situated view of personal epistemology such that different situations require different epistemic processes.

Barnes and colleagues (2003) investigated differences in adults' beliefs about the importance of reliability criteria and their use of that criteria. Participants were first asked to rank-order 12 evaluation criteria (e.g., currency, source authority, design) by perceived importance and then they were asked to apply these criteria to rate the quality of authentic websites. The three websites were classified as high-, medium-, and low-quality. When rank-ordering the criteria, participants rated accuracy and expertise criteria as most important, whereas design features and ease of use were rated least important. However, in practice, participants primarily utilized currency criteria (e.g., design and aesthetic) to select high-quality information. Many researchers have found that individuals use both content-based and epistemic criteria to establish reliability with varying degrees of success (Mason et al., 2011; Wiley et al., 2009; Zhang & Duke, 2011). Despite researchers' favourable position toward epistemic criteria, no clear relationship between individuals' criteria use and their source evaluations exist. To complicate matters further, Halverson and colleagues (2010) found participants using epistemic criteria in a troubling way.

Halverson and colleagues (2010) asked university students to draft reports on a socio-scientific topic to assess how they selected, evaluated and used online sources. The researchers observed participants limited evaluation processes as they selected the most used websites to

answer the report prompt. However, to establish reliability, more than half the students assessed the source's credibility, followed by the accuracy, objectivity and/or perspective of information presented. Despite the prevalence of these epistemic criteria in students' reports, the researchers found that many participants described the sources they selected as objective and credible even when the websites provided biased data about the topic, which the researchers asserted was due to students' topic-specific beliefs. The researchers also frequently observed participants apply the content-based criteria, readability, during their source evaluations. Despite using a variety of terminology, the criteria observed by Halverson and colleagues (2010) have also been documented with similar frequencies in larger datasets (e.g., Kąkol et al., 2017) and smaller, qualitative studies (e.g., Subramaniam et al., 2015). Table 2 lists the epistemic and non-epistemic evaluation criteria documented by researchers. The terminology used by researchers is captured in the micro-level categories and the broader criteria these categories refer to are summarized with macro-level categories.

Of the 25 articles that captured evaluation criteria, most researchers observed individuals employ important epistemic criteria to some degree, including indicators that they examine the authority and accuracy of information. Unsurprisingly, many studies that have manipulated information sources have altered these features (e.g., E. H. Jung et al., 2016; W. S. Jung et al., 2018; Mason et al., 2011; Thon & Jucks, 2017). Alongside authority and accuracy, researchers documented individuals' reflections on the content of the information. Of the top three macro-level categories, content-focused features, such as design and relevance, were observed most frequently (e.g., Kiili et al., 2008; Ulyshen et al., 2015). Yet, some researchers also observed high rates of epistemic criteria use (Halverson et al., 2010; Kąkol & Nielek, 2015). To improve

EC during source evaluations, researchers have developed interventions that support important epistemic processes.

Source Evaluation Training

Researchers have designed instructional interventions (e.g., Hämäläinen et al., 2020; Nokes et al., 2007; Wiley et al., 2009) and computer-based tutorials (e.g., Barzilai, Mor-Hagani, et al., 2020; Britt & Aglinskas, 2002; Pretorius, 2018) to train students on source evaluations, integrations, or a combination of both these skills. Using short- and longer-term designs, researchers have conducted interventions within regular classroom activities (e.g., Clark et al., 2017) or as distinct courses (e.g., McGrew et al., 2019). They have used a variety of methods to train students on source evaluation skills, including declarative knowledge (e.g., Barzilai & Ka'adan, 2017; Mason et al., 2014; Wiblom, 2019), contrasting cases (Braasch et al., 2013; Bråten et al., 2019; Hämäläinen et al., 2020), co-construction of criteria and practice (Barzilai, Mor-Hagani, et al., 2020; Clark et al., 2017; McGrew et al., 2019; Stadtler et al., 2016; Walraven et al., 2013). Some researchers have criticized the use of declarative knowledge or checklist approaches to source evaluation instruction (e.g., Breakstone et al., 2018; Menchen-Trevino & Hargittai, 2011), explaining that this method supports blind skepticism. To address this concern, some researchers have trained students on heuristics instead (e.g., Britt & Aglinskas, 2002; McGrew et al., 2019; Nokes et al., 2007). Regardless of approach, interventions focused on authority and accuracy criteria. Three interventions aimed at improving source evaluation skills are highlighted. Further, Table 3 summarizes the skills individuals were trained on and the criteria they used during their source evaluations.

Table 2

Evaluation Criteria Identified from the Literature

Macro-level categories	**Micro-level categories**	**Articles** (%, *n*)
Accuracy of information[1,2,4,6-10,12-17,19,20-24]	Able to corroborate information, quality of evidence, accuracy of evidence, presence of references, quality of references, links to noncredible sites, third party endorsement	84%, *n* = 21
Author transparency[1,15,19]	Availability of contact information	12%, *n* = 3
Author or publisher's purpose[1,2,4,6-8,12,13,15,17,19,22,24]	absence of bias, absence of advertisements, author's purpose, objectivity	52%, *n* = 13
Authority of source[1-5,7-10,12-15,17-21,24]	Author expertise, author's notoriety, venue reputation	76%, *n* = 19
Content of webpage[1,4,6-8,11,13,15-17,19,22,24]	Comprehensibility, correct citation practices, graphics, media use, readability, scope of content, uniqueness, writing quality	52%, *n* = 13
Currency of information[1,3,7,11,14,15,17,19,24]	Freshness of data, publication date, up-to-dateness	36%, *n* = 9
Design and aesthetics[1,4-8,11,13- 15,17,19,22,23]	Layout, appeal, presentation	56%, *n* = 14
Links[1,4,8,14,15]	Number of links, broken links, outbound links, presence of hyperlinks	20%, *n* = 5
Relevance of information[3,6,11,14,16,17]	Topicality, usefulness, type of webpage	24%, *n* = 6
Webpage usability[1,4,13,14,17,22]	Ease of use, accessibility	24%, *n* = 6
Individual differences[7,18,21]	Site familiarity, user goals, personal beliefs, interest, confidence in claims	12%, *n* = 3
Miscellaneous[1,4,13,14]	Download time, popularity, social presence, intended audience	16%, *n* = 4

[1]Barnes and colleagues (2003)
[2]Barzilai and Zohar (2012)
[3]Cheng and Tsai (2020)
[4]Chiagouris and colleagues (2008)
[5]Cugelman and colleagues (2009)
[6]Cunningham and Johnson (2016)
[7]Dochterman and Stamp (2010)
[8]Eysenbach (2002)
[9]Ferguson and colleagues (2012)
[10]Flanagin and Metzger (2007)
[11]Gerjets and colleagues (2011)
[12]Greene and colleagues (2014, 2018)
[13]Halverson and colleagues (2010)
[14]Iding and Klem (2005)
[15]Kąkol and colleagues (2017)
[16]Kalichman and colleagues (2006)
[17]Kiili and colleagues (2008)
[18]Lombardi and colleagues (2014)
[19]Madden and colleagues (2012)
[20]Mason and colleagues (2010)
[21]Meyer and colleagues (2010)
[22]Robertson-Lang and colleagues (2011)
[23]Starling and colleagues (2018)
[24]Ulyshen and colleagues (2015)

Drawing from a multiple documents approach, Britt and Aglinskas (2002) examined secondary school and college students' sourcing skills in three studies. In Study 1, students were randomly assigned to one of two conditions: one group received comprehension-focused instructions and the other group received sourcing instructions to attend to source features. Participants were given printed-out texts to examine and asked to answer questions about key aspects of the texts, evaluate the texts, and create arguments on two topics covered in the texts. Overall, the researchers found that source evaluation skills were low for both groups and at both academic levels. Participants were also observed inappropriately using evidence to answer skills questions. The researchers developed and implemented a computer-based tutorial aimed at improving source evaluation skills in studies 2A and 2B. Like Study 1, participants were split into two groups: one group received training and the other served as a control group. The training group learned about sourcing, contextualization, and corroboration and practiced these important domain-specific heuristics over a two-day period. Both groups completed a similar task to Study 1 where they read printed-out texts and answered questions as an indicator of their sourcing ability. To assess performance, the researchers added the participants' mentions of source features in notes to the number of questions they answered correctly. In both studies, the training group performed better on the skills assessment, suggesting that the tutorial was effective. In Study 3, both groups received the same materials presented in two formats: as distinct documents in the tutorial or as one document in the control condition. After examining the document(s), participants were asked to complete the skills assessment and draft a short essay about the historical topic. Like the previous studies, the training group outperformed the control group on the skills assessment. In essays, both groups mentioned an equal amount of background information from the document(s) they read; however, the training group provided

more source-based information from primary and secondary sources. The researchers explained that this finding may be due individuals' differentiation of documents when presented separately (e.g., salience). Britt and colleagues' later work (2013) has described how the boundaries of documents (and webpages) are blurred due to similar formatting, which makes source evaluations more difficult. Also using a multiple documents approach, Salmerón and colleagues (2018) have argued that reading real documents (e.g., textbook), rather than printed-out ones, provides haptic experiences that may support their source evaluation skills. In accordance with the situated perspective of personal epistemology, Salmerón and colleagues (2018) posited that individuals' abilities may have been underestimated due to researchers' methodological choices.

Through a short-term instructional intervention, Mason and colleagues (2014) provided secondary school students with declarative knowledge about online source evaluations. The researchers assigned participants to one of two conditions, where they either received written instruction based on Wiley and colleagues' (2009, Study 2) SEEK material or no training on source evaluation. The SEEK (Source, Evaluation, Explanation, Knowledge) material instructed participants to assess source authoritativeness (e.g., author's expertise or purpose) and content accuracy (e.g., plausibility, corroboration). Over two sessions, participants in both groups were asked to read sources from an offline digital environment that ranged in authoritativeness and viewpoint, rank the reliability of each website, and provide a justification for their ranking. Like Barzilai, Thomm, and colleagues' (2020) study described in the previous section, participants also constructed a written argument that was coded for their argument structure. Beyond prior knowledge and argumentative reasoning skills, Mason and colleagues (2014) found that participants who received training were more accurate when rank-ordering the most reliable websites, and they used SEEK criteria to justify their ranking for the most and least reliable

websites. However, in the first session, there were no differences between the groups' overall ranking accuracy. In the second session, the training group performed better during overall ranking of websites, and they demonstrated better source integration in their written essays. Specifically, participants in the training group with higher prior knowledge produced more complex arguments. Mason and colleagues' more recent work (2018) has found that individuals struggle with identifying reliable and unreliable sources, even if they can provide a relevant justification for their ranking.

McGrew and colleagues (2019) incorporated training into a university-level critical thinking course that aimed to foster students' evaluation and synthesis of evidence. The researchers randomly assigned course sections to the training or control condition. Using a pre/posttest design to assess the efficacy of the training, all participants were asked to complete four tasks before and after the training was scheduled to occur. Providing full access to the internet for each task, the researchers had students assess the reliability of a single article and then evaluate two websites to determine which provided a better starting point for research on the topic. The participants also completed tasks that focused on source evaluations on social media. After the pretest, students in the training group received two 75-minute lessons where they were introduced to and practiced lateral reading and corroboration strategies, such as scrutinizing the author and the evidence provided in an article by conducting independent web searches to assess the veracity of claims. The control group received instruction as usual. To examine participants' improvements from pretest to posttest, their responses on all tasks were coded as beginning, emerging, or proficient to calculate a composite score. The researchers found that the training group was twice as likely to perform better than the control group on the

four posttest tasks. Unfortunately, the researchers did not provide information about participants' changes on each task, but instead provided overall comparisons.

These three intervention studies have demonstrated that effective source evaluation and integration skills can be developed and implemented with immediate improvements. Table 3 provides additional examples of intervention designs with findings. Of the 14 intervention articles that captured source evaluations and/or integrations, most researchers reported improvement on one or more skills participants were trained on. Primarily using rank-ordering and justification tasks, researchers demonstrated that various training methods can improve individuals' ability to differentiate between reliable and unreliable sources. However, Barzilai and Ka'adan (2017), who focused on source integration after evaluation, were the only researchers to include a delayed task and demonstrate training effects after four weeks. Next, examinations of source integrations and arguments are reviewed.

Source Integration and Argumentation

Since source evaluation is an important antecedent to their effective integration of information (Rouet & Britt, 2011), researchers have also examined how knowledge is constructed after evaluation (i.e., creative processes, Barzilai & Chinn, 2018). This epistemic process occurs when individuals synthesize, create and communicate information from multiple sources (Barzilai et al., 2018). Relationships between sources may be implicit and require individuals to actively construct, rearrange and combine information from multiple sources to generate a reasonable conclusion. Some researchers have related source integration to EC in the context of comprehension and argumentation (e.g., Bråten et al., 2013; Mason et al., 2014; Wiley et al., 2009). The following studies illuminate how evaluation may influence integration.

Table 3

Illustrative Examples of Interventions with Findings

Researchers	Intervention	Relevant Measures	Findings
		Printed-out documents	
Nokes et al. (2007) Secondary school students	Four instructional interventions were developed to examine the influence of text type (traditional text vs. multiple texts) and instruction type (content instruction vs. heuristic instruction).	• Written response (accuracy) • Rank-ordering (reliability, usefulness)	Students that used multiple texts used more corroboration than those who used textbooks. Further, students that used multiple texts to study heuristics were better at sourcing than all other students. Students in all groups rarely used contextualization as a heuristic.
Braasch et al. (2013) Secondary school students	Three modules where students contrasted two student strategy protocols for evaluating a source were developed to represent sophisticated and acritical source evaluation strategies.	• Rank-ordering (usefulness) • Justification of rank-ordering • Written response (comprehension)	Students that participated in the contrasting cases modules were better able to discriminate between more and less reliable sources. They also incorporated more source features to justify their rank-ordering decisions. Students that receive training also included more correct conceptions from the more reliable documents in their essays.
		Offline environment	
Wiley et al. (2009) Secondary school students (Study 2)	Three-page declarative SEEK material developed that explained what and how to attend to source features during source evaluations.	• Rank-ordering (reliability) • Justification of rank-ordering • Written response (comprehension)	Students that read the declarative materials were better able to differentiate between more and less reliable sources and explain their reasoning for rank-ordering decisions. They were also more likely to have integrated essays with fewer erroneous conceptions.
Barzilai & Ka'adan (2017) Secondary school students	Two experimental scaffolds (strategic & metastrategic) were developed to support students' source evaluations and integrations.	• Written response (integration)	Students that received either experimental scaffold improved their integration skills in an immediate and 4-week delayed task. Students in the metastrategic group improved immediately, whereas students in the strategic group over time.

Researchers	Intervention	Relevant Measures	Findings
		Online environment	
Pretorius (2018) University students	Online self-discovery tutorial developed to teach to students to examine content-based and epistemic features of sources.	• Reliability scores • Justification of reliability score	At the beginning of the tutorial, most students incorrectly believed that a personal blog was reliable, justifying their decision on relevance and subjective criteria. Most students learned to identify a reliable source by the end of the tutorial and use relevance and epistemic criteria to justify their decisions.
Hämäläinen et al., (2020) Elementary school students	Three modules which consisted of explicit teaching, modeling, practicing and/or reflecting on searching, evaluating and synthesizing information were developed.	• Justification of reliability scores • Written response (justification)	Students that received training better justified their reliability scores using references to the source and other features, such as the date or website design. Both groups included equal references to argument or content-based features in their justifications. Students in both groups included similar amounts of justifications in their essays.

Using a multiple documents approach, De La Paz and Felton (2010) examined the efficacy of training secondary school students to examine authors, biases and evidence in preparation to plan and draft an argumentative essay. All participants were asked to examine the reliability of historical texts and construct an argument on the topic. The researchers assessed the structure and quality of the written arguments using Toulmin's (1958) method, which focuses on knowledge claims. De La Paz and Felton (2010) calculated each participant's total number of claims and examined the development of each claim in their argument. The researchers found that the training group was more likely to score better and to have developed their claims. Further, the training group was more historically accurate and elaborated more on both their claims and rebuttals. Other researchers that have conducted interventions have emphasized different elements of Toulmin's (1958) method in their coding of argument structures (e.g., reconciling multiple perspectives, Barzilai, Mor-Hagani, et al., 2020; Mason et al., 2014).These researchers found that participants were better able to integrate sources with scaffolds or after training, respectively.

Emphasizing the role of comprehension in source integration, Barzilai and Eshet-Alkalai (2015) documented university students' argumentation in the context of examining either conflicting or converging perspectives on a controversial topic. The researchers presented participants with expert blogposts in an offline environment and asked them to rate the reliability of each blog on a Likert scale, identify the author's viewpoint on the topic, justify their rating of the source's reliability, and construct an argument on the topic. Participants' justifications were coded for mentioning author viewpoints as an indicator of reliability, and their arguments were coded for argument structure (e.g., one-sided argument) and source use. The researchers found that participants' viewpoint comprehension positively impacted their source integration. As

described in Table 3, Barzilai and colleagues' subsequent work (Barzilai, Mor-Hagani, et al., 2020) designed epistemic scaffolds based on the principles of the Apt-AIR framework (Barzilai & Chinn, 2018). The training group used slightly more evaluation criteria and outperformed the control group on the integration task, with most participants forming two-side arguments and reconciling both perspectives on the topic.

These studies have demonstrated that individuals' EC is related to their integration of multiple sources. None of the selected articles examined both EC and source integration in a written argument using an online environment. Education researchers often include measures that can illuminate the relationships between evaluative and creative epistemic processes; however, their emphasis on curated learning environments has undermined understanding these processes in authentic environments. Next, contributions and limitations of the reviewed literature are presented.

Discussion and Future Directions

Researchers' divergent findings using differing tasks, instructions and learning environments have supported a situated approach to investigating EC during online source evaluations. Barzilai and Chinn's (2018) Apt-AIR framework provides a valuable structure to summarize results of this literature review. Beginning with epistemic aims, researchers have demonstrated individuals' difficulty identifying epistemically valuable products through rank-ordering tasks (e.g., Mason et al., 2014, 2018). However, those who can identify meaningful questions seek sources that differ in reliability and employ epistemic metacognitive planning and monitoring to establish reliability (Cho et al., 2018). During source evaluations, individuals primarily rely on non-epistemic criteria to evaluate sources (Kiili et al., 2008; Ulyshen et al., 2015), even if they report that epistemic ideals are more important (Barnes et al., 2003) or

frequently refer to these criteria (Kąkol & Nielek, 2015; Mason et al., 2018). Individuals have demonstrated different levels of scrutiny for distinct source types (Flanagin & Metzger, 2007; Pretorius, 2018). Unfortunately, they may not know what and how to appropriately use important epistemic ideals, such as objectivity and bias (Halverson et al., 2010). On and off the internet, interventions have successfully supported EC in source evaluations and integrations (McGrew, 2020; McGrew et al., 2019; Pretorius, 2018).

Whereas Barzilai and Chinn's (2018) framework acknowledged that individuals perform differently as a result of their task and context, their description and illustrative examples for examining website reliability do not delineate what and when epistemic processes are appropriate to achieve epistemic ends. Broadly, when individuals examine the reliability of a website, the epistemic properties they use as indicators of trustworthiness are authority, accuracy, and purpose (see Table 3 in Barzilai & Chinn, 2018). Barzilai and Chinn's (2018) vague descriptions of source evaluations may be accounted for by their previous research using offline environments (e.g., Barzilai et al., 2015; Barzilai, Mor-Hagani, et al., 2020; Barzilai, Thomm, et al., 2020). Due to the situated nature of EC, further distinction between epistemic processes in different environments is needed with illustrative examples to elaborate on which epistemic processes are appropriate and when. Based on the categories in Table 3 (this chapter), relevant epistemic ideals and processes beyond examining explicit indicators of reliability are suggested to extend research.

Evaluation criteria can directly and indirectly examine the epistemic properties of knowledge claims, sources and arguments. Source evaluations can also occur within and across sources. Evaluation criteria are highly related, such that indicators of authority may also provide insight about accuracy or purpose. To deem a source trustworthy based on its authority,

individuals may examine the author's expertise and the publisher's reputation. An author's credentials, experience, or previous work may be used as an indicator of their authority. Similarly, a publisher's authority may be assessed based on indicators of its reputation such as its level of factual reporting, biases, or funding. Trustworthiness based on a source's accuracy may be determined by looking at the quality of the sources and evidence included or linked to by the author. Individuals may also attempt to corroborate the quality and accuracy of the evidence presented by looking at third-part endorsements or other sources on the topic. A source may be considered trustworthy based on its purpose after examining indicators of the author's or publisher's objectivity, such as assessing their relevant biases or motivations. Individuals may use a combination of these three categories or even combine these categories with content-based categories to determine a source's reliability. Whereas the three categories identified by Barzilai and Chinn (2018) clearly fall under EC, categories related to a source's currency and relevance may also influence epistemic processes during online source evaluations. For example, individuals may consider the freshness of data to determine the quality of evidence or the presence and quality of advertisements to determine a website's purpose. These cues provide important information about the context to which individuals are adapting their EC.

As Chinn and colleagues have recently argued (Barzilai & Chinn, 2018; Chinn & Sandoval, 2018), some epistemic processes are useful in multiple domains, tasks, or topics, and others are not. Whereas multiple documents researchers have contributed greatly to education researchers' understandings of EC during offline source evaluations, their methodologies do reflect key aspects of online evaluations, such as the option to look for external sources or deduct a website's purpose from its advertisements. These researchers have been criticized for asserting research implications for online environments (Wopereis & van Merriënboer, 2011). If the

present literature review were reduced to solely represent online environments, only 22 articles (of 73) would have been selected for detailed review, with two articles from EC researchers (Greene et al., 2014, 2018). There is a bidirectional relationship between understandings of EC and online source evaluation behaviours (Tsai, 2004). Further research is needed in authentic online environments to better understand EC and how criteria previously deemed non-epistemic and unsophisticated may support adaptive EC in an unfiltered environment.

General Conclusions

Source evaluations on the internet have become an increasingly important avenue of research due to the potential impact at the individual and societal level. While post-truth era dynamics are not new, the affordances of technology to access, create, and distribute knowledge is new. Everyone is responsible for navigating the sea of information created by these affordances. As such, strong source evaluation skills are necessary to swim in this non-linear environment. Researchers have documented individuals' varying cognitive processes during source evaluations. They have also successfully supported development of crucial evaluation skills; however, many questions remain about epistemic processes in online environments. Additional research is needed to address this gap in the literature.

References

Alexander, P. A. (2016). The arguments for and the reasoning about epistemic cognition. In J. A. Greene, W. A. Sandoval, & I. Bråten (Eds.), *Handbook of epistemic cognition* (pp. 112–122). Routledge.

*Alkali, Y. E., & Amichai-Hamburger, Y. (2004). Experiments in digital literacy. *CyberPsychology & Behavior*, *7*(4), 421–429.

*Argelagós, E., & Pifarré, M. (2012). Improving information problem solving skills in secondary education through embedded instruction. *Computers in Human Behavior*, *28*(2), 515–526. https://doi.org/10.1016/j.chb.2011.10.024

*Barnes, M. D., Penrod, C., Neiger, B. L., Merrill, R. M., Thackeray, R., Eggett, D. L., & Thomas, E. (2003). Measuring the relevance of evaluation criteria among health information seekers on the internet. *Journal of Health Psychology*, *8*(1), 71–82. https://doi.org/10.1177/1359105303008001436

Barzilai, S., & Chinn, C. A. (2018). On the goals of epistemic education: Promoting apt epistemic performance. *Journal of the Learning Sciences*, *27*(3), 353–389. https://doi.org/10.1080/10508406.2017.1392968

Barzilai, S., & Chinn, C. A. (2020). A review of educational responses to the "post-truth" condition: Four lenses on "post-truth" problems. *Educational Psychologist*, *55*(3), 107–119. https://doi.org/10.1080/00461520.2020.1786388

*Barzilai, S., & Eshet-Alkalai, Y. (2015). The role of epistemic perspectives in comprehension of multiple author viewpoints. *Learning and Instruction*, *36*, 86–103. https://doi.org/10.1016/j.learninstruc.2014.12.003

*Article included in detailed review

*Barzilai, S., & Ka'adan, I. (2017). Learning to integrate divergent information sources: The interplay of epistemic cognition and epistemic metacognition. *Metacognition and Learning*, *12*(2), 193–232.

*Barzilai, S., Mor-Hagani, S., Zohar, A. R., Shlomi-Elooz, T., & Ben-Yishai, R. (2020). Making sources visible: Promoting multiple document literacy with digital epistemic scaffolds. *Computers & Education*, *157*, 103980. https://doi.org/10.1016/j.compedu.2020.103980

*Barzilai, S., Thomm, E., & Shlomi-Elooz, T. (2020). Dealing with disagreement: The roles of topic familiarity and disagreement explanation in evaluation of conflicting expert claims and sources. *Learning and Instruction*, *69*, 101367. https://doi.org/10.1016/j.learninstruc.2020.101367

*Barzilai, S., Tzadok, E., & Eshet-Alkalai, Y. (2015). Sourcing while reading divergent expert accounts: Pathways from views of knowing to written argumentation. *Instructional Science*, *43*(6), 737–766.

Barzilai, S., & Weinstock, M. (2015). Measuring epistemic thinking within and across topics: A scenario-based approach. *Contemporary Educational Psychology*, *42*, 141–158. https://doi.org/10.1016/j.cedpsych.2015.06.006

*Barzilai, S., & Zohar, A. (2012). Epistemic thinking in action: Evaluating and integrating online sources. *Cognition and Instruction*, *30*(1), 39–85. https://doi.org/10.1080/07370008.2011.636495

Barzilai, S., & Zohar, A. (2014). Reconsidering personal epistemology as metacognition: A multifaceted approach to the analysis of epistemic thinking. *Educational Psychologist*, *49*(1), 13–35. https://doi.org/10.1080/00461520.2013.863265

Barzilai, S., & Zohar, A. (2016). Epistemic (meta) cognition: Ways of thinking about knowledge and knowing. In J. A. Greene, W. A. Sandoval, & I. Bråten (Eds.), *Handbook of epistemic cognition* (pp. 409–424). Routledge.

Barzilai, S., Zohar, A. R., & Mor-Hagani, S. (2018). Promoting integration of multiple texts: A review of instructional approaches and practices. *Educational Psychology Review*, 1–27.

*Bates, B. R., Romina, S., Ahmed, R., & Hopson, D. (2006). The effect of source credibility on consumers' perceptions of the quality of health information on the Internet. *Medical Informatics and the Internet in Medicine*, *31*(1), 45–52. https://doi.org/10.1080/14639230600552601

Baxter Magolda, M. B. (2004). Evolution of a constructivist conceptualization of epistemological reflection. *Educational Psychologist*, *39*(1), 31–42.

Belenky, M. F., Clinchy, B. M., Goldberger, N. R., Tarule, J. M., & others. (1986). *Women's ways of knowing: The development of self, voice, and mind* (Vol. 15). Basic books.

Bendixen, L. D. (2002). A process model of epistemic belief change. In B. K. Hofer & P. R. Pintrich (Eds.), *Personal epistemology: The psychology of beliefs about knowledge and knowing* (pp. 191–208). Lawrence Erlbaum Associates Publishers.

Bendixen, L. D., & Rule, D. C. (2004). An integrative approach to personal epistemology: A guiding model. *Educational Psychologist*, *39*(1), 69–80.

*Braasch, J. L. G., Bråten, I., Strømsø, H. I., Anmarkrud, Ø., & Ferguson, L. E. (2013). Promoting secondary school students' evaluation of source features of multiple documents. *Contemporary Educational Psychology*, *38*(3), 180–195. https://doi.org/10.1016/j.cedpsych.2013.03.003

Brante, E. W., & Strømsø, H. I. (2018). Sourcing in text comprehension: A review of interventions targeting sourcing skills. *Educational Psychology Review*, *30*(3), 773–799.

Bråten, I., Brante, E. W., & Strømsø, H. I. (2019). Teaching sourcing in upper secondary school: A comprehensive sourcing intervention with follow-up data. *Reading Research Quarterly*, *54*(4), 481–505.

Bråten, I., Ferguson, L. E., Anmarkrud, Ø., & Strømsø, H. I. (2013). Prediction of learning and comprehension when adolescents read multiple texts: The roles of word-level processing, strategic approach, and reading motivation. *Reading and Writing*, *26*(3), 321–348.

Bråten, I., Strømsø, H. I., & Andreassen, R. (2016). Sourcing in professional education: Do text factors make any difference? *Reading and Writing*, *29*(8), 1599–1628.

Bråten, I., Strømsø, H. I., & Britt, M. A. (2009). Trust matters: Examining the role of source evaluation in students' construction of meaning within and across multiple texts. *Reading Research Quarterly*, *44*(1), 6–28.

Bråten, I., Strømsø, H. I., & Samuelstuen, M. S. (2005). The relationship between internet-specific epistemological beliefs and learning within internet technologies. *Journal of Educational Computing Research*, *33*(2), 141–171. https://doi.org/10.2190/E763-X0LN-6NMF-CB86

Breakstone, J., McGrew, S., Smith, M., Ortega, T., & Wineburg, S. (2018). Why we need a new approach to teaching digital literacy. *Phi Delta Kappan*, *99*(6), 27–32.

*Britt, M. A., & Aglinskas, C. (2002). Improving students' ability to identify and use source information. *Cognition and Instruction*, *20*(4), 485–522.

Britt, M. A., Rouet, J.-F., & Braasch, J. L. (2013). Documents as entities. In A. Britt, S. Goldman, & J.-F. Rouet (Eds.), *Reading–from words to multiple texts* (pp. 160-179). Routledge.

Bromme, R., Kienhues, D., & Porsch, T. (2010). Who knows what and who can we believe? Epistemological beliefs are beliefs about knowledge (mostly) to be attained from others. In L. D. Bendixen & F. C. Feucht (Eds.), *Personal epistemology in the classroom: Theory, research, and implications for practice* (pp. 163-194). Cambridge University Press.

*Camiling, M. K. S. (2019). Ehealth literacy of high school students in the philippines. *IAFOR Journal of Education*, *7*(2), 69–87. https://doi.org/10.22492/ije.7.2.04

Cartiff, B. M., Duke, R. F., & Greene, J. A. (2020). The effect of epistemic cognition interventions on academic achievement: A meta-analysis. *Journal of Educational Psychology*. https://doi.org/10.1037/edu0000490

Chandler, M., Boyes, M., & Ball, L. (1990). Relativism and stations of epistemic doubt. *Journal of Experimental Child Psychology*, *50*(3), 370–395.

Chandler, M. J., Hallett, D., & Sokol, B. W. (2002). Competing claims about competing knowledge claims. In B. K. Hofer & P. R. Pintrich (Eds.), *Personal epistemology: The psychology of beliefs about knowledge and knowing* (pp. 145–168). Lawrence Erlbaum Associates Publishers.

*Cheng, Y.-H., & Tsai, C.-C. (2020). Online research behaviors of engineering graduate students in Taiwan. *Journal of Educational Technology & Society*, *20*(1), 169–179.

*Chiagouris, L., Long, M. M., & Plank, R. E. (2008). The consumption of online news: The relationship of attitudes toward the site and Credibility. *Journal of Internet Commerce, 7*(4), 528–549. https://doi.org/10.1080/15332860802507396

Chinn, C. A., Buckland, L. A., & Samarapungavan, A. (2011). Expanding the dimensions of epistemic cognition: Arguments from philosophy and psychology. *Educational Psychologist, 46*(3), 141–167. https://doi.org/10.1080/00461520.2011.587722

Chinn, C. A., & Rinehart, R. W. (2016). Epistemic cognition and philosophy: Developing a new framework for epistemic cognition. In J. A. Greene, W. A. Sandoval, & I. Bråten (Eds.), *Handbook of epistemic cognition* (pp. 460–478). Routledge.

Chinn, C. A., Rinehart, R. W., & Buckland, L. A. (2014). Epistemic cognition and evaluating information: Applying the AIR model of epistemic cognition. In D. N. Rapp & J. L.G. Braasch (Eds.), *Processing inaccurate information: Theoretical and applied perspectives from cognitive science and the educational sciences* (pp. 425–453). MIT Press.

Chinn, C. A., & Sandoval, W. A. (2018). Epistemic cognition and epistemic development. In F. Fischer, C. E. Hmelo-Silver, S. R. Goldman, & P. Resiman (Eds.), *International handbook of the learning sciences* (pp. 46–55). Routledge.

*Cho, B.-Y., Woodward, L., & Li, D. (2018). Epistemic processing when adolescents read online: A verbal protocol analysis of more and less successful online readers. *Reading Research Quarterly, 53*(2), 197–221.

*Clark, J. C., Rawson, C., Hobbs, L., Oughtred, C., Hayes, K., Kelly, L., & Higgins, J. (2017). A pedagogy for epistemic agency in the judgment of accuracy and reliability. *Qwerty-Open and Interdisciplinary Journal of Technology, Culture and Education, 11*(2), 27–47.

*Cugelman, B., Thelwall, M., & Dawes, P. (2009). The dimensions of web site credibility and their relation to active trust and behavioural impact. *Communications of the Association for Information Systems, 24*. https://doi.org/10.17705/1CAIS.02426

*Cunningham, A., & Johnson, F. (2016). Exploring trust in online health information: A study of user experiences of patients.co.uk. *Health Information & Libraries Journal, 33*(4), 323–328. https://doi.org/10.1111/hir.12163

*De La Paz, S., & Felton, M. K. (2010). Reading and writing from multiple source documents in history: Effects of strategy instruction with low to average high school writers. *Contemporary Educational Psychology*, *35*(3), 174–192. https://doi.org/10.1016/j.cedpsych.2010.03.001

DeBacker, T. K., Crowson, H. M., Beesley, A. D., Thoma, S. J., & Hestevold, N. L. (2008). The Challenge of measuring epistemic beliefs: An analysis of three self-report instruments. *The Journal of Experimental Education*, *76*(3), 281–312. https://doi.org/10.3200/JEXE.76.3.281-314

*Dochterman, M. A., & Stamp, G. H. (2010). The determination of Web credibility: A thematic analysis of Web user's judgments. *Qualitative Research Reports in Communication*, *11*(1), 37–43. https://doi.org/10.1080/17459430903514791

*Eastin, M. S., Yang, M.-S., & Nathanson, A. I. (2006). Children of the net: An empirical exploration into the evaluation of Internet content. *Journal of Broadcasting & Electronic Media, 50*(2), 211–230. https://doi.org/10.1207/s15506878jobem5002_3

Elby, A., & Hammer, D. (2010). Epistemological resources and framing: A cognitive framework for helping teachers interpret and respond to their students' epistemologies. In L. D. Bendixen & F. C. Feucht (Eds.), *Personal epistemology in the classroom: Theory,*

research, and implications for practice (Vol. 4, pp. 409–434). Cambridge University Press.

Eshet, Y. (2012). Thinking in the digital era: A revised model for digital literacy. *Issues in Informing Science and Information Technology*, *9*(2), 267–276.

Eshet-Alkalai, Y., & Chajut, E. (2009). Changes over time in digital literacy. *CyberPsychology & Behavior*, *12*(6), 713–715.

*Eysenbach, G. (2002). How do consumers search for and appraise health information on the world wide web? Qualitative study using focus groups, usability tests, and in-depth interviews. *BMJ*, *324*(7337), 573–577. https://doi.org/10.1136/bmj.324.7337.573

*Ferguson, L. E., Bråten, I., & Strømsø, H. I. (2012). Epistemic cognition when students read multiple documents containing conflicting scientific evidence: A think-aloud study. *Learning and Instruction*, *22*(2), 103–120.

*Flanagin, A. J., & Metzger, M. J. (2007). The role of site features, user attributes, and information verification behaviors on the perceived credibility of web-based information. *New Media & Society*, *9*(2), 319–342. https://doi.org/10.1177/1461444807075015

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American Psychologist*, *34*(10), 906.

Flavell, J. H., Miller, P. H., & Miller, S. A. (2002). *Cognitive development* (4th ed., Vol. 338). Prentice-Hall Englewood Cliffs, NJ.

*Gerjets, P., Kammerer, Y., & Werner, B. (2011). Measuring spontaneous and instructed evaluation processes during Web search: Integrating concurrent thinking-aloud protocols and eye-tracking data. *Learning and Instruction*, *21*(2), 220–231. https://doi.org/10.1016/j.learninstruc.2010.02.005

Greene, J. A. (2016). Interacting epistemic systems within and beyond the classroom. In J. A. Greene, W. A. Sandoval, & I. Bråten (Eds.), *Handbook of epistemic cognition* (pp. 265–278). Routledge.

Greene, J. A., Azevedo, R., & Torney-Purta, J. (2008). Modeling epistemic and ontological cognition: Philosophical perspectives and methodological directions. *Educational Psychologist, 43*(3), 142–160. https://doi.org/10.1080/00461520802178458

Greene, J. A., Cartiff, B. M., & Duke, R. F. (2018). A meta-analytic review of the relationship between epistemic cognition and academic achievement. *Journal of Educational Psychology*. https://doi.org/10.1037/edu0000263

*Greene, J. A., Copeland, D. Z., Deekens, V. M., & Seung, B. Y. (2018). Beyond knowledge: Examining digital literacy's role in the acquisition of understanding in science. *Computers & Education*, *117*, 141–159.

Greene, J. A., Sandoval, W. A., & Bråten, I. (2016). *Handbook of epistemic cognition*. Routledge.

*Greene, J. A., Yu, S. B., & Copeland, D. Z. (2014). Measuring critical components of digital literacy and their relationships with learning. *Computers & Education*, *76*, 55–69. https://doi.org/10.1016/j.compedu.2014.03.008

Greeno, J. G. (2015). Commentary: Some prospects for connecting concepts and methods of individual cognition and of situativity. *Educational Psychologist, 50*(3), 248–251. https://doi.org/10.1080/00461520.2015.1077708

*Greer, J. D. (2003). Evaluating the credibility of online information: A test of source and advertising influence. *Mass Communication and Society*, *6*(1), 11–28. https://doi.org/10.1207/S15327825MCS0601_3

Gui, M., & Argentin, G. (2011). Digital skills of internet natives: Different forms of digital literacy in a random sample of northern Italian high school students. *New Media & Society*, *13*(6), 963–980.

*Halverson, K. L., Siegel, M. A., & Freyermuth, S. K. (2010). Non-science majors' critical evaluation of websites in a biotechnology course. *Journal of Science Education and Technology*, *19*(6), 612–620. https://doi.org/10.1007/s10956-010-9227-6

*Hämäläinen, E. K., Kiili, C., Marttunen, M., Räikkönen, E., González-Ibáñez, R., & Leppänen, P. H. T. (2020). Promoting sixth graders' credibility evaluation of Web pages: An intervention study. *Computers in Human Behavior*, *110*, 106372. https://doi.org/10.1016/j.chb.2020.106372

Hammer, D., & Elby, A. (2002). On the form of a personal epistemology. In B. K. Hofer and P. R. Pintrich (Eds.), *Personal epistemology: The psychology of beliefs about knowledge and knowing* (pp. 169–190). Lawrence Erlbaum Associates Publishers.

*Harris, P. R., Sillence, E., & Briggs, P. (2009). The effect of credibility-related design cues on responses to a Web-based message about the breast cancer risks from alcohol: Randomized controlled trial. *Journal of Medical Internet Research*, *11*(3), e37. https://doi.org/10.2196/jmir.1097

Hirsh, S. G. (1999). Children's relevance criteria and information seeking on electronic resources. *Journal of the American Society for Information Science*, *50*(14), 1265–1283.

Hofer, B. K. (2000). Dimensionality and disciplinary differences in personal epistemology. *Contemporary Educational Psychology*, *25*(4), 378–405.

Hofer, B. K. (2001). Personal epistemology research: Implications for learning and teaching. *Educational Psychology Review*, *13*(4), 353–383.

Hofer, B. K. (2004). Epistemological understanding as a metacognitive process: Thinking aloud during online searching. *Educational Psychologist, 39*(1), 43–55.

Hofer, B. K. (2016). Epistemic cognition as a psychological construct: Advancements and challenges. In J. A. Greene, W. A. Sandoval, & I. Bråten (Eds.), *Handbook of epistemic cognition* (pp. 31–50). Routledge.

Hofer, B. K., & Pintrich, P. R. (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research*, *67*(1), 88–140.

Hootsuite & We Are Social (2020). *Digital 2020*. Hootsuite. https://hootsuite.com/resources/digital-2020

*Iding, M., & Klemm, E. B. (2005). Pre-service teachers critically evaluate scientific information on the World Wide Web. *Computers in the Schools*, *22*(1–2), 7–18. https://doi.org/10.1300/J025v22n01_02

*Iordanou, K., Muis, K. R., & Kendeou, P. (2018). Epistemic perspective and online epistemic processing of evidence: Developmental and domain Differences. *The Journal of Experimental Education*, 1–21.

*Jung, E. H., Walsh-Childers, K., & Kim, H.-S. (2016). Factors influencing the perceived credibility of diet-nutrition information web sites. *Computers in Human Behavior*, *58*, 37–47. https://doi.org/10.1016/j.chb.2015.11.044

*Jung, W. S., Chung, M.-Y., & Rhee, E. S. (2018). The Effects of Attractiveness and Source Expertise on Online Health Sites. *Health Communication*, *33*(8), 962–971. https://doi.org/10.1080/10410236.2017.1323364

Kąkol, M., & Nielek, R. (2015). What affects web credibility perception? An analysis of textual justifications. *Computer Science, 16,* 295-310.

*Kąkol, M., Nielek, R., & Wierzbicki, A. (2017). Understanding and predicting Web content credibility using the Content Credibility Corpus. *Information Processing & Management*, *53*(5), 1043–1061. https://doi.org/10.1016/j.ipm.2017.04.003

*Kalichman, S. C., Cherry, C., Cain, D., Weinhardt, L. S., Benotsch, E., Pope, H., & Kalichman, M. (2006). Health information on the Internet and people living with HIV/AIDS: Information evaluation and coping styles. *Health Psychology*, *25*(2), 205–210. https://doi.org/10.1037/0278-6133.25.2.205

Kammerer, Y., Amann, D. G., & Gerjets, P. (2015). When adults without university education search the Internet for health information: The roles of Internet-specific epistemic beliefs and a source evaluation intervention. *Computers in Human Behavior*, *48*, 297–309. https://doi.org/10.1016/j.chb.2015.01.045

Kammerer, Y., Bråten, I., Gerjets, P., & Strømsø, H. I. (2013). The role of Internet-specific epistemic beliefs in laypersons' source evaluations and decisions during Web search on a medical issue. *Computers in Human Behavior*, *29*(3), 1193–1203. https://doi.org/10.1016/j.chb.2012.10.012

*Kammerer, Y., Kalbfell, E., & Gerjets, P. (2016). Is this information source commercially biased? How contradictions between web pages stimulate the consideration of source information. *Discourse Processes*, *53*(5–6), 430–456. https://doi.org/10.1080/0163853X.2016.1169968

Kienhues, D., Bromme, R., & Stahl, E. (2008). Changing epistemological beliefs: The unexpected impact of a short-term intervention. *British Journal of Educational Psychology, 78*(4), 545–565.

Kienhues, D., Ferguson, L., & Stahl, E. (2016). Diverging information and epistemic change. In J. A. Greene, W. A. Sandoval, & I. Bråten (Eds.), *Handbook of epistemic cognition* (pp. 318–330). Routledge.

*Kiili, C., Laurinen, L., & Marttunen, M. (2008). Students evaluating Internet sources: From versatile evaluators to uncritical readers. *Journal of Educational Computing Research, 39*(1), 75–95. https://doi.org/10.2190/EC.39.1.e

*Kim, M. (2015). Partisans and controversial news online: Comparing perceptions of bias and credibility in news content from blogs and mainstream media. *Mass Communication and Society, 18*(1), 17–36. https://doi.org/10.1080/15205436.2013.877486

King, P. M., & Kitchener, K. S. (1994). *Developing reflective judgment: Understanding and promoting intellectual growth and critical thinking in adolescents and adults.* Jossey-Bass.

*Kiousis, S., & Dimitrova, D. V. (2006). Differential impact of Web site content: Exploring the influence of source (public relations versus news), modality, and participation on college students' perceptions. *Public Relations Review, 32*(2), 177–179. https://doi.org/10.1016/j.pubrev.2006.02.013

Kitchener, R. F. (2011). Personal epistemology and philosophical epistemology: The view of a philosopher. In J. Elen, E. Stahl, R. Bromme, & G. Clarebout (Eds.), *Links between beliefs and cognitive flexibility* (pp. 79–103). Springer.

Kitchner, K. S. (1983). Cognition, metacognition, and epistemic cognition. *Human Development*, *26*(4), 222–232.

Knight, S., Rienties, B., Littleton, K., Mitsui, M., Tempelaar, D., & Shah, C. (2017). The relationship of (perceived) epistemic cognition to interaction with resources on the internet. *Computers in Human Behavior*, *73*, 507–518. https://doi.org/10.1016/j.chb.2017.04.014

Kuhn, D. (1991). *The skills of argument*. Cambridge University Press.

Kuhn, D. (1999). A developmental model of critical thinking. *Educational Researcher*, *28*(2), 16–46.

Kuhn, D., Cheney, R., & Weinstock, M. (2000). The development of epistemological understanding. *Cognitive Development*, *15*, 309–328. https://doi.org/10.1016/S0885-2014(00)00030-7

Kuhn, D., & Weinstock, M. (2002). What is epistemological thinking and why does it matter? In B. K. Hofer & P. R. Pintrich (Eds.), *Personal epistemology: The psychology of beliefs about knowledge and knowing* (pp. 121–144). Lawrence Erlbaum Associates Publishers.

Leu, D. J., Kinzer, C. K., Coiro, J., Castek, J., & Henry, L. A. (2017). New literacies: A dual-level theory of the changing nature of literacy, instruction, and assessment. *Journal of Education*, *197*(2), 1–18.

*Liao, Q. V., & Fu, W.-T. (2014). Age differences in credibility judgments of online health information. *ACM Transactions on Computer-Human Interaction*, *21*(1), 1–23. https://doi.org/10.1145/2534410

*Lombardi, D., Seyranian, V., & Sinatra, G. M. (2014). Source effects and plausibility judgments when reading About climate change. *Discourse Processes*, *51*(1–2), 75–92. https://doi.org/10.1080/0163853X.2013.855049

Louca, L., Elby, A., Hammer, D., & Kagey, T. (2004). Epistemological resources: Applying a new epistemological framework to science instruction. *Educational Psychologist, 39*(1), 57–68. https://doi.org/10.1207/s15326985ep3901_6

Lucassen, T., & Schraagen, J. M. (2011). Factual accuracy and trust in information: The role of expertise. *Journal of the American Society for Information Science and Technology*, *62*(7), 1232–1242.

*Lucassen, T., & Schraagen, J. M. (2013). The influence of source cues and topic familiarity on credibility evaluation. *Computers in Human Behavior*, *29*(4), 1387–1392. https://doi.org/10.1016/j.chb.2013.01.036

*Madden, A. D., Ford, N., Gorrell, G., Eaglestone, B., & Holdridge, P. (2012). Metacognition and web credibility. *The Electronic Library, 30*(5), 671–689. https://doi.org/10.1108/02640471211275710

Maier, J., & Richter, T. (2013). How nonexperts understand conflicting information on social science issues. *Journal of Media Psychology, 25*(1), 14-26. https://doi.org/10.1027/1864-1105/a000078

Maier, J., & Richter, T. (2016). Effects of text-belief consistency and reading task on the strategic validation of multiple texts. *European Journal of Psychology of Education*, *31*(4), 479–497.

*Mason, L., Ariasi, N., & Boldrin, A. (2011). Epistemic beliefs in action: Spontaneous reflections about knowledge and knowing during online information searching and their

influence on learning. *Learning and Instruction*, *21*(1), 137–151. https://doi.org/10.1016/j.learninstruc.2010.01.001

*Mason, L., Boldrin, A., & Ariasi, N. (2010). Searching the Web to learn about a controversial topic: Are students epistemically active? *Instructional Science*, *38*(6), 607–633.

Mason, L., Junyent, A. A., & Tornatora, M. C. (2014). Epistemic evaluation and comprehension of web-source information on controversial science-related topics: Effects of a short-term instructional intervention. *Computers & Education*, *76*, 143–157. https://doi.org/10.1016/j.compedu.2014.03.016

Mason, L., Scrimin, S., Tornatora, M. C., Suitner, C., & Moè, A. (2018). Internet source evaluation: The role of implicit associations and psychophysiological self-regulation. *Computers & Education*, *119*, 59–75. https://doi.org/10.1016/j.compedu.2017.12.009

McGrew, S. (2020). Learning to evaluate: An intervention in civic online reasoning. *Computers & Education*, *145*, 103711. https://doi.org/10.1016/j.compedu.2019.103711

*McGrew, S., Smith, M., Breakstone, J., Ortega, T., & Wineburg, S. (2019). Improving university students' web savvy: An intervention study. *British Journal of Educational Psychology*, bjep.12279. https://doi.org/10.1111/bjep.12279

Menchen-Trevino, E., & Hargittai, E. (2011). Young adults' credibility assessment of Wikipedia. *Information, Communication & Society*, *14*(1), 24–51.

Meppelink, C. S., Smit, E. G., Fransen, M. L., & Diviani, N. (2019). "I was right about vaccination": Confirmation bias and health literacy in online health information seeking. *Journal of Health Communication*, *24*(2), 129–140. https://doi.org/10.1080/10810730.2019.1583701

Merk, S., Rosman, T., Muis, K. R., Kelava, A., & Bohl, T. (2018). Topic specific epistemic beliefs: Extending the Theory of Integrated Domains in Personal Epistemology. *Learning and Instruction*, *56*, 84–97.

Merk, S., Rosman, T., Rueß, J., Syring, M., & Schneider, J. (2017). Pre-service teachers' perceived value of general pedagogical knowledge for practice: Relations with epistemic beliefs and source beliefs. *Plos One*, *12*(9), e0184971.

Merk, S., Schneider, J., Bohl, T., Kelava, A., & Syring, M. (2017). Teacher students' epistemic beliefs about general pedagogical knowledge: Topic-, source-and context specificity. *Journal for Educational Research Online/Journal Für Bildungsforschung Online*, *9*(1), 169–189.

*Meyer, H. K., Marchionni, D., & Thorson, E. (2010). The ournalist behind the news: Credibility of straight, collaborative, opinionated, and blogged "news." *American Behavioral Scientist, 54*(2), 100–119. https://doi.org/10.1177/0002764210376313

Muis, K. R. (2007). The role of epistemic beliefs in self-regulated learning. *Educational Psychologist, 42*(3), 173–190. https://doi.org/10.1080/00461520701416306

Muis, K. R., Bendixen, L. D., & Haerle, F. C. (2006). Domain-generality and domain-specificity in personal epistemology research: Philosophical and empirical reflections in the development of a theoretical framework. *Educational Psychology Review*, *18*(1), 3–54.

Muis, K. R., Chevrier, M., & Singh, C. (2018). The role of epistemic emotions in personal epistemology and self-regulated learning. *Educational Psychologist*, 1–20.

Muis, K. R., & Duffy, M. C. (2013). Epistemic climate and epistemic change: Instruction designed to change students' beliefs and learning strategies and improve achievement. *Journal of Educational Psychology*, *105*(1), 213.

Muis, K. R., Etoubashi, N., & Denton, C. A. (2020). The catcher in the lie: The role of emotions and epistemic judgments in changing students' misconceptions and attitudes in a post-truth era. *Contemporary Educational Psychology, 62*, 101898. https://doi.org/10.1016/j.cedpsych.2020.101898

Muis, K. R., Trevors, G., & Chevrier, M. (2016). Epistemic climate for epistemic change. In J. A. Greene, W. A. Sandoval, & I. Bråten (Eds.), *Handbook of epistemic cognition* (pp. 331–359). Routledge.

Muis, K. R., & Singh, C. (2018). The three faces of epistemic thinking in self-regulated learning. In D. H. Schunk & J. A. Greene (Eds.), *Handbook of self-regulated of learning and performance* (pp. 434–456). Routledge/Taylor & Francis Group.

Murphy, P., Alexander, P., & Muis, K. R. (2012). Knowledge and knowing: The journey from philosophy and psychology to human learning. *Educational Psychology Handbook, 1*, 189–226.

Murphy, P. K., Alexander, P. A., Greene, J. A., & Edwards, M. N. (2007). Epistemological threads in the fabric of conceptual change research. In S. Vosniadou, A. Baltas, & X. Vamvakoussi (Eds.), *Advances in learning and instruction series. Reframing the conceptual change approach in learning and instruction* (p. 105–122). Elsevier Science.

*Nokes, J. D., Dole, J. A., & Hacker, D. J. (2007). Teaching high school students to use heuristics while reading historical texts. *Journal of Educational Psychology*, *99*(3), 492–504. https://doi.org/10.1037/0022-0663.99.3.492

Perfetti, C. A., Rouet, J.-F., & Britt, M. A. (1999). Toward a theory of documents representation. In H. van Oostendorp & S. R. Goldman (Eds.), *The construction of mental representations during reading* (p. 99–122). Lawrence Erlbaum Associates Publishers

Perry, W. G. (1970). *Forms of intellectual and ethical development in the college years: A scheme*. Holt, Rinehart and Winston.

*Pretorius, L. (2018). Experiential and self-discovery learning in digital literacy: Developing the discernment to evaluate source reliability. *College & Undergraduate Libraries*, *25*(4), 388–405. https://doi.org/10.1080/10691316.2018.1530626

*Rains, S. A., & Karmikel, C. D. (2009). Health information-seeking and perceptions of website credibility: Examining Web-use orientation, message characteristics, and structural features of websites. *Computers in Human Behavior*, *25*(2), 544–553. https://doi.org/10.1016/j.chb.2008.11.005

Ricco, R., Schuyten Pierce, S., & Medinilla, C. (2010). Epistemic beliefs and achievement motivation in early adolescence. *The Journal of Early Adolescence*, *30*(2), 305–340.

Richter, T., & Maier, J. (2017). Comprehension of multiple documents with conflicting information: A two-step model of validation. *Educational Psychologist*, *52*(3), 148–166. https://doi.org/10.1080/00461520.2017.1322968

Richter, T., & Schmid, S. (2010). Epistemological beliefs and epistemic strategies in self-regulated learning. *Metacognition and Learning*, *5*(1), 47–65.

*Robertson-Lang, L., Major, S., & Hemming, H. (2011). An exploration of search patterns and credibility Issues among older adults seeking online health information. *Canadian Journal on Aging / La Revue Canadienne Du Vieillissement*, *30*(4), 631–645. https://doi.org/10.1017/S071498081100050X

Rouet, J.-F., & Britt, M. A. (2011). Relevance processes in multiple document comprehension. In M. T. McCrudden, J. P. Magliano, & G. Schraw (Eds.), *Text relevance and learning from text* (p. 19–52). IAP Information Age Publishing.

Rouet, J.-F., Britt, M. A., & Durik, A. M. (2017). RESOLV: Readers' representation of reading contexts and tasks. *Educational Psychologist, 52*(3), 200–215. https://doi.org/10.1080/00461520.2017.1329015

Royce, J. (1978). Three ways of knowing and the scientific world view. *Methodology and Science, 11*, 146–164.

Royce, J. R. (1959). The search for meaning. *American Scientist, 47*(4), 515–535.

Rule, D. C., & Bendixen, L. D. (2010). The integrative model of personal epistemology development: Theoretical underpinnings and implications for education. In L. Bendixen & F. C. Feucht (Eds.), *Personal epistemology in the classroom: Theory, research, and implications for practice* (pp.94-123). Cambridge University Press.

Ryu, S., & Sandoval, W. A. (2012). Improvements to elementary children's epistemic understanding from sustained argumentation. *Science Education, 96*(3), 488–526.

Sabourin, J. L., Shores, L. R., Mott, B. W., & Lester, J. C. (2013). Understanding and predicting student self-regulated learning strategies in game-based learning environments. *International Journal of Artificial Intelligence in Education*, *23*(1–4), 94–114.

Salmerón, L., Gil, L., & Bråten, I. (2018). Effects of reading real versus print-out versions of multiple documents on students' sourcing and integrated understanding. *Contemporary Educational Psychology, 52*, 25–35. https://doi.org/10.1016/j.cedpsych.2017.12.002

Sandoval, W. A. (2005). Understanding students' practical epistemologies and their influence on learning through inquiry. *Science Education*, *89*(4), 634–656.

Sandoval, W. A. (2009). In defense of clarity in the study of personal epistemology. *The Journal of the Learning Sciences*, *18*(1), 150–161. https://doi.org/10.1080/10508400802581700

Sandoval, W. A. (2012). Situating epistemological development. *International Society of the Learning Sciences (ISLS)*, 347 – 354.

Sandoval, W. A. (2014). Science education's need for a theory of epistemological development. *Science Education, 98*(3), 383–387.

Sandoval, W. A. (2017). Situating practices of epistemic cognition. In T. G. Amin & O. Levrini (Eds.), *Converging Perspectives on Conceptual Change* (1st ed., pp. 253–260). Routledge. https://doi.org/10.4324/9781315467139-30

Sandoval, W. A., & Çam, A. (2011). Elementary children's judgments of the epistemic status of sources of justification. *Science Education, 95*(3), 383–408.

Sandoval, W.A., Greene, J. A., & Bråten, I. (2016). Understanding and promoting thinking about knowledge: Origins, issues, and future directions of research on epistemic cognition. *Review of Research in Education*, *40*(1), 457–496. https://doi.org/10.3102/0091732X16669319

Sandoval, W. A., & Redman, E. H. (2015). The contextual nature of scientists' views of theories, experimentation, and their coordination. *Science & Education, 24*(9–10), 1079–1102.

Schacter, J., Chung, G. K., & Dorr, A. (1998). Children's Internet searching on complex problems: Performance and process analyses. *Journal of the American Society for Information Science, 49*(9), 840–849.

Schommer, M. (1990). Effects of beliefs about the nature of knowledge on comprehension. *Journal of Educational Psychology*, *82*(3), 498.

Schommer-Aikins, M. (2004). Explaining the epistemological belief system: Introducing the embedded systemic model and coordinated research approach. *Educational Psychologist, 39*(1), 19–29.

Sinatra, G. M. (2016). Thoughts on knowledge about thinking about knowledge. In J. A. Greene, W. A. Sandoval, & I. Bråten (Eds.), *Handbook of epistemic cognition* (pp. 479–491). Routledge.

Sosa, E. (2015). *Judgment and agency*. Oxford University Press, USA.

*Stadtler, M., Scharrer, L., Macedo-Rouet, M., Rouet, J.-F., & Bromme, R. (2016). Improving vocational students' consideration of source information when deciding about science controversies. *Reading and Writing*, *29*(4), 705–729. https://doi.org/10.1007/s11145-016-9623-2

Stanford History Education Group. (2020). *COR*. Civic Online Reasoning. https://cor.stanford.edu

*Starling, M. S., Cheshire, C., Deardorff, J., & Nuru-Jeter, A. (2018). Late adolescent user experiences with online sexual health resources: A qualitative study. *American Journal of Sexuality Education*, *13*(4), 487–503. https://doi.org/10.1080/15546128.2018.1505572

Strømsø, H. I., & Bråten, I. (2010). The role of personal epistemology in the self-regulation of internet-based learning. *Metacognition and Learning*, *5*(1), 91–111. https://doi.org/10.1007/s11409-009-9043-7

*Subramaniam, M., Taylor, N. G., St. Jean, B., Follman, R., Kodama, C., & Casciotti, D. (2015). As simple as that?: Tween credibility assessment in a complex online world. *Journal of Documentation*, *71*(3), 550–571. https://doi.org/10.1108/JD-03-2014-0049

*Swasy, A., Tandoc, E., Bhandari, M., & Davis, R. (2015). Traditional reporting more credible than citizen news. *Newspaper Research Journal*, *36*(2), 225–236. https://doi.org/10.1177/0739532915587298

*Tandoc, E. C., & Thomas, R. J. (2017). Readers value objectivity over transparency. *Newspaper Research Journal*, *38*(1), 32–45. https://doi.org/10.1177/0739532917698446

*Taylor, W. D., Johnson, G., Ault, M. K., Griffith, J. A., Rozzell, B., Connelly, S., Jensen, M. L., Dunbar, N. E., & Ness, A. M. (2015). Ideological group persuasion: A within-person study of how violence, interactivity, and credibility features influence online persuasion. *Computers in Human Behavior*, *51*, 448–460. https://doi.org/10.1016/j.chb.2015.04.071

The News Literacy Project. (2020). *Checkology*. Checkology. https://get.checkology.org

*Thon, F. M., & Jucks, R. (2017). Believing in expertise: How authors' credentials and language use influence the credibility of online health information. *Health Communication*, *32*(7), 828–836. https://doi.org/10.1080/10410236.2016.1172296

*Treise, D., Walsh-Childers, K., Weigold, M. F., & Friedman, M. (2003). Cultivating the science Internet audience: Impact of brand and domain on source credibility for science information. *Science Communication*, *24*(3), 309–332. https://doi.org/10.1177/1075547002250298

Trevors, G. J., Muis, K. R., Pekrun, R., Sinatra, G. M., & Muijselaar, M. M. (2017). Exploring the relations between epistemic beliefs, emotions, and learning from texts. *Contemporary Educational Psychology*, *48*, 116–132. https://doi.org/10.1016/j.cedpsych.2016.10.001

Tsai, C.-C. (2004). Beyond cognitive and metacognitive tools: The use of the Internet as an 'epistemological' tool for instruction. *British Journal of Educational Technology*, *35*(5), 525–536.

*Ulyshen, T. Z., Koehler, M. J., & Gao, F. (2015). Understanding the connection between epistemic beliefs and Internet searching. *Journal of Educational Computing Research*, *53*(3), 345–383.

Van Strien, J. L., Brand-Gruwel, S., & Boshuizen, H. P. (2014). Dealing with conflicting information from multiple nonlinear texts: Effects of prior attitudes. *Computers in Human Behavior*, *32*, 101–111.

*Walraven, A., Brand-Gruwel, S., & Boshuizen, H. P. A. (2013). Fostering students' evaluation behaviour while searching the internet. *Instructional Science*, *41*(1), 125–146. https://doi.org/10.1007/s11251-012-9221-x

*Wiblom, J. (2019). Developing students' critical reasoning about online health information: A capabilities approach. *Res Sci Educ*, 24.

*Wiley, J., Goldman, S. R., Graesser, A. C., Sanchez, C. A., Ash, I. K., & Hemmerich, J. A. (2009). Source evaluation, comprehension, and learning in Internet science inquiry tasks. *American Educational Research Journal*, *46*(4), 1060–1106. https://doi.org/10.3102/0002831209333183

Wineburg, S. S. (1991). On the reading of historical texts: Notes on the breach between school and academy. *American Educational Research Journal*, *28*(3), 495–519.

Wopereis, I. G., & van Merriënboer, J. J. (2011). Evaluating text-based information on the World Wide Web. *Learning and Instruction*, *21*(2), 232–237.

Bridging Text

Chapter 2 presented theoretical and empirical explorations of personal epistemology to better understand the relationship between EC and source evaluations. The purpose was to review conceptualizations of EC (Part 1) and examine how researchers have investigated source evaluations in distinct learning contexts (Part 2). Based on this literature review, a situated view of' EC is adopted to account for researchers' diverse findings about source evaluations. This review also identified that online source evaluations have been underexplored by education researchers. To extend research efforts, Chapter 2 identified authentic online environments as a promising avenue to better understand individuals' situated EC during source evaluations.

The following chapter presents the first two phases of a multiphase mixed methods research program aimed at developing an online source evaluation training. Adopting Barzilai and Chinn's (2018) Apt-AIR framework, in phase 1, college students were interviewed about their epistemic ideals and reliable epistemic processes to infer their epistemic metacognitive knowledge about source evaluations. Four college students also completed a source evaluation task where they compared the reliability of two online sources: a news article and an original research article. In phase 2, a larger group of college students' epistemic ideals were captured in a source evaluation task where they compared the reliability of two authentic news articles. These studies contribute to better understanding epistemic ideals and reliable processes and to developing a source evaluation training to improve online source evaluations.

Chapter 3

Manuscript 1

En-garde: Source evaluations in the digital age

Denton, C. A., Muis, K. R., Armstrong, S., & Dubé, A. (in preparation). En-garde: Source evaluations in the digital age. *Contemporary Educational Psychology*.

Abstract

Students have difficulty assessing the quality of information. They often rely on content-focused criteria to make reliability assessments and, as a result, may accept inaccurate information. Despite the impact of poor source evaluation skills, educational researchers have not widely examined source evaluation behaviours in authentic environments or tasks. Students' epistemic cognition, or their thinking about the epistemic properties of specific knowledge claims and sources, is one promising avenue to better understand their source evaluation behaviours. This manuscript reports two studies that explored students' epistemic thinking. In Study 1, college students (n = 12) reported their reliability criteria in focus group interviews. Four of these participants (n = 4) also examines the reliability of an online news article. Grounded theory was used to infer students' epistemic ideals and describe their reliable epistemic processes. In Study 2, students (n = 43) rank-ordered two news articles and justified how they assigned each article its' rank in a written response. Most students were able to accurately rank-order the articles using relevant epistemic processes. Cluster analysis was used to characterize the evaluation criteria used. Surprisingly, more participants that justified their decisions using relevance criteria accurately rank-ordered the articles. The role of direct and indirect indicators of reliability are discussed through the lens of the Apt-AIR framework of epistemic thinking.

Keywords: epistemic cognition, source evaluation, internet

Access to the internet has changed the way students interact with the world around them. With more opportunities than ever to access, create, and share content, internet users can be seekers and sources of information. The pervasiveness of these roles has reinvigorated educational efforts to specifically develop students' ability to discern whether information is reliable. Researchers have documented students' difficulty gauging the reliability of sources (Braasch et al., 2013; Halverson et al., 2010; Mason et al., 2011). Specifically, researchers have identified the tendency to rely on content-focused features, such as comprehensibility (Machackova & Smahel, 2018; Subramaniam et al., 2015), and surface-level epistemic features, such as publisher (Bråten et al., 2009), to assess reliability. One way to foster adaptive source evaluations on the internet is by improving epistemic cognition (Greene & Yu, 2016; Sinatra & Chinn, 2012). Epistemic cognition refers to thinking about the acquisition, justification and use of knowledge (Hofer, 2016).

Students that do not engage in this specific form of thinking are more susceptible to accept and disseminate false information, which can have local and societal impacts (Chinn & Barzilai, 2018). For example, believing that the installation of 5G towers lead to the pandemic may influence the daily safety measures observed as well as voting decisions. Undoubtedly, the potential impact of improving students' epistemic cognition has stimulated research to better understand the nature of their thinking on the internet (e.g., Cho et al., 2018; Greene et al., 2014, 2018) and to boost this crucial 21st century skill (e.g., Barzilai et al., 2020; Mason et al., 2014; Wiley et al., 2009). Yet, to explore this digital skill, education researchers have primarily conducted studies offline (e.g., Mason et al., 2018) or used curated materials that may not reflect authentic information found on the internet (e.g., E. H. Jung et al., 2016; Thon & Jucks, 2017). Given the situated nature of students' epistemic cognition (Sandoval, 2017), the implications of

such studies may not apply to source evaluations on the internet. To address this gap in the literature, we investigated college students' epistemic thinking about source evaluations using the Apt-AIR framework.

Literature Review

Critically assessing the quality of online information requires engaging in a variety of cognitive and metacognitive processes. These processes can consist of epistemic thinking or cognitive and metacognitive thinking about the epistemic properties of specific information, knowledge claims, and sources (Barzilai & Zohar, 2014). For example, a student may start by examining surface-level features, such as the content's publication date (cognition), and then assess the alignment of the information with their task (metacognition). Next, they may examine the author's expertise (epistemic cognition) and monitor the results of their evaluation to move forward accordingly (epistemic metacognition). According to Barzilai and Zohar (2014), a student's epistemic thinking processes interact such that their epistemic ideals influence the reliable epistemic processes they engage in. In their Apt-AIR framework, Barzilai and Chinn (2018) elaborated on the cognitive and metacognitive aspects of students' epistemic aims, ideals, and reliable processes during source evaluations.

Situated Epistemic Thinking

Educational theorists' have pushed for a situated view of personal epistemology to account for researchers' context-dependent findings (Sandoval, 2014, 2017; Hammer & Elby, 2002; Elby & Hammer, 2010). Barzilai and Chinn (2018) addressed this call in their Apt-AIR framework, in which they integrated their previous theories: the AIR model (Chinn et al., 2011, 2014; Chinn & Rinehart, 2016) and the Multifaceted Framework of epistemic thinking (Barzilai & Zohar, 2014, 2016). The Apt-AIR framework acknowledged that some epistemic thinking

may apply to multiple domains, whereas other epistemic thinking remains domain-specific. Chinn and Sandoval (2018) refined this position, explaining that students' epistemic processes may appear similar between domains or context, but the details of the processes differ substantially. For example, students engage in source evaluations in science and history contexts. However, they can engage in different reliable epistemic processes to evaluate trustworthiness. Given the variety of situations students encounter information, they are required to competently and adaptively apply appropriate epistemic aims, ideals, and processes to obtain epistemic achievements (Barzilai & Chinn, 2018). Students' apt use of epistemic processes support their ability to accurately evaluate and create information. To illuminate the intersections of students' cognitive and metacognitive processes presented in the Apt-AIR framework, we describe Barzilai's and Chinn's previous theoretical work separately.

In the AIR model, Chinn and colleagues (2014, 2016) described the cognitive processes that surround achievement of an epistemic aim. Their model included epistemic aims, ideals, and reliable epistemic processes. Epistemic aims refer to the objectives and importance a student sets for their cognition or action (e.g., knowledge, Chinn et al., 2014), and their aims can influence how they process information (Greene et al., 2014, 2018). For example, a student's epistemic aim may be to determine whether they can use information from an unfamiliar health website to make an informed health decision. Epistemic ideals refer to the criteria or standards students use to examine whether their epistemic aims have been met (e.g., adequacy of evidence, Chinn et al., 2014). Chinn and colleagues (2014) explained that a students' epistemic ideals are the criteria that they use to justify their acceptance or rejection of an epistemic product (e.g., claim or entire webpage, Barzilai & Chinn, 2018). To assess information quality, a student may enact reliable epistemic processes, such as consistency checking or integrating multiple sources, to achieve

their aims or produce epistemic products (Barzilai & Zohar, 2014; Richter & Schmid, 2010b). Whereas Chinn and colleagues (2014, 2016) focused on epistemic achievements, Barzilai and Zohar (2014, 2016) emphasized the antecedents of successful achievements.

Barzilai and Zohar's framework (2014, 2016) contributed cognitive and metacognitive aspects of epistemic thinking to the Apt-AIR model. Their framework described cognitive epistemic strategies and processes that can be used to scrutinize specific knowledge claims and sources. Following Flavell and colleagues (1979, Flavell et al., 2002), Barzilai and Zohar (2014, 2016) also delineated three aspects of epistemic metacognition: epistemic metacognitive skills, epistemic metacognitive knowledge, and epistemic metacognitive experiences. Epistemic metacognitive skills refer to a student's planning, monitoring and evaluating of the epistemic strategies and processes they engage in (Barzilai & Chinn, 2018; Barzilai & Zohar, 2014, 2016). For example, Cho and colleagues (2018) found that students employed planning and monitoring to integrate multiple perspectives and examine the accuracy of knowledge claims and sources to establish reliability. Epistemic metacognitive knowledge refers to a student's metacognitive knowledge about the nature of knowledge and knowing (Barzilai & Chinn, 2018; Barzilai & Zohar, 2014, 2016). During a source evaluation, a student's metacognitive knowledge that online information is created for a variety of purposes may stimulate their evaluation of an author's resulting biases. Their epistemic beliefs about knowledge in general may influence the types of processes they engage in as well as their epistemic metacognitive experiences. Epistemic metacognitive experiences refer to a student's emotions that are evoked as they build knowledge (Barzilai & Chinn, 2018; Barzilai & Zohar, 2014, 2016). For example, a student that believes the nature of knowledge is complex or uncertain may experience less anxiety when confronted with conflicting perspectives than an student who does not hold those beliefs (Muis et al., 2015).

Taken together, Barzilai and Chinn's (2018) theoretical work illuminates how students' epistemic thinking could influence the quality of their source evaluations.

Source Evaluations on the Internet

When examining the reliability of information online, students may compare a source's content-based, design-based, and epistemic features to their tacit or explicit epistemic ideals. For example, to be deemed trustworthy, a student may adopt the epistemic ideal that a reliable health website cites high-quality evidence to support its' claims. To examine whether this epistemic ideal has been met, the student may scrutinize the sources cited in a reference list or click on embedded hyperlinks to see where that evidence came from. Researchers have documented students' use of a variety of evaluation criteria during source evaluations, frequently noting students use of epistemic ideals (e.g., author's expertise, message accuracy, or purpose, Halverson et al., 2010, Ulyshen et al., 2015), content-based (e.g., Barnes et al., 2003; Kiili et al., 2008) and design-focused criteria (e.g., Gerjets et al., 2011; Cunningham & Johnson, 2016). Despite students' reliance on epistemic and non-epistemic evaluation criteria, some researchers have suggested that students do not consider epistemic features at all when evaluating the reliability of new information (e.g., Bråten et al., 2016; Wineburg, 1991) or use limited epistemic ideals to justify their acceptance or rejection of information (e.g., Barzilai & Eshet-Alkalai, 2015; Britt & Aglinskas, 2002; Greene et al., 2014, 2018). Yet, other researchers have observed high rates of students' epistemic ideal use (e.g., Kąkol et al., 2017; Halverson et al., 2010).

Mason and colleagues (2011) asked students to think out loud as they examined eight curated webpages presented in an offline environment. The researchers varied the webpages' authoritativeness, position toward the topic, and the evidence provided to gather students' spontaneous reflections about the sources. Their analyses revealed that most students reflected

on at least one epistemic ideal while examining the webpages, such as whether the source and its' evidence were scientific. Mason and colleagues' earlier work (2010) acknowledged that students require new skills to evaluate the authority or accuracy of internet sources, yet these researchers continued to design offline environments to assess such ideals and behaviours (Mason et al., 2011, 2018). Like Mason and colleagues, education researchers have predominantly examined epistemic ideals in controlled offline environments, including multiple documents contexts (e.g., Braasch et al., 2013; Bråten et al., 2009; Wiley et al., 2009) and hypermedia environments (e.g., Barzilai et al., 2020). As a result, findings about epistemic cognition in curated contexts have been inappropriately extended to a distinct environment—the unfiltered quagmire of the internet. Consequently, source evaluation trainings have been developed based on findings from these controlled environments (Mason et al., 2014; Wiley et al., 2009; Zhang & Duke, 2011), which undermines the efficacy of these trainings for internet source evaluations.

Whereas researchers using online environments have documented higher rates of students' epistemic ideal use (e.g., Kąkol et al., 2017), Halverson and colleagues (2010) identified university students' inappropriate use of epistemic ideals to evaluate online sources. To establish reliability, the researchers observed more than half the students employ important epistemic ideals, including assessing the source's credibility, followed by its' accuracy, objectivity and/or perspective of information presented, alongside content-based criteria. Despite the prevalence of epistemic ideals in students' written reports, the researchers highlighted the discrepancy between students' descriptions of selected sources as objective and credible and the contents of the source (e.g., biased data). The researchers attributed this finding to students' topic-specific beliefs; however, their metacognitive knowledge about what, when and how to use

these epistemic ideals may have also played a role in students' inaccurate website assessments. Although similarities between online and offline source evaluations exist, the prevalence of students' appropriate epistemic ideal use during online source evaluations is unclear. Barzilai and Chinn (2018) have outlined key guidelines to assess students' epistemic processes and developed offline interventions (Barzilai et al., 2020) to assess epistemic scaffolds using their guidelines; however, further educational research is needed to better understand the variety and use of students' epistemic processes in environments and tasks that more closely represent their online experiences.

The Present Studies

The purpose of the present research was to examine college students' epistemic thinking about authentic source evaluations. College students were selected because research has demonstrated students' limited use of appropriate epistemic ideals during source evaluations (Braasch et al., 2013; Halverson et al., 2010). In Study 1, students' metacognitive knowledge about epistemic ideals and processes were collected in focus group interviews. Four students' epistemic ideals were also examined during their assessment of two source types. In Study 2, students' epistemic ideals were investigated during their evaluation of two news articles.

The following research questions guided the studies:

1. What characterizes college students' epistemic metacognitive knowledge about source evaluations on the internet?
2. How do college students' epistemic ideals contribute to their overall source evaluations?

Based on previous findings, we expect students to describe a variety of epistemic ideals and processes related to epistemic and non-epistemic information features. Findings from Study 1

will be used to develop coding schemes for Study 2. We hypothesize that students will emphasize non-epistemic criteria to assess reliability. For students that rely on epistemic ideals, we expect them to outperform those who rely on content-based criteria.

Methods

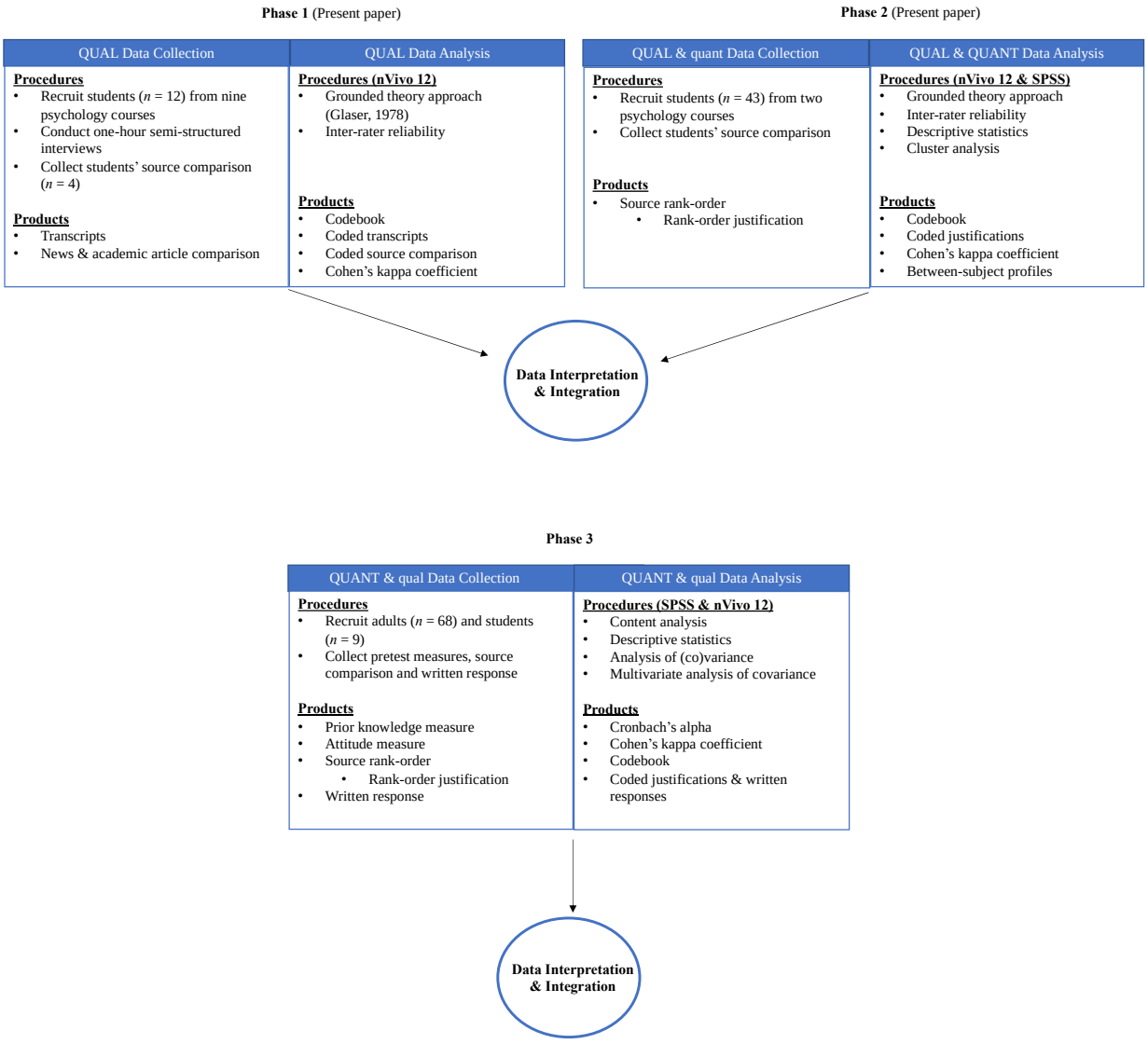
Study Design

As a research team, we approached our investigation from a pragmatist perspective, drawing on the strengths of diverse frameworks to understand students' source evaluations (Johnson & Onwuegbuzie, 2004). Following Creswell and Plano Clark's (2017) guidelines, we used a multiphase mixed methods design to assess students' strengths and weaknesses during source evaluations. A flow diagram depicting the design is presented in Note. . We use two notations to represent how emphasis was placed on each data collection and analysis method (Creswell & Plano Clark, 2011; Morse, 2003). For example, our use of "QUAL" in phase 2 indicates the emphasis on qualitative methods, whereas our use of "QUANT" indicates emphasis on quantitative analysis.

In phase 1, we conducted Study 1 using focus group interviews to examine students' epistemic metacognitive knowledge and a source evaluation task to collect measures of their reliable epistemic processes. In phase 2, we conducted Study 2 by gathering additional measures of students' reliable processes. Each phase was independently analyzed prior to integration of the results. The current investigation represents two phases of a larger study aimed at developing and implementing a training to improve online source evaluations.

Figure 1

Schematic of Multiphase Mixed Methods Design



Note. Figure adapted from White and colleagues' (2019) Figure 1.

Research Context and Setting

The studies took place at a public college in Québec that offers pre-university and career programs. Approximately 6700 students are enrolled in two-year pre-university programs each semester. Students can choose concentrations in arts and sciences, liberal arts, social sciences,

and visual arts, among others. In 2017, about 83% of enrolled students were 17-20 years of age. The student population represented more than 85 nationalities, with about 65% of students' reporting English as their mother tongue.

To complete phase 1 data collection, focus group interviews were conducted at the college in February 2020. Due to the campus closure associated with COVID-19, phase 1 data collection was completed remotely. To complete phase 2 data collection, students completed one in-class source evaluation task in February 2020.

Researcher Positionality

Our research team approached these studies with varied connections to the research setting. The second and third author had prior relationships with the college and the instructors involved in the research, which undoubtedly accelerated the establishment of trust between instructors and researchers. The first author also had a prior relationship with one of the instructors involved in the research. These connections provided insider knowledge to the researchers, which may have influenced our interpretations. To reduce potential biases related to our established connections, the first author was directly involved in recruitment and data collection at the college, whereas the other three co-authors engaged solely in data coding, analysis, and interpretation of results.

All four co-authors identify as highly educated, non-Hispanic White individuals. In terms of sociodemographic characteristics, three members of the research team more closely resemble the students' instructors rather than those involved in the study. Our lived experiences and backgrounds may notably differ from the students involved in our research. As such, our chosen position in relation to the study setting and participants is outsider with substantial knowledge of the study context. To reduce the influence of our experiences on interpretation of students'

source evaluation behaviours, we engaged in reflexivity exercises throughout data analysis and integration (Lutrell, 2010).

Study 1

Participants

Five psychology instructors were approached to include their classes in the research. Forty students from nine psychology classes were recruited through in-class announcements, and eighteen students volunteered to participate in focus group interviews. Twelve students attended one of the interview sessions. Following Morgan's (1988) recommendations, between 4 and 12 participants were scheduled for each session. Participants were asked a series of questions about their background and experience with source evaluations (e.g., How have your instructors discussed source evaluations in your courses?). Using this information, a profile was compiled for each participant, which can be found in Table 1. All cancellations occurred due to illness or unforeseen schedule conflicts. As compensation for participants' time, they were entered in a draw to win $100.

Procedure

After providing written consent, participants were contacted by email to sign up for a one-hour interview session. After attending the interview, four students completed a source comparison task as part of their course.

Table 1

Focus Group Participant Descriptions by Class

Participant pseudonym	**Age**	**Sex**	**Program**	**Year of study**	**Previous experience with source evaluations**
Class 1					
Jose	18	Male	Social sciences	2nd	Enrolled in literature course that examines reliability and "truth" in American non-fiction (e.g., memoirs)
Cameron	18	Male	Social sciences	2nd	Completed research methods course that examined the process of finding reliable sources and reducing bias
Class 2					
Sharon	19	Female	Social sciences	2nd	Attended six lessons presented by college's librarians on finding peer reviewed sources; Completed same course as Cameron
Michelle	18	Female	Arts & sciences	2nd	Evaluated primary and secondary sources for literature course term paper
Class 3					
Amanda	17	Female	Social sciences	1st	Completed sociology course that explored problem-solving using multiple perspectives
Charles	18	Male	Social sciences	2nd	Attended lesson presented by high school librarian about finding sources in French
Class 4					
Dolores	17	Female	Liberal arts	1st	Attended lesson presented by college's librarians about the CRAAP test
Betty	18	Female	Social sciences	2nd	Taught younger sibling about importance of authority when evaluating controversial evidence
Vera	17	Female	Liberal arts	1st	Evaluated multiple perspectives for term paper on controversial topic; Attended same lesson as Dolores
Josephine	18	Female	Social sciences	2nd	Enrolled in social psychology course that examines the role of attitudes and bias in behavior
Class 5					
Will	18	Male	Social sciences	2nd	Used multiple forums, with varying levels of reliability, to answer personal inquiries
Class 6					
Jennie	18	Female	Social sciences	2nd	Enrolled in different section of the same course as Josephine

Focus Group Interviews

To better understand how students evaluate sources, participants were asked about their experiences examining information on the internet. The semi-structured interview questions were adapted from Barzilai and Zohar's (2012) interview protocol and included nine open-ended questions that explored their (a) criteria for establishing reliability (e.g., What features does a reliable website have?), (b) procedure for establishing reliability (e.g., What would you do if you found two websites that made conflicting claims?), and (c) beliefs about the influence of student differences on source evaluations (e.g., How do biases influence how information is created and interpreted?). Focus group interviews were audio-recorded and transcribed verbatim for qualitative analysis. Participants' epistemic metacognitive knowledge was inferred from their responses. See Appendix A for the full interview protocol.

Using the grounded theory approach, one transcript was independently evaluated by the first and third authors over three stages: initial, focused and theoretical coding (Glaser, 1978). First, the raters examined the transcription line-by-line to identify emerging themes brought up by participants. Emerging themes included stating evaluation criteria, describing the evaluation procedure and identifying student differences. These themes were discussed to develop a preliminary focused coding scheme. See Table 2 for a list of selected codes with illustrative examples. In the second phase, the raters synthesized larger segments of the text, examining each segment and constantly comparing that incident to previously coded segments (Glaser & Strauss, 2017). The raters again examined any disagreements to revise the coding scheme. Using the updated coding scheme, the first author coded the remainder of the transcripts in NVivo 12 and added novel codes to the coding scheme.

To establish inter-rater reliability, the two raters coded the initial transcript a third time using the revised coding scheme. Their agreement, as measured by Cohen's kappa coefficient, was initially established at .62 and all disagreements were discussed before another round of coding, which established the raters' agreement at .79, with substantial agreement (Landis & Koch, 1977). All disagreements were resolved and used to inform the final coding scheme. The first author reanalyzed the remainder of the transcripts using constant comparison. In the final phase, the raters integrated the focused codes using a combination of Glaser's (1978) process and dimension coding families. See Appendix B for the full coding scheme.

Source Comparison Task

Four interview participants also submitted a source comparison task as part of their course. For this task, participants were asked to find a news article about a research study in an unreliable publication (e.g., The National Enquirer) and then assess the reliability of the source and content. They were encouraged to use online databases and the internet to support or refute the claims made in the article chosen. Participants were asked to document their research and give an overall assessment of the news article's reliability in a written response of approximately 1250 words (see Appendix C).

Table 2

Selected Interview Codes with Illustrative Examples

Micro-codes	Illustrative examples (Participant pseudonym)
Macro-code: Stating evaluation criteria	
1. Author expertise	1. Well it can have like the… like words about the author like "Oh he studied this for this, and oh he's got a bachelor's degree in this" and then you're kind of just like "oh, okay he knows what he's talking about." (Amanda)
2. Corroboration	2. Yeah, you could but in like really good research you're at least going to have one other person to support and say like "oh I found this as well." (Sharon)
3. Currency	3. So we have to look at the currency. (Dolores)
4. Design	4. Another thing is like, if you read through the article, this is a really particular thing that bothers me. But they cut the text up like, between pictures and quotes and ads. And then there's also the fact that even then, their using up a lot of space to make it seem longer and more professional but they're really saying something really simple and they're not really communicating anything, they're just word vomiting and what they're saying doesn't make, necessarily make coherent sense. It's just there, this is information and it may be a little biased and it's just not professional. (Dolores)
5. Evidence quality	5. Well like in terms of that I think if it's such a divided opinion maybe look at how they came to the conclusions and sort of then decide which method is more reliable. (Michelle)
6. Funding	6. I mean I think it plays a really big role. Like sometimes there's like when you look on websites, you'll see like popup ads and um especially on my, like when you get into like really untrustworthy ones, they have a lot of them and they have flashy and really catching titles. Like even I'm like "I so don't want to see that" and that's sort of how they make money, by you clicking on it. So, when I see like a lot of ads I sort of, not trust it. (Josephine)
7. Objectivity	7. Academic articles, no colour, black and white. It's just like facts, this is why they're so trustworthy right, there's no sugar coating on it. (Cameron)
8. Peer review	8. Exactly. You have to… A book or an academic article has to go through a process before being put out into the world but um… someone writing for Buzzfeed or someone making a video in their basement ranting about something they don't like, that's not going through anyone else it's just them and their information and what they want to say and then it's out there and anyone can see it. And if you don't think about that in context you can easily think "well these two pieces of information, I found them in the same place so even though one is a book and one is a, not a journal article but just a random Internet article, then they're about the same thing, I found them in the same spot, then they're probably about the same value". But you have to think about the process that one of them had to go through. Like a book had to be written, and then edited by the person, and the probably edited by an actual editor, and then had to be approved by a publisher and then… (Vera)
9. Purpose	9. The person has nothing to gain usually is trustworthy. (Josephine)

10. Tone	10. Ooh. I read more about it, try to see other people's perspectives about it or maybe it's the way they worded it that made it seem fishy. (Amanda)
11. Truthfulness	11. Um, well if they are honest about where they get their information from so they're gonna cite where they got the information. (Betty)
12. Type	12. Let's say if I read um a blog online, and then I read a book right after, I'd more likely believe the book because, from what I've been taught, it's something that's more valid than something just written by like I remember I don't know if I, one of my old teachers say " you never know who's writing on the Internet, it could be under a pseudonym or anything" and it's a lot safer to trust in a book than something online. (Jose)
13. Venue	13. Just like, they've proven themselves to not be faulty and they've been giving accurate information in the past. (Sharon)
14. Writing quality	14. sometimes the quality of the writing you can find, if there's a lot of spelling mistakes or something or improper punctuation, I look for that sometimes and you can kind of tell that it wasn't written properly. (Jose)
	Macro-code: Describing evaluation procedure
1. Assess bias(es)	1. And then by detecting that, check what… check comparatively what biases there are. I was wrong when you said you can't completely check bias. Because you can check… we know that there are certain biases that exist that are common, and we can be aware of that and we can see it but its only by comparing it to other sources that we know don't have that bias, that we can tell what's a bias and what isn't. (Dolores)
2. Assess currency	2. Usually I'll search it up, and I'll see what pops up. I'll click the first one cause that's what most of us does I think, and I'll start reading and then if the website seems like a bit, I don't know why but outdated, I don't know why but I won't really use the information. If it looks more modern and you can find who wrote it even if they have a little word about the author that's nice too. (Amanda)
3. Assess the source	3. But it turns out the source was completely unreliable, and that's exactly what they want to say. And they're like "oh it's okay it's a 2011 research study, don't worry about it, this guy said it". But like, who is this guy? You search up this guy, he's no one. And then it just goes on. (Charles)
4. Boost disciplinary knowledge	4. Trying to find as many unfamiliar terms in whatever research you find, like try and find the definitions for the terms you're unfamiliar with, I guess? (Michelle)
5. Compare multiple sources	5. Um, but generally if I'm actually researching something and I want to make sure that all of my information is viable I just, if I think a site looks suspicious I just like, I keep the information in mind but I go look at other sources and see if their stories match up. That's basically the extent of what I do. (Dolores)
6. Compare multiple perspectives	6. I try to, I always try to think of it from like, the opposite point of view. Especially when it comes to arguments, or like solving arguments between people. Like, okay you might think you're right, but have you considered it from this point of view so why the other person you're arguing with might obviously also think they're right. (Sharon)
7. Evaluate evidence	7. Well like in terms of that I think if it's such a divided opinion maybe look at how they came to the conclusions and sort of then decide which method is more reliable. (Michelle)

Macro-code: Identifying individual differences	
Individual differences	So long as, it's about bias, cause so long as we keep bias, cause like it has to do with where bias comes from cause we all have our own personal biases through our own personal experiences. But so long as we keep that bias in mind, I don't think we're in danger of falling prey to it [our own biases]. (Dolores)

Following Halverson and colleagues (2010), we assessed the reliability of each news article chosen by participants. Their written responses were independently analyzed by the first and third author using the evaluation criteria described during the focus group interviews, including venue's reputation, author's writing quality, and perceived biases. We also identified an additional emerging theme: source's target audience. The raters' agreement, as measured by Cohen's kappa coefficient, was initially established at .60. Each disagreement was carefully examined and discussed before another round of coding, which established the raters' agreement at .75, with substantial agreement (Landis & Koch, 1977). The first author reanalyzed the four written responses using constant comparison. See Appendix D for full coding scheme.

Study 1 Results

To answer research question 1, we present indicators of participants' epistemic metacognitive knowledge related to source evaluations on the internet. Focus group interviews were conducted on weekdays between 10:00 and 16:00 hours and lasted between 44 - 58 minutes (M = 50 minutes). Group sizes were between 1 - 5 participants each (M = 2.4). Whereas participants identified a wide range of evaluation criteria and processes, we highlight their descriptions of (a) author's credibility, (b) venue's design and purpose, (c) source's evidence, and (d) biases. To answer research question 2, we elaborate on participants' epistemic ideals from a source comparison task.

Focus Group Interviews

Students Assess Author's Credibility First

All five groups emphasized assessing author's credibility, an important epistemic ideal and process, before proceeding to evaluate other aspects of the source. The publication venue's *transparency* about the author played an important role in two participants' assessments of two

specific source types. For example, Will explained their critical approach toward information available on forums, “I don’t really trust them [forums] because it could be anyone and it is anyone.” Similarly, Amanda questioned “the whole Wikipedia thing, you know. Anyone could have written it and you wouldn’t know because there’s no like list of authors.” When asked about the ease of finding information about the author to assess their credibility, Josephine explained, “Um, I mean if it’s a trustworthy source, then usually it is [easy to find information]. But if not, usually like it’s hard to find the author or something.” In online environments, Greene (2016) highlighted that the difficulty of source evaluations is the lack of transparency or availability of information to make reliability judgments. To navigate this epistemically unfriendly environment, Josephine explained their broad approach toward information:

> Well I think we just, like for me I just always keep my guard up. Like I never trust anything. Cause that’s the thing it could be fake, it could be a fake person, it could be a straight up lie. We don’t know the motives of the person behind this [information] so… Never trusting anything you see, double check things when something sounds a little weird, a little fishy and also understand that like most people are biased and have like, nothing is objective really online.

Although Josephine described “never trusting anything,” they also suggested evaluating information through corroboration because, as Dolores explained, “You can’t operate thinking that everything is completely unreliable, so there is still a highest standard of reliability.” Vera added, “Right, as close as you can get to perfection without actually achieving it.”

When author’s information is available, seven participants identified that knowing the *author’s credentials* or experience were important indicators of their expertise. Amanda used their own experience to depict how the information they could provide about psychology compared to someone with a degree:

> Like I could write about let’s say psychology, but I have no degrees in this field, I’ve never really studied it, here’s what I know. But I’m not an expert so my information could be kind of off or not accurate compared to like someone who’s got a bachelor’s in

psychology.

Josephine supported this explanation, stating that credible authors are "People with degrees, like who actually studied it in school, not just as like for fun." To assess an author's expertise Amanda and Jose suggested investigating the author's experience and previous work.

Three participants specifically described the *author's reputation* as an influential indicator of their credibility. Jose and Michelle brought up the reception of an author's work to assess their reputation. Jose suggested, "Maybe you could look up the author and see what other stuff they've done in the past, you could see if there's any piece of news about them being scandalous or anything." Michelle added a list of questions to ask when examining an author's work:

> Uh, relevance as well as like, was their previous writing particularly notable? Did a lot of people find it a good article? Or did... was there like backlash on the article? Does it seem like yeah, it's peer reviewed but it's not as good quality as some other peer reviewed article?

Whereas Jose and Michelle described processes, Cameron introduced an important aspect of reliability judgements: previous experience with the source. They explained, "Well, you look at the source first. Like if it's a source [author] you know for a fact is mostly reliable, then it's most likely reliable." Sharon refined Cameron's explanation, stating:

> Well reliable is kind of like they like they've proven themselves. Like they haven't lied in anything else, or if they have it's been like minor discrepancies. […] Just like, they've proven themselves to not be faulty and they've been giving accurate information in the past.

Beyond an author's credentials and experience, which is commonly manipulated by researchers (e.g., W. S. Jung et al., 2018; Mason et al., 2011; Thon & Jucks, 2017), participants highlighted both internal and external indicators of an author's credibility. Researchers have found that

students with low prior knowledge or topic familiarity rely on internal source features to establish reliability (E. W. Jung et al., 2016; Lucassen & Schraagen, 2013).

Students Use Design Features to Determine a Venue's Purpose

Like their assessments of author credibility, participants also described evaluating publishing venue's credibility, specifically news venues'. These assessments ranged from looking at the publishing venue to investigating indicators of its' purpose. Vera explained their emphasis on the information's *venue* during source evaluations, explaining:

> I like to look at where things come from, if I'm seeing, I guess an article, I'll check to see which news company put it out. Because if it's something like Breitbart or the Daily Mirror or Buzzfeed, it's not going to be as reliable as some other things [venues].

Cameron identified the primary purpose of news venues as selling ad space by gaining readers:

> Well, news article are made to just, first of all [to] make money, hook viewers in and as a sort of propaganda for the political views. So of course, there's going to be colour and a bunch of things that are supposed to make you angry or make you sad about something.

Four participants associated the *presence of advertisements* with this primary purpose. Josephine described how the presence of advertisements reduced their perception of trustworthiness:

> I mean I think it plays a really big role. Like sometimes there's like when you look on websites, you'll see like popup ads and um especially on my, like when you get into like really untrustworthy ones, they have a lot of them and they have flashy and really catching titles. Like even I'm like "I so don't want to see that" and that's sort of how they make money, by you clicking on it. So, when I see like a lot of ads I sort of, not trust it. […] I think well also, well I said it before, ads. If there are like a lot of pop up ads, I don't trust it. If there's just a couple [of ads] and they're normal things like let's say like companies that you know, like big companies, then I'd trust it.

Josephine distinguished the trustworthiness of websites that have a large quantity of popup ads with catchy titles from websites that have fewer popup ads that come from well-known companies. Dolores, who also noted the quantity of advertisements, added:

> I look like, I personally like to look at the volume and quality of the ads that are being shown to me, because if there are a lot of them and they're like fishy and dodgy looking it definitely lowers the, um, the uh reliability of the site a lot. Um, and then the more

> there are, the more there are the more they're [the venue are] being paid by all, the more they need to be paid by all these smaller companies and perhaps illegal companies, so…

Dolores pointed out that venues receive funding from advertisers on their website, and Vera delineated how a venue's purpose could influence the accuracy of the information published:

> And again, because a lot of news sites now, can only support themselves by people clicking on their articles, they're more inclined to make them eye-catching or make them short or um, mis-quote studies, or pick something that doesn't really have anything to do with the story they're telling to put on the article so that people will click on it and read it. And it's a big thing, a big thing for surveys especially, where like someone will study, uh I think it was someone studying chocolate, or like flavanol levels in chocolate in pregnancy. And then a news site covered it and said that you should eat chocolate during pregnancy and that was their title, which was not what the study said. But because it had the word study and chocolate in it, and pregnancy, it was kind of related so people could get away with it. So, because it's all very motivated by […] personal interests. And by trying to get people invested in things, everything is going to be very dramatic and very eye-catching. And it's difficult to tell what's being eye catching because its important and people need to know about it, and what's being eye catching because someone wants money.

Beyond advertisements, four participants indicated that design features played a role in their reliability assessments. Whereas Josephine associated "nicely formatted" websites with ones that could be trusted, and Amanda preferred "more modern, not as loud" websites, Dolores explained that the layout of the webpage may distract from the content:

> Another thing is like, if you read through the article, this is a really particular thing that bothers me. But they cut the text up like, between pictures and quotes and ads. And then there's also the fact that even then, their using up a lot of space to make it seem longer and more professional but they're really saying something really simple and they're not really communicating anything, they're just word vomiting and what they're saying doesn't make, necessarily make coherent sense. It's just there, this is information and it may be a little biased and it's just not professional.

Vera extended Dolores' layout description, stating that modern formatting does not automatically signify reliability:

> Modern doesn't necessarily mean reliable source. Because again, a lot of very good journalism has been done through things like documentaries which are very visual mediums. But it's the sort of thing where if you're writing a piece and you break it up with videos related to that piece specifically its um… Going back to the Syrian refugee

> crisis thing, I did so much research for that I need to talk about it somewhere. There was one that I saw, it was a *New Yorker* thing, and it was an article, but it was an article about a documentary. So obviously they included clips from the documentary, quotes from the documentary and links to the entire documentary at the end so I could get the information from it firsthand. Lots of images, lots of pictures, lots of quotes. And that was a reliable source because what they were talking about in the pictures and ads, ads… not ads, videos were taken from something relevant. If it's, like you're going through an article and they have a bunch of reaction memes just interspersed throughout it for the sake of comedy, like Buzzfeed does, then it's more like 'oh they're trying to cater to people with smaller attention spans or trying to grab your attention again and make you stay and look at the funny pictures kinda thing, so it depends on the context.

Taken together, a venue's design features are closely tied to their primary purpose. Researchers have recently documented the deleterious impact that some advertisements have on students' reliability assessments of news websites (e.g., native ad vs. display ad, Aribarg & Schwartz, 2020; Krouwer et al., 2020). Yet, these features have not been explored in relation to students' epistemic cognition on the internet.

Students Compare Sources to Evaluate Evidence

Eleven participants reported that *corroboration* was an important aspect of determining whether a website is reliable. They described internal indicators of reliability, such as reference list and peer review process, and external comparisons with fact-checking websites and other online sources. Amanda outlined the influence of corroboration on the evaluation process:

> Well, a good rule of thumb that I've always learned is if multiple people are saying the same thing, then it should be the right thing. So, if this website's saying this and this website is not saying the same thing then maybe there's some… something fishy, but if the both of them are saying quite similar things then I guess that the information is correct.

Jose agreed, explaining that failure to corroborate a claim can be an indicator that an evaluation is needed: "say for an online article, you read about one thing and then you go search it up and if you read something that matches it, but not quite you have to, I start to question it." When encountering conflicting information, Michelle suggested, "I think if it's such a divided opinion

maybe look at how they came to the conclusions and sort of then decide which method is more reliable." Cameron summarized this process as "see[ing] the trail from where they [the author under scrutiny] got the original information." Unfortunately, as Jennie described, attempting to corroborate evidence can become confusing and lead to giving up on the process:

> Well what I do is I try to verify in other places. If what this is true or if this is only that person or that website that is saying that. So, I try to inform myself, and sometimes it's really confusing cause one thing says one and then another thing says another thing and I'm like I just give up I'm like "I'm not going to believe anything," cause It's really hard, you could get really confused.

To help reduce the likelihood of experiencing what Jennie described, two students elaborated on ways to internally examine the quality of a source's evidence. Within sources, Betty distinguished trustworthy sources' *citation practices* from untrustworthy sources:

> Um, well if they are honest about where they get their information from so they're gonna cite where they got the information. Cause a lot of times you're getting it [the information] like second hand, third hand, whatever. But if you can see the trail from where they got the original information. Like if you were lying, you'd be afraid to show your sources, if you have any.

Vera identified the source's *peer review process* as another internal factor to differentiate the quality of sources:

> You have to… A book or an academic article has to go through a process before being put out into the world but um… someone writing for Buzzfeed or someone making a video in their basement ranting about something they don't like, that's not going through anyone else it's just them and their information and what they want to say and then it's out there and anyone can see it. And if you don't think about that in context you can easily think "well these two pieces of information, I found them in the same place so even though one is a book and one is a, not a journal article but just a random internet article, then they're about the same thing, I found them in the same spot, then they're probably about the same value". But you have to think about the process that one of them had to go through. Like a book had to be written, and then edited by the person, and then probably edited by an actual editor, and then had to be approved by a publisher and then…

Whereas Betty and Vera identified internal processes, Dolores elaborated on the contribution and limits of corroboration during a source evaluation:

> So, what I found was very helpful was I'd find a source and read through it, figure out if it seemed somewhat legit and then I would go to a fact checker site and I would see how they rated it and why. And it's obviously like fact checking a fact checking site it just, there comes a point where you have to be like okay, I've done my due diligence, this [source] is as trustworthy as something I can find can be and then I would use it as a source.

When students assess multiple sources, they must reconcile the similarities and differences between sources, their claims, and the support presented (Barzilai & Zohar, 2012). All ages struggle to engage in this epistemic process (Eshet-Alkalai & Chajut, 2009). Often students may decide not to engage further with information if it contradicts their attitude and they may not assess the validity of their own beliefs (Hart et al., 2009). Bråten and colleagues (2011) delineated the impact of failing to integrate perspectives, explaining that a student may espouse false beliefs from biased sources.

Students Detect Bias by Evaluating Multiple Perspectives

All five groups brought up that individual differences, such as biases, attitudes and purpose, influence how information is both evaluated and created. Acknowledging the difficulty of integrating multiple perspectives, Sharon offered the following example:

> I try to, I always try to think of it [the topic] from like, the opposite point of view. Especially when it comes to arguments, or like solving arguments between people. Like, "okay you might think you're right, but have you considered it from this point of view?," so why the other person you're arguing with might obviously also think they're right. Like, it goes back to, I've seen this I don't remember where I was but, like this guy was holding up a book to the camera and it says, "This book is green," but the back [of the book] is white. And so, you're thinking, "no, the book is white," but then he turns around and the book is green. So, for him the book is green, but for us it's white. So, it's like you really have to see it from the other person's point of view, which can be completely hard.

In Sharon's example, the colour of the book serves as evidence of the claim for the author and the evaluator. Each students' perspective is independently valid and through Sharon's integration

they can observe the accuracy on either side. The participants identified their own biases as potential inhibitors of their source integration. Vera described the powerful role student differences play as antecedents to students' source evaluations:

> Because as a person, you're never going to be unbiased 'cause you have your own beliefs, you have the way you were brought up, you have the environment you were brought up in, you have the purpose of what you are writing, or what you're researching or what you think about a certain subject, and it's always going to be there in the back of your mind. So even if you find a source and it is very reliable and it meets all the qualifications of the CRAAP test and it's objectively a very good source, if it disagrees with you then there's going to be some little sensation in the back of your mind going "something is wrong with this" and what's wrong with it is the fact that it doesn't agree with you. But depending on the situation you could take that as "ok I'm wrong and I need to look at other sources to see what matches up with my thoughts and if it's as reliable as this or if the general consensus is this is right, I was wrong," [Path A] or you could say "I like my opinions better than learning things, so I'm going to close my computer and continue believing what I believe" [Path B]. And it depends on the situation and the person, where they're going to go from there.

As Vera explained, students' cognitive differences may influence who and what they choose to believe. Even after deeming a source reliable, as Vera pointed out, a student may choose to disregard information that does not align with their existing attitude on a topic. Rather than discounting information that conflicts with existing attitudes, Dolores suggested that students assess their attitude by "Try[ing] to find as many pieces of information as possible to find, and based on the information that you find figure out [what] your opinion on the thing [topic] is, and then kind of move on from there."

Six participants discussed their critical approach when examining the presence of *author biases*. Dolores described the role of surveying multiple sources to assess author biases:

> We as readers need to keep in mind that that person [the author] is going to have biases that we can't necessarily detect immediately, and that's why we look at other sources to see what those biases may be […] in this situation I personally don't think there's any way you can actually like check what someone's biases are cause they're presenting you information. So, you… in this case you need to look comparatively. You need to compare it to other sources. See what other... sources in the other… biases in the other sources you can detect. And then by detecting that, check what… check comparatively what biases

> there are. I was wrong when you [I] said you can't completely check bias. Because you can check… we know that there are certain biases that exist, that are common, and we can be aware of that and we can see it, but it's only by comparing it to other sources that we know don't have that bias that we can tell what's a bias and what isn't [a bias].

As Dolores pointed out, an author's bias can be identified in relation to other sources that may hold distinct biases and perspectives. Three participants also identified that an author's biases can be determined by assessing their *tone.* Vera provided an example:

> Yeah, so, or if someone is writing like into a lot of anti-feminist and alt-right things because I… sometimes you need to get angry. And they'll often just phrase it as "females" and then just "men", and that sort of like, change in terminology is just like "oh okay, I see what you're thinking from that" and you can usually pick up [their biases]. It's usually not that obvious, but you can pick up just by like phrasing, tone, and the way that they're addressing the audience too.

The participants tapped into important indicators of author's implicit biases. Researchers have associated students' source evaluation skills with their ability to integrate sources (Rouet & Britt, 2011). To better understand this relationship, four of the participants completed a task where they evaluated and integrated two sources in a written response.

Source Comparison Task

To answer research question 2, four participants' epistemic ideals were inferred from their assessments of a news article's reliability. Participants' news articles were retrieved from websites that report neuroscience research (e.g., *Neuroscience News*), featuring studies from science-based multidisciplinary journals (e.g., *Nature*). Jose and Sharon indicated that the news articles they chose inaccurately portrayed the academic article, whereas Cameron and Michelle determined that their news articles were credible after comparing the source to its academic counterpart. The characteristics of each participant's evaluation are summarized and assessed.

Jose Identified the Author's Bias

Jose's source comparison was characterized by their evaluation of the news article's purpose, which differed from their emphasis on non-epistemic evaluation criteria during their interview. Jose criticized the news article they chose based on the author's evident biases:

> This article [news article] briefly mentions that 'A research team parsed 68 current cannabis studies last year on behalf of the Canadian government,' but in this article he [the author] only talks about one [*bias*], which seems kind of biased against the legalization of cannabis, this article only mentions the cons of the drug but none of the benefits, he doesn't link these benefits nor does he talk about them inside his media article [*argument*]. By writing this he wants to convince the audience that cannabis is ultimately bad for you and should not be legalized [*purpose*], even the vocabulary he used makes this evident, such as 'potheads' [*bias, tone*]. Maybe he even wanted to try and divert voters from voting Liberal in the election [*purpose*], but this is just crazy speculation so it most likely is not true [*epistemic judgment*].

In this passage, Jose argued that the news article did not meet their epistemic standard for a trustworthy source because the author presented a one-sided argument using biased evidence and terminology. Jose identified that other perspectives were missing from the author's argument, an important reliable epistemic process, and accurately deemed the news article unreliable. We noted that the authors' bias extended to their source selection, which heavily relied on other news articles reporting on research studies and linked to their previous articles on the topic.

Michelle Identified the New Article's Problematic Tone

Like Jose, Michelle's source comparison was characterized by their examination of the sources' accuracy and purpose. Michelle noted the difference between each author's tone:

> Another difference between the articles were their tone. The media article had a much more exciting tone [*tone*], using sensational language such as "key role" or mentioning that one of the researchers shared a Nobel Prize [*bias*]. They also had the article tell a story and focused more on how the researchers came to their discovery than the discovery itself [*tone*]. The original research had a more objective tone to it [*tone*]. Instead of using persuasive language, the researchers provided the data and explained their reasoning [*objectivity*]. This differs from the media article as it allows the reader to come to their own conclusions.

Michelle identified that the news article's author used labelling and sensational language to persuade the reader of the importance of the researchers' discovery. Yet, Michelle explained:

> For the most part, the two articles gave the same information without any grave inaccuracies. This is most like[ly] due to the fact that the media article heavily relied on quotes from the authors [*evidence*] of the original research article.

Michelle considered the differences in tone acceptable and accurately rated the news article as a reliable representation of the academic article. We noted that the tone Michelle identified revolved around the formation of the research team and the research process, rather than the team's findings.

Sharon Analyzed the News Article's Evidence

Sharon also noted that the sources conveyed similar information; however, Sharon explained that the news article used the academic article as evidence of a claim different than the academic article's purpose:

> The "popular magazine", although it had a lot of congruent information and statistics found in the original experiment, the article also brought up other research not mentioned in the original study as well as referenced other works [*uniqueness*]. The magazine made a connection between reading fiction and the amount of television sets found in households these days. Furthermore, they made references to an article which made a correlation between the presence of television in a child's life and the connection it has with the theory of mind, being their ability to understand the feeling's and beliefs of others. All of these elements do not present themselves in the initial study aside from a brief mention of the theory of mind.

They added that the news article's use of information from the academic article was not "the most accurate either" as it oversimplified key elements of the research. Given the discrepancy between the academic article's purpose and how it was used by the news article, Sharon deemed the news article unreliable. After reviewing both articles, we noted that the news article referenced a blog from the research institution, rather than the original academic article, which quoted the lead researcher. Sharon's assertion that that news article oversimplified the research article's content may have stemmed from a judgment of the news article's peripheral coverage of

the academic article's topic. We rated the news article as reliable, judging the simplification acceptable for the news article's target audience.

Cameron Assessed the Target Audience

Unlike the other participants, Cameron attributed differences between the sources to their distinct target audiences:

> The academic paper was more complex, used more scientific language [*tone*] and was more precise overall. It had graphs and images and a lot of proof [*evidence*] and overall the information was just better and more complete in its information about the retrosplenial cortex. The media article on neuroscience news however could be read by a larger audience than the research paper [*target audience*].

Despite these differences, Cameron explained that the news article summarized the academic article well, stating:

> Even if the media article is a small summary of the full research [*uniqueness*], it seems like it is quite accurate in its simplification. The method by which the scientists tested the mice and the key results were well and clearly explained by the media article [*writing quality*].

They accurately rated the news article as reliable. We noted that the news article included the research article's abstract and details to easily access the original source.

Whereas the participants' use of evaluation criteria differed as they assessed the sources, they exhibited epistemic ideals consistent with those brought up in the focus group interviews, adding only "target audience" to the coding scheme for Study 2. To understand how evaluation criteria contributed to the accuracy of reliability rank-ordering, in Study 2, we asked participants to compare two authentic news articles.

Study 2

Participants

A similar sample (n = 43, 28 female) was recruited from the study site for Study 2. The sample primarily represented students enrolled in their second year (73%) of a social sciences program (86%). Students self-identified as White or Caucasian (57%), Asian (16%), Black or

African Canadian (13%), mixed race (11%), or Indigenous (3%). Their mean age was 18.7 years (*SD* = 0.9). Participants volunteered to take part in the research and were entered in a draw to win $100 as compensation for their time.

Procedure

After providing written consent, participants were given an activity packet which included two news articles and a rank-ordering and justification form. Participants were instructed to read both articles, rank the sources from most to least reliable, and then explain why they assigned each source its rank. Participants were given a total of 45 minutes to complete the tasks.

Materials

Participants were presented printed-out versions of two news articles that used similar sources but held distinct perspectives about medically assisted death regulations in Québec. The articles differed in terms of author expertise, publication venue and currency. See Table 3 for a description of the articles. The articles were 750 words and 975 words, with Flesch-Kincaid readability ease scores of 45.5 and 48.2, respectively. From top to bottom, each article included the: (1) title, (2) author description, which included the author's name and job title, (3) publication type, (4) publication venue, (5) publication date, and (6) body of the text. The order of the articles was randomized to eliminate order effects.

For both articles, the first and third authors evaluated indicators of the source's currency, relevance, authority, accuracy, and purpose to rank-order the articles based on their reliability. Their agreement, as measured by Cohen's kappa coefficient, was established at .67, with substantial agreement (Landis & Koch, 1977). Despite disagreements, both authors ranked the articles in the same order. Each disagreement was carefully discussed and resolved.

Table 3

News Article Descriptions

Article	Title	Author Description	Publication Venue	Publication Date	Main Claims	Position regarding medically assisted death
1	Quebec court strikes down restriction to medically assisted dying law, calls it unconstitutional	Tu Thanh Ha, **National Reporter** Kelly Grant, **Health Reporter**	The Globe and Mail, **Health section**	**September 11, 2019**	The lack of access to medically assisted death in Quebec is **harmful and a denial of rights** to those suffering from debilitating, lifelong conditions.	**For**
2	Quebec won't challenge medically assisted death court ruling	Philip Authier, **Journalist**	Montreal Gazette, **News section**	**October 3, 2019**	The government will not contest the ruling of a Quebec court to allow the medically assisted deaths of two students and **will look into relaxing the laws**.	**Neutral**

Note. Differences influencing evaluations are in bold.

Measures of Epistemic Cognition

Overall Rank-ordering

Participants were asked to rank-order the articles from the most to the least reliable. They were provided a randomized list of the articles and a blank space to write the rank assigned. Their rankings were compared to the authors' rank-order, and correct rankings were given a score of 1, whereas incorrect rankings were given a score of 0 (Mason et al., 2014, 2018).

Justification for Source Evaluation

Participants were also asked to justify their rank-ordering in a written response. Two independent raters qualitatively analyzed participants' responses for evaluation criteria using the coding scheme from Study 1 (see Table 2). The coding scheme was tested by the first and third authors using seven participant responses (16% of sample). Their agreement, as measured by Cohen's kappa coefficient, was established at .75 for the first round. Each disagreement was carefully examined and discussed to update the justification coding scheme (i.e., add micro-code "date," see Table 4). A second round of coding was performed using another seven participant responses (16% of sample), and the final inter-rater reliability was established at .91, with almost perfect agreement (Landis & Koch, 1977). The first author coded the remainder of the justification responses in NVivo 12. One point was awarded for each evaluation criteria mentioned (Mason et al., 2014, 2018). The micro-codes were summed to create six macro-level categories for cluster analysis: currency, relevance, authority, accuracy, purpose, and other. See Appendix E for the full coding scheme.

Table 4

Selected Justification Codes with Illustrative Examples

Micro-level category	Description	Example (Participant pseudonym)
	Macro-level category: Currency	
Date	An evaluative decision about a rank based on participant's consideration of various publication dates.	They also provided the original publication date. (Lillie) Furthermore, it is a more recent account of the event, which does not always guarantee reliability, but it may present a more accurate depiction of the current event. (Grace)
	Macro-level category: Relevance	
Alignment	An evaluative decision about a rank based on participant's consideration of webpage's alignment with their task definition. Participant may mention any of the following: 1. topic relevance (e.g., central or peripheral to task) 2. appropriateness of evidence for task (e.g., type of evidence)	1. The article has details on the topic but I feel it's too long and not focused enough on the topic. (Bessie) 2. It may be less effective in 'understanding concerns about medically-assisted death' because it lacks a breadth of perspectives. (Owen)
Uniqueness	An evaluative decision about a rank based on participant's consideration of how unique the information presented on the webpage is in relation to other sources supplied in activity.	It read as though this article was a summary of Article 1 [more reliable article]. (Nicole)

	Macro-level category: Authority	
Author expertise	An evaluative decision about a rank based on participant's consideration of author or organization's expertise related to the topic.	This article is written by one author who is a general reporter and doesn't have the same standards/qualifications (Katherine) The authors' title is health reporter, thus specifying they might have more knowledge about this subject (Hannah)
Venue	An evaluative decision based on participant's consideration of the body that published the information (e.g., publisher's reputation.	The venue of publication seems like a reliable source (Myrtle) Article 1 [the more reliable article] is a more well-known source. (Sam)
	Macro-level category: Accuracy	
Evidence	An evaluative decision about a rank based on participant's consideration of the sources of information provided as evidence. Participant may mention any of the following: 1. type of evidence 2. source of evidence 3. quality of evidence	1. Seems to have more scientific research to back up its claim. (Christine) 2. It lacks a reference page, so it doesn't credit the information it's using. (Elijah) 3. Less credible speakers with quotes. (Alexander)
Corroboration	An evaluative decision about a rank based on participant's consideration of whether information can be corroborated or not in other supplied sources.	Objectively verifiable statements. (Samuel)
	Macro-level category: Purpose	
Argument	An evaluative decision based on participant's consideration of author or organization's argument structure (e.g., one sided or two sided argument).	It quotes people from both sides of the question. (Lucille) It also does not give multiple points of view on the situation. (Mia)
Objectivity	An evaluative decision based on participant's consideration of source's objectivity. Participant may mention any of the following: 1. subjective and opinion-based writing 2. fact-driven and objective writing with facts, statistics, or research	1. Based on the title alone, the one from the Montreal Gazette makes it seem like they have a strong one-sided opinion and it seems like they really wrote the article for clicks. (Kayla) 2. Was more informative and neutral compared to the first article. (Abigail)

Macro-level category: Other		
Length	An evaluative decision about a rank based on learner's consideration of the length of the webpages.	This article seemed to be short and not as detailed as the first one. (Abigail) The article was fairly short. (Lucas)
Writing quality	An evaluative decision about a rank based on learner's consideration of the readability of the webpage based on its syntax or spelling.	The wording in this article is easier to understand and follow along. (Ada) Typo within the second paragraph. (Taylor)

Study 2 Results

To answer research question 2, we present participants' rank-ordering justifications as indicators of their epistemic ideals. A hierarchical cluster analysis using the macro-level categories (currency, relevance, authority, accuracy, purpose, and other) yielded four distinct clusters of homogenous groups of cases (see Table 5). Further, a multinomial logistic regression was used (Press & Wilson, 1978) to determine the most parsimonious number of clusters and placed 100% of participants back into their original four clusters. The clusters (from largest to smallest) were labelled as the (1) accuracy group (Acc, *n* = 18), (2) accuracy, purpose, and other group (APO, *n* = 13), (3) relevance group (Rel, *n* = 9), and (4) authority group (Auth, *n* = 3). The accuracy group was characterized by their use of epistemic ideals that focused on the evidence presented. The APO group used a combination of content-based and epistemic criteria, such as writing quality, evidence, and objectivity to justify their rank-ordering of the articles. The relevance group was characterized by their use of content-based criteria that focused on the alignment of the content with their task and the content's uniqueness. Lastly, the authority group was characterized by their use of epistemic ideals that focused on the author or venue's expertise and reputation.

Table 5

Evaluation Criteria Use by Cluster Group

	Cluster			
Criteria	Acc (*n* = 18)	APO (*n* = 13)	Rel (*n* = 9)	Auth (*n* = 3)
Currency	.06 (.06)	.46 (.18)	.33 (.24)	.00 (.00)
Relevance	.83 (.23)	.38 (.18)	2.11 (.20)	.00 (.00)
Authority	.89 (.21)	.23 (.12)	.55 (.24)	3.00 (.58)
Accuracy	2.56 (.27)	1.77 (.36)	.22 (.15)	.00 (.00)
Purpose	.44 (.17)	1.38 (.33)	.89 (.39)	.00 (.00)
Other	.28 (.11)	2.08 (.21)	.55 (.24)	1.00 (1.00)

Note. Acc = accuracy group; APO = accuracy, purpose and other group; Rel = relevance group; Auth = authority group. Standard errors are presented in parentheses.

Regardless of the criteria they used, most participants (70%) produced an accurate rank-ordering of the articles. Participants that accurately rank-ordered the articles used more criteria overall (M = 5.6, SD = 2.2) to justify their rank-ordering decisions than their counterparts (M = 4.5, SD = 1.5). To examine the relationship between participants' justifications and their rank-ordering accuracy, a Kruskal Wallis test revealed that there were no statistically detectable differences in accuracy between the clusters, $\chi^2(3) = 3.826$, $p = .281$, with a mean rank accuracy of 26.1 for the relevance group, followed by 22.5 for the accuracy group, followed by 20.2 for the APO group, and 14.1 for the authority group. Surprisingly, the relevance group performed slightly better than the other groups.

General Discussion

Access to the internet has altered students' relationship with information. On a larger scale than ever before students are responsible for identifying high quality information. In Study 1, participants reported a variety of criteria and processes that they use as direct and indirect indicators of a source's reliability. Participants also described epistemic ideals that could be examined within a single source (i.e., vertical reading) and across multiple sources (i.e., lateral reading). To directly assess the epistemic properties of a source, participants highlighted looking at the author's credentials or experience as well as verifying evidence and determining author biases through corroboration. They described indirectly examining the properties by looking at the presence of advertisements and the author's tone. When assessing a news article, four participants used indicators in different ways to determine their source's reliability, with three participants accurately rating their sources. In Study 2, of the participants that accurately evaluated the printed-out authentic sources, those who justified their rank-ordering using relevance criteria outperformed those that did not.

The findings from these studies provide illustrative examples and empirical evidence of the epistemic ideals and processes described by Barzilai and Chinn (2018) in their Apt-AIR framework. Many education researchers have focused on epistemic ideals that explicitly assess the epistemic properties of a source's authority, accuracy, and purpose. The criteria captured in the present studies reflect the epistemic and non-epistemic evaluation criteria students have described in previous interviews (e.g., evidence or bias, Barzilai & Zohar, 2012) or used in rank-ordering and justification tasks (e.g., author credentials, Braasch et al., 2013; Wiley et al., 2009). Without instruction, some researchers have documented difficulty accurately assessing the reliability of sources and limited use of epistemic ideals to justify reliability judgements (Mason et al., 2014; Pretorius, 2018). Other researchers have highlighted students' frequent use of important epistemic ideals with limited success (Halverson et al., 2010).

Given the situated nature of epistemic cognition, Sandoval and Redman (2015) described these distinct findings as contextualized rather than divergent. For rank-ordering and justification tasks, as used in Study 2, similar designs using manipulated texts or offline environments have found that students rely on content-focused features and struggle to accurately rank sources with varying levels of reliability (Kiili et al., 2008; McGrew et al., 2018). Yet, the findings from Study 2 suggest that participants who accurately rank-ordered the authentic articles examined content-focused criteria to determine the sources' reliability. As Barzilai and Chinn (2018) suggested, the appropriateness of students' epistemic processes is determined by the success of their use. Participants that accurately rank-ordered the articles used appropriate ideals to achieve the task's aim of selecting the most reliable source. The present findings suggest that further research is required using authentic source evaluation tasks and environments. The results from the focus group interviews also point toward incorporating online environments in future

research to better understand direct and indirect indicators of reliability. Several participants introduced an epistemic ideal that they use to indirectly assess epistemic properties in an online environment: presence of advertisements as an indicator of purpose. Although the impact of advertisements has not been widely observed in education research (Eysenbach, 2002; Kąkol et al., 2017; Robertson-Lang et al., 2011), researchers in other fields have documented different types of advertisements' negative influence on perceptions of website reliability (Aribarg & Schwartz, 2020; Krouwer et al., 2020). Due to the controlled nature of many education researchers' studies, where curated texts and offline environments are used, advertisements have not been explored in relation to source evaluations. Using more authentic environments and tasks may illuminate the role of non-epistemic criteria, which are widely used, in source evaluations.

Limitations

Our promising foundation to develop a source evaluation training does not come without limitations. In Study 1, focus groups were small due to last-minute cancellations. Consequently, the researcher served as a participant rather than a facilitator in two groups (Nunkoosing, 2005). As a result, the first author's contributions about source evaluations may have influenced responses, reducing authenticity. These participants may have also been influenced by the first author's status due to similarities with their instructors.

In Study 2, printed-out versions of two news articles were used in the evaluation task. Like previous research, this task did not allow participants to look for external sources or examine features unique to digital environments to establish the sources' reliability. However, the content of the news articles was identical to the original text online, rather than a modified text created by the researchers (e.g., E. H. Jung et al., 2016). These materials more closely

reflected the type of information that can be found on the internet, even if it was not presented in an online environment.

Conclusions

Students struggle to examine the reliability of information online. Using both direct and indirect indicators of reliability, they assess a variety of features with just as many possible outcomes. The features identified in the present study were indicators of currency, relevance, authority, accuracy and purpose—criteria included in a popular evaluation tool, the CRAAP test. Although the CRAAP test describes important epistemic and non-epistemic properties of source evaluations, research is needed to better understand how these properties are weighed in authentic online environments.

Acknowledgements

We thank the participating instructors and Cara Singh for their support with data collection during the studies. This research was supported by a grant from the Social Sciences and Humanities Research Council (SSHRC). The content is the sole responsibility of the authors and does not necessarily represent the official views of SSHRC.

References

Aribarg, A., & Schwartz, E. M. (2020). Native advertising in online news: Trade-offs among clicks, brand recognition, and website trustworthiness. *Journal of Marketing Research, 57*(1), 20–34. https://doi.org/10.1177/0022243719879711

Barnes, M. D., Penrod, C., Neiger, B. L., Merrill, R. M., Thackeray, R., Eggett, D. L., & Thomas, E. (2003). Measuring the relevance of evaluation criteria among health information seekers on the Internet. *Journal of Health Psychology*, *8*(1), 71–82. https://doi.org/10.1177/1359105303008001436

Barzilai, S., & Chinn, C. A. (2018). On the goals of epistemic education: Promoting apt epistemic performance. *Journal of the Learning Sciences*, *27*(3), 353–389. https://doi.org/10.1080/10508406.2017.1392968

Barzilai, S., & Eshet-Alkalai, Y. (2015). The role of epistemic perspectives in comprehension of multiple author viewpoints. *Learning and Instruction, 36*, 86–103.

Barzilai, S., Mor-Hagani, S., Zohar, A. R., Shlomi-Elooz, T., & Ben-Yishai, R. (2020). Making sources visible: Promoting multiple document literacy with digital epistemic scaffolds. *Computers & Education*, *157*, 103980. https://doi.org/10.1016/j.compedu.2020.103980

Barzilai, S., & Zohar, A. (2012). Epistemic thinking in action: Evaluating and integrating online sources. *Cognition and Instruction, 30*(1), 39–85. https://doi.org/10.1080/07370008.2011.636495

Barzilai, S., & Zohar, A. (2014). Reconsidering personal epistemology as metacognition: A multifaceted approach to the analysis of epistemic thinking. *Educational Psychologist, 49*(1), 13–35. https://doi.org/10.1080/00461520.2013.863265

Barzilai, S., & Zohar, A. (2016). Epistemic (meta) cognition: Ways of thinking about knowledge and knowing. In J. A. Greene, W. A. Sandoval, & I. Bråten (Eds.), *Handbook of epistemic cognition* (pp. 409–424). Routledge.

Braasch, J. L. G., Bråten, I., Strømsø, H. I., Anmarkrud, Ø., & Ferguson, L. E. (2013). Promoting secondary school students' evaluation of source features of multiple documents. *Contemporary Educational Psychology*, *38*(3), 180–195. https://doi.org/10.1016/j.cedpsych.2013.03.003

Bråten, I., Strømsø, H. I., & Andreassen, R. (2016). Sourcing in professional education: Do text factors make any difference? *Reading and Writing*, *29*(8), 1599–1628.

Bråten, I., Strømsø, H. I., & Britt, M. A. (2009). Trust matters: Examining the role of source evaluation in students' construction of meaning within and across multiple texts. *Reading Research Quarterly*, *44*(1), 6–28.

Breakstone, J., McGrew, S., Smith, M., Ortega, T., & Wineburg, S. (2018). Why we need a new approach to teaching digital literacy. *Phi Delta Kappan*, *99*(6), 27–32.

Britt, M. A., & Aglinskas, C. (2002). Improving students' ability to identify and use source information. *Cognition and Instruction*, *20*(4), 485–522.

Chinn, C. A., Buckland, L. A., & Samarapungavan, A. (2011). Expanding the dimensions of epistemic cognition: Arguments from philosophy and psychology. *Educational Psychologist*, *46*(3), 141–167. https://doi.org/10.1080/00461520.2011.587722

Chinn, C. A., & Rinehart, R. W. (2016). Epistemic cognition and philosophy: Developing a new framework for epistemic cognition. In J. A. Greene, W. A. Sandoval, & I. Bråten (Eds.), *Handbook of epistemic cognition* (pp. 460–478). Routledge.

Chinn, C. A., Rinehart, R. W., & Buckland, L. A. (2014). Epistemic cognition and evaluating information: Applying the AIR model of epistemic cognition. In D. N. Rapp & J. L.G. Braasch (Eds.), *Processing inaccurate information: Theoretical and applied perspectives from cognitive science and the educational sciences* (pp. 425–453). MIT Press.

Chinn, C., & Sandoval, W. A. (2018). Epistemic cognition and epistemic development. In F. Fischer, C. E. Hmelo-Silver, S. R. Goldman, & P. Resiman (Eds.), *International handbook of the learning sciences* (pp. 46–55). Routledge.

Cho, B.-Y., Woodward, L., & Li, D. (2018). Epistemic processing when adolescents read online: A verbal protocol analysis of more and less successful online readers. *Reading Research Quarterly, 53*(2), 197–221.

Cresswell, J. W., & Plano Clark, V. L. (2011). *Designing and conducting mixed methods research* (2nd ed.). Sage Publications.

Cresswell, J. W., & Plano Clark, V. L. (2017). *Designing and conducting mixed methods research* (3rd ed.). Sage Publications.

Cunningham, A., & Johnson, F. (2016). Exploring trust in online health information: A study of user experiences of patients.co.uk. *Health Information & Libraries Journal, 33*(4), 323–328. https://doi.org/10.1111/hir.12163

Elby, A., & Hammer, D. (2010). Epistemological resources and framing: A cognitive framework for helping teachers interpret and respond to their students' epistemologies. I In L. D. Bendixen & F. C. Feucht (Eds.), *Personal epistemology in the classroom: Theory, research, and implications for practice* (Vol. 4, pp. 409–434). Cambridge University Press.

Eshet-Alkalai, Y., & Chajut, E. (2009). Changes over time in digital literacy. *CyberPsychology & Behavior, 12*(6), 713–715.

Eysenbach, G. (2002). How do consumers search for and appraise health information on the world wide web? Qualitative study using focus groups, usability tests, and in-depth interviews. *BMJ, 324*(7337), 573–577. https://doi.org/10.1136/bmj.324.7337.573

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American Psychologist, 34*(10), 906.

Flavell, J. H., Miller, P. H., & Miller, S. A. (2002). *Cognitive development* (4th ed., Vol. 338). Prentice-Hall Englewood Cliffs, NJ.

Gerjets, P., Kammerer, Y., & Werner, B. (2011). Measuring spontaneous and instructed evaluation processes during Web search: Integrating concurrent thinking-aloud protocols and eye-tracking data. *Learning and Instruction, 21*(2), 220–231. https://doi.org/10.1016/j.learninstruc.2010.02.005

Glaser, B. G. (1978). *Theoretical sensitivity: Advances in the methodology of grounded theory*. The Sociology Press.

Glaser, B. G., & Strauss, A. L. (2017). *Discovery of grounded theory: Strategies for qualitative research*. Routledge.

Greene, J. A. (2016). Interacting epistemic systems within and beyond the classroom. In J. A. Greene, W. A. Sandoval, & I. Bråten (Eds.), *Handbook of epistemic cognition* (pp. 265–278). Routledge.

Greene, J. A., Copeland, D. Z., Deekens, V. M., & Seung, B. Y. (2018). Beyond knowledge: Examining digital literacy's role in the acquisition of understanding in science. *Computers & Education, 117*, 141–159.

Greene, J. A., & Yu, S. B. (2016). Educating critical thinkers: The role of epistemic cognition. *Policy Insights from the Behavioral and Brain Sciences*, *3*(1), 45–53. https://doi.org/10.1177/2372732215622223

Greene, J. A., Yu, S. B., & Copeland, D. Z. (2014). Measuring critical components of digital literacy and their relationships with learning. *Computers & Education*, *76*, 55–69. https://doi.org/10.1016/j.compedu.2014.03.008

Halverson, K. L., Siegel, M. A., & Freyermuth, S. K. (2010). Non-science majors' critical evaluation of websites in a biotechnology course. *Journal of Science Education and Technology*, *19*(6), 612–620. https://doi.org/10.1007/s10956-010-9227-6

Hammer, D., & Elby, A. (2002). On the form of a personal epistemology. In B. K. Hofer and P. R. Pintrich (Eds.), *Personal epistemology: The psychology of beliefs about knowledge and knowing* (pp. 169–190). Lawrence Erlbaum Associates Publishers.

Hart, W., Albarracín, D., Eagly, A. H., Brechan, I., Lindberg, M. J., & Merrill, L. (2009). Feeling validated versus being correct: A meta-analysis of selective exposure to information. *Psychological Bulletin*, *135*(4), 555.

Hofer, B. K. (2016). Epistemic cognition as a psychological construct: Advancements and challenges. In J. A. Greene, W. A. Sandoval, & I. Bråten (Eds.), *Handbook of epistemic cognition* (pp. 31–50). Routledge.

Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher*, *33*(7), 14–26. https://doi.org/10.3102/0013189X033007014

Jung, E. H., Walsh-Childers, K., & Kim, H.-S. (2016). Factors influencing the perceived credibility of diet-nutrition information web sites. *Computers in Human Behavior, 58,* 37–47. https://doi.org/10.1016/j.chb.2015.11.044

Jung, W. S., Chung, M.-Y., & Rhee, E. S. (2018). The effects of attractiveness and source expertise on online health sites. *Health Communication, 33*(8), 962–971. https://doi.org/10.1080/10410236.2017.1323364

Kąkol, M., Nielek, R., & Wierzbicki, A. (2017). Understanding and predicting Web content credibility using the Content Credibility Corpus. *Information Processing & Management, 53*(5), 1043–1061. https://doi.org/10.1016/j.ipm.2017.04.003

Kiili, C., Laurinen, L., & Marttunen, M. (2008). Students evaluating internet sources: From versatile evaluators to uncritical readers. *Journal of Educational Computing Research, 39*(1), 75–95. https://doi.org/10.2190/EC.39.1.e

Krouwer, S., Poels, K., & Paulussen, S. (2020). Moving towards transparency for native advertisements on news websites: A test of more detailed disclosures. *International Journal of Advertising*, *39*(1), 51–73. https://doi.org/10.1080/02650487.2019.1575107

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159. https://doi.org/10.2307/2529310

Lucassen, T., & Schraagen, J. M. (2013). The influence of source cues and topic familiarity on credibility evaluation. *Computers in Human Behavior*, *29*(4), 1387–1392. https://doi.org/10.1016/j.chb.2013.01.036

Luttrell, W. (2010). *Qualitative educational research: Readings in reflexive methodology and transformative practice*. Routledge.

Machackova, H., & Smahel, D. (2018). The perceived importance of credibility cues for the assessment of the trustworthiness of online information by visitors of health-related websites: The role of individual factors. *Telematics and Informatics*, *35*(5), 1534–1541. https://doi.org/10.1016/j.tele.2018.03.021

Mason, L., Ariasi, N., & Boldrin, A. (2011). Epistemic beliefs in action: Spontaneous reflections about knowledge and knowing during online information searching and their influence on learning. *Learning and Instruction*, *21*(1), 137–151. https://doi.org/10.1016/j.learninstruc.2010.01.001

Mason, L., Boldrin, A., & Ariasi, N. (2010). Searching the Web to learn about a controversial topic: Are students epistemically active? *Instructional Science*, *38*(6), 607–633.

Mason, L., Junyent, A. A., & Tornatora, M. C. (2014). Epistemic evaluation and comprehension of web-source information on controversial science-related topics: Effects of a short-term instructional intervention. *Computers & Education*, *76*, 143–157. https://doi.org/10.1016/j.compedu.2014.03.016

Mason, L., Scrimin, S., Tornatora, M. C., Suitner, C., & Moè, A. (2018). Internet source evaluation: The role of implicit associations and psychophysiological self-regulation. *Computers & Education*, *119*, 59–75. https://doi.org/10.1016/j.compedu.2017.12.009

Morgan, D. L. (1988). *Focus groups as qualitative research* (Vol. 16). Sage Publications.

Morse, J. M. (2003). Principles of mixed methods and multimethod research design. In A. Tashakkori, & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 189–208). Sage Publications.

Muis, K. R., Pekrun, R., Sinatra, G. M., Azevedo, R., Trevors, G., Meier, E., & Heddy, B. C. (2015). The curious case of climate change: Testing a theoretical model of epistemic

beliefs, epistemic emotions, and complex learning. *Learning and Instruction*, *39*, 168–183.

Nunkoosing, K. (2005). The problems with interviews. *Qualitative Health Research*, *15*(5), 698–706.

Press, S. J., & Wilson, S. (1978). Choosing between logistic regression and discriminant analysis. *Journal of the American Statistical Association*, *73*(364), 699–705.

Richter, T., & Schmid, S. (2010). Epistemological beliefs and epistemic strategies in self-regulated learning. *Metacognition and Learning*, *5*(1), 47–65.

Robertson-Lang, L., Major, S., & Hemming, H. (2011). An exploration of search patterns and credibility Issues among older adults seeking online health information. *Canadian Journal on Aging / La Revue Canadienne Du Vieillissement*, *30*(4), 631–645. https://doi.org/10.1017/S071498081100050X

Rouet, J.-F., & Britt, M. A. (2011). In M. T. McCrudden, J. P. Magliano, & G. Schraw (Eds.), *Text relevance and learning from text* (p. 19–52). IAP Information Age Publishing.

Sandoval, W. A. (2014). Science education's need for a theory of epistemological development. *Science Education*, *98*(3), 383–387.

Sandoval, W. A. (2017). Situating practices of epistemic cognition. In T. G. Amin & O. Levrini (Eds.), *Converging Perspectives on Conceptual Change* (1st ed., pp. 253–260). Routledge. https://doi.org/10.4324/9781315467139-30

Sandoval, W. A., & Redman, E. H. (2015). The contextual nature of scientists' views of theories, experimentation, and their coordination. *Science & Education*, *24*(9–10), 1079–1102.

Sinatra, G. M., & Chinn, C. A. (2012). Thinking and reasoning in science: Promoting epistemic conceptual change. In *APA educational psychology handbook: Application to learning and teaching* (Vol. 3, pp. 257–282). American Psychological Association.

Subramaniam, M., Taylor, N. G., St. Jean, B., Follman, R., Kodama, C., & Casciotti, D. (2015). As simple as that?: Tween credibility assessment in a complex online world. *Journal of Documentation*, *71*(3), 550–571. https://doi.org/10.1108/JD-03-2014-0049

Thon, F. M., & Jucks, R. (2017). Believing in expertise: How authors' credentials and language use influence the credibility of online health information. *Health Communication*, *32*(7), 828–836. https://doi.org/10.1080/10410236.2016.1172296

Ulyshen, T. Z., Koehler, M. J., & Gao, F. (2015). Understanding the connection between epistemic beliefs and Internet searching. *Journal of Educational Computing Research*, *53*(3), 345–383.

White, A. M., DeCuir-Gunby, J. T., & Kim, S. (2019). A mixed methods exploration of the relationships between the racial identity, science identity, science self-efficacy, and science achievement of African American students at HBCUs. *Contemporary Educational Psychology*, *57*, 54–71. https://doi.org/10.1016/j.cedpsych.2018.11.006

Wiley, J., Goldman, S. R., Graesser, A. C., Sanchez, C. A., Ash, I. K., & Hemmerich, J. A. (2009). Source evaluation, comprehension, and learning in Internet science inquiry tasks. *American Educational Research Journal*, *46*(4), 1060–1106. https://doi.org/10.3102/0002831209333183

Wineburg, S. S. (1991). On the reading of historical texts: Notes on the breach between school and academy. *American Educational Research Journal*, *28*(3), 495–519.

Zhang, S., & Duke, N. K. (2011). The impact of instruction in the WWWDOT framework on students' disposition and ability to evaluate web sites as sources of information. *The Elementary School Journal, 112*(1), 132–154. https://doi.org/10.1086/660687

Appendices

Appendix A

Focus Group Interview Protocol

Epistemic ideals

1. Can you tell me about a time that you had to evaluate information?
    a. What criteria did you use to evaluate whether that information was reliable?
    b. Where did you learn about that criteria?
    c. How have your instructors discussed source evaluations?
2. How do you know if a piece of information is reliable?
    a. What features does it have? For example, do you look at who the author is or when the information was last updated?
    b. What does an unreliable website have?
3. What makes something true online?

Reliable epistemic processes

4. When do you have to evaluate information?
5. How do you identify false or misleading information?
6. Do you think people use the same criteria to evaluate information consistently?
7. Is it possible for a piece of information to be completely objective or unbiased?
    a. Are there degrees of bias?
    b. How are biases shown in information online? What can we look for?
    c. How do our own biases affect how we interpret information?
    d. If all people are biased by nature, can they do anything to reduce or minimize the influence of their biases?
8. What would you do if you found two websites that made opposite claims? For example, imagine you were looking up information on eye-witness testimony. Site A says that eye-witness testimony is inaccurate because recalling memories is like putting puzzle pieces together and Site B says that eye-witness testimony is accuracy because our memories play back like a video. What would you do?
9. If two websites make opposite claims, can only one site be correct or could both be correct?
    a. **Only one website can be correct:** Why? How can you tell which one is correct?
    b. **Both websites can be correct:** Why? Could one website be more correct than the other or are they equally correct?
    c. **One website cannot be more correct:** Why?
    d. **Equally correct:** Why? How can you tell which is more correct?

Appendix B

Integrated Interview Coding Scheme with Illustrative Examples

Macro-coding category & definition	**Micro-codes**	**Illustrative examples (Participant pseudonym)**
	What is the process being described and how is it defined?	
Defining term: describing the meaning or scope of a specific term. *Meaning* refers to the definition of the term. *Scope* refers to the boundaries of the term.	1. Bias 2. CRAAP Test 3. Clickbait 4. Digital literacy 5. Media literacy 6. Peer reviewed 7. Reliable 8. Truth 9. Values	1. Your opinion. It's your opinion, that's it. (Will) 2. So we have to look at currency… what's the R? (Dolores) 3. Yeah, just obnoxious text that just… it relies on shock value to get you to try and consume whatever it is they're trying to get you to consume. So, like, you see it on YouTube all the time, these titles with lots of emojis and all caps and exclamation marks like, 'My uber driver almost kidnapped me!" or something, and it makes you go 'what?" and you kind of want to click on it, but at the same time as soon as you become aware of the fact that it's probably like the uber driver took a wrong turn at some point and this person overreacted. And its, yeah… (Vera) 4. well it can be like video information, pictures, images, graphs, all sorts of stuff, I guess, any type of information that you can access online. (Amanda) 5. I've heard more the term media literacy, so like just different forms of media not just like technology wise, so like in books too. But like knowing what you're reading, knowing what you're seeing, like pictures, videos now a days. (Betty) 6. I think it means that someone higher up in the educational food chain has read it and said this is true. (Jose) 7. Well reliable is kind of like they like they've proven themselves. Like they haven't lied in anything else, or if they have it's been like minor discrepancies. (Sharon) 8. Yeah, but it's pretty self-explanatory isn't it, truth is the reality of something. (Dolores) 9. So, I think your values are most importantly how you think about the world. (Betty)
Identifying evaluation process: Providing step-by-step procedure for evaluating information or elaborating on specific step(s) of the evaluation process.	1. Assess bias(es) 2. Assess currency 3. Assess the source	1. And then by detecting that, check what… check comparatively what biases there are. I was wrong when you said you can't completely check bias. Because you can check… we know that there are certain biases that exist that are common, and we can be aware of that and we can see it but its only by comparing it to other sources that we know don't have that bias, that we can tell what's a bias and what isn't. (Dolores) 2. Usually I'll search it up, and I'll see what pops up. I'll click the first one cause that's what most of us does I think, and I'll start reading and then if the website seems like a bit, I don't know why but outdated, I don't know why but I won't really use the information. If it looks more modern and you can find who wrote it even if they have a little word about the author that's nice too. (Amanda) 3. But it turns out the source was completely unreliable, and that's exactly what they want to say. And they're like "oh it's okay it's a 2011 research study, don't worry about it, this guy

			said it". But like, who is this guy? You search up this guy, he's no one. And then it just goes on. (Charles)
	4.	Boost disciplinary knowledge	4. Trying to find as many unfamiliar terms in whatever research you find, like try and find the definitions for the terms you're unfamiliar with, I guess? (Michelle)
	5.	Compare multiple sources	5. Um, but generally if I'm actually researching something and I want to make sure that all of my information is viable I just, if I think a site looks suspicious I just like, I keep the information in mind but I go look at other sources and see if their stories match up. That's basically the extent of what I do. (Dolores)
	6.	Compare multiple perspectives	6. I try to, I always try to think of it from like, the opposite point of view. Especially when it comes to arguments, or like solving arguments between people. Like, okay you might think you're right, but have you considered it from this point of view so why the other person you're arguing with might obviously also think they're right. (Sharon)
	7.	Evaluate evidence	7. Well like in terms of that I think if it's such a divided opinion maybe look at how they came to the conclusions and sort of then decide which method is more reliable. (Michelle)
<u>Stating reliability criteria:</u> Describing the meaning or scope of a specific reliability criteria. *Meaning* refers to the definition of the criteria. *Scope* refers to the boundaries of the criteria.	1.	Author expertise	1. Well it can have like the… like words about the author like "Oh he studied this for this, and oh he's got a bachelor's degree in this" and then you're kind of just like "oh, okay he knows what he's talking about." (Amanda)
	2.	Author transparency	2. Like the whole Wikipedia thing you know. Anyone could've just written it and you wouldn't know cause there's no like, list of authors. (Amanda)
	3.	Corroboration	3. Yeah, you could but in like really good research you're at least going to have one other person to support and say like "oh I found this as well." (Sharon)
	4.	Currency	4. So we have to look at the currency. (Dolores)
	5.	Design	5. Another thing is like, if you read through the article, this is a really particular thing that bothers me. But they cut the text up like, between pictures and quotes and ads. And then there's also the fact that even then, their using up a lot of space to make it seem longer and more professional but they're really saying something really simple and they're not really communicating anything, they're just word vomiting and what they're saying doesn't make, necessarily make coherent sense. It's just there, this is information and it may be a little biased and it's just not professional. (Dolores)
	6.	Evidence quality	6. Well like in terms of that I think if it's such a divided opinion maybe look at how they came to the conclusions and sort of then decide which method is more reliable. (Michelle)
	7.	Evidence quantity	7. Also depending how many sources you have, cause like if you have a massive long essay with like multiple examples but all those examples come from one person you can like… like you question it, like why is it only using information from this person. Whereas if you have a massive essay with multiple examples from multiple different people all saying the same thing, you'll… well I don't know about you guys, but I'll have a tendency to believe it more. (Sharon)
	8.	Funding	8. I mean I think it plays a really big role. Like sometimes there's like when you look on websites, you'll see like popup ads and um especially on my, like when you get into like really untrustworthy ones, they have a lot of them and they have flashy and really catching

		titles. Like even I'm like "I so don't want to see that" and that's sort of how they make money, by you clicking on it. So, when I see like a lot of ads I sort of, not trust it. (Josephine)
	9. Objectivity	9. Academic articles, no colour, black and white. It's just like facts, this is why they're so trustworthy right, there's no sugar coating on it. (Cameron)
	10. Peer review	10. Exactly. You have to… A book or an academic article has to go through a process before being put out into the world but um… someone writing for Buzzfeed or someone making a video in their basement ranting about something they don't like, that's not going through anyone else it's just them and their information and what they want to say and then it's out there and anyone can see it. And if you don't think about that in context you can easily think "well these two pieces of information, I found them in the same place so even though one is a book and one is a, not a journal article but just a random Internet article, then they're about the same thing, I found them in the same spot, then they're probably about the same value". But you have to think about the process that one of them had to go through. Like a book had to be written, and then edited by the person, and the probably edited by an actual editor, and then had to be approved by a publisher and then… (Vera)
	11. Popularity	11. And if more people view it, I am more inclined to believe it. (Jose)
	12. Purpose	12. The person has nothing to gain usually is trustworthy. (Josephine)
	13. Retrieval location	13. Uh well everything you get off like um say JSTOR or the library. (Jose)
	14. Tone	14. Ooh. I read more about it, try to see other people's perspectives about it or maybe it's the way they worded it that made it seem fishy. (Amanda)
	15. Truthfulness	15. Um, well if they are honest about where they get their information from so they're gonna cite where they got the information. (Betty)
	16. Type	16. Let's say if I read um a blog online, and then I read a book right after, I'd more likely believe the book because, from what I've been taught, it's something that's more valid than something just written by like I remember I don't know if I, one of my old teachers say " you never know who's writing on the Internet, it could be under a pseudonym or anything" and it's a lot safer to trust in a book than something online. (Jose)
	17. Venue	17. Just like, they've proven themselves to not be faulty and they've been giving accurate information in the past. (Sharon)
	18. Writing quality	18. sometimes the quality of the writing you can find, if there's a lot of spelling mistakes or something or improper punctuation, I look for that sometimes and you can kind of tell that it wasn't written properly. (Jose)
<u>Identifying skills:</u> Describing the meaning or scope of a specific evaluation skill. *Meaning* refers to the definition of the skill.	1. Critical thinking	1. Personally, its if you're able to look at certain posts and not just believe everything that's popping up in front of you. (Sharon)
	2. Faster processing speed	2. Everything's out there, you don't know what to believe what not to believe. It's easier with the internet when you're digitally literate to process information faster, quicker, same thing. (Charles)

Scope refers to the boundaries of the skill.	3.	Open mindedness	3. It's just keeping an open mind, not necessarily always doing your research but being aware of what's happening. (Sharon)
	4.	Manipulating information	4. Oh god, I know for like a fact that like editing and like copying and pasting things into videos that weren't there beforehand is a massive thing now. Well not just that but in photos too. (Sharon)
	5.	Knowing the language	5. Well online we text differ-… we write differently I think so we try to like "love" or things like that. (Jennie)
<u>Providing an example</u>: Describing an example context, criteria or argument that has either occurred or may occur.	1.	Media example	1. Yeah, just obnoxious text that just… it relies on shock value to get you to try and consume whatever it is they're trying to get you to consume. So, like, you see it on YouTube all the time, these titles with lots of emojis and all caps and exclamation marks like, 'My uber driver almost kidnapped me!" or something, and it makes you go 'what?" and you kind of want to click on it, but at the same time as soon as you become aware of the fact that it's probably like the uber driver took a wrong turn at some point and this person overreacted. (Vera)
	2.	Personal example	2. Evidence. Um, for, recording is key for evidence. In my car I have a dashcam. I got hit by car and the guy told me "no you hit me", I told him no I have a camera on and he pissed his pants. So, eyewitness or a videotape is completely um… (Will)
	3.	Mixed example	3. Well… like if I were to talk, I'm very fact based so I read stuff and I'm like "oh this makes so much more sense" whereas people are like "I'm just here". So, whenever I have an argument with someone that doesn't really read much it's like here are the facts, explain me how this isn't happening, and well like. I guess they just don't want to choose to believe it, it's happening, I guess. (Amanda)
		How does the process develop?	
<u>Describing previous experience with process</u>: Describing previous experience learning about evaluating information or evaluating information.	1.	Course assignment	1. My English teacher actually. We had to integrate a secondary source for our paper. Um, and she actually took it a step further of not just taking a peer-reviewed source but actually evaluating who wrote that source and what their other research looks like to see if it would fall in line with what you're looking at. (Michelle)
	2.	Course discussion	2. Well since I've come to cégep I know like they've really taught me like they hammered in what what sources are reliable and what isn't. (Jose)
Important to distinguish evaluations in an academic context and a non-academic context.	3.	Personal	3. my dad especially had it drilled in my brain that you have to question where things are coming from and what is motivating people, and that is well, he's not in school, he's a teacher but he's not in school so he doesn't have to do research on things very often, but it's still a skill that he has and a skill that he instilled in me and my brother and it was this thing of "okay, think about where things are coming from", and he's especially, like he hates advertising and we would listen like on long car rides we would listen to this CBC program called 'under the influence' which is all about advertising and how they trick you and what they do and their techniques and all that. (Vera)
	4.	Other	4. I think I've heard the term before, I can't really remember what. (Michelle)

Establishing need for evaluation: Identifying a context in which information requires (or does not require) evaluation	Evaluation need		So that's what like sparked my… my little uh, I guess my learning process. Like they had the image and the type of cloud, so I was like "Oh, that's that one, and this is that one". I didn't really think about redoing the search to make sure it was credible, I guess. (Amanda)
		How, when, and why does the process change? What are the consequences of the process?	
Identifying external factors: Describing tools or external factors that influence the evaluation process. These factors are not assumed in the evaluation procedure.	1. Accessibility	1.	'Cuz it's so much easier to just go look on the Internet because it's accessible everywhere because everyone has a phone. It requires more effort to go and search out for a book. (Jose)
	2. Algorithm	2.	It's like what is the most linked to. So, like, let's say there's a recipe you're looking for. So, someone who wrote a recipe, and then that recipe gets cited a lot, google is like hey people like this people are like trying to get to this. (Betty)
	3. Comments section	3.	Yeah, well say like you read a newspaper or online because that's how it's done, that is, and then you see some of the comments saying "oh that's not actually what happened, this, this, this, and this happened." (Jose)
	4. Cookies	4.	Not just that but they look at your past searches as well, so like if you have a tendency of I don't know, always looking up things in the colour red, if you type in the word 'coat' like 10 to 1 you're probably going to get a red coat as one of your top hits. (Sharon)
	5. Fact checking sites	5.	Personally, I like fact checker sites, so um there's… I had to do a research project last year on the Syrian refugee crisis, and that's obviously a very inflammatory issue, so it's hard to find information that is unbiased and presenting things in a very like clean, scientific way so I can actually get good information. So, what I found was very helpful was I'd find a source and read through it, figure out if it seemed somewhat legit and then I would go to a fact checker site and I would see how they rated it and why. And it's obviously like fact checking a fact checking site it just, there comes a point where you have to be like okay, I've done my due diligence this is as trustworthy as something I can find can be and then I would use it as a source. (Vera)
	6. Google Scholar	6.	Uh, well, there is a function on google called google scholar which will um whittle down the amount of articles or pages that you can read. And uh basically just focuses on scholarly journals or magazines and stuff. And you can filter it out to check which dates you're looking for, what languages, um yeah. (Michelle)
	7. Individual differences	7.	So long as, it's about bias, cause so long as we keep bias, cause like it has to do with where bias comes from cause we all have our own personal biases through our own personal experiences. But so long as we keep that bias in mind, I don't think we're in danger of falling prey to it [our own biases]. (Dolores)
	8. Keywords	8.	To find exactly what you're looking for without any of the stuff you don't want. To filter everything out without actually having to filter through the 6 million google searches or articles. (Charles)
	9. Linked social media platforms	9.	I know I went, like I know a lot of social media platforms are linked and everything. (Sharon)

10. Purpose of search	10. I think I tend to, like cause that was for a school project, so like I tend to put a lot more effort into things where I know I'm going to have to do like plagiarism checks, and I know it has to be good quality overall because it's a research paper than if I'm trying to figure out information just for my own purposes. I will still go through a process to make sure that the information I'm getting is good, but it probably won't be as rigorous. (Vera)

Appendix C

News and Research Article Comparison Prompt

Distinguishing Between Credible and Non-Credible Sources

Assignment: Find an article in a non-credible publication such as The National Enquirer. Then research the topic of the article in an attempt to support or refute the claims made in the article. Search online databases, the library catalog, and/or the Internet. Document your research – both what you were able to find and what you were unable to find. Indicate what evidence you uncovered, and from what source. Finally, give an overall assessment of the article's credibility.

Purpose: Learn to critically analyze claims and to question the credibility of sources. If searching through multiple research tools (e.g., catalog vs. databases), understand differences in content in and search strategies for these different tools.

Length: ~1250 words

Assignments MUST be handed in during class time on their respective due date. Late assignments will be penalized up to 10% per day. All assignments will be handed back to the student before the end of the semester. All referencing needs to be in APA format.

Appendix D

Source Comparison Coding Scheme with Illustrative Examples

Micro-level codes	Description	Example (Participant pseudonym)
	Macro-level category: Relevance	
Uniqueness	An evaluative decision based on participant's consideration of the information included or excluded.	The "popular magazine", although it had a lot of congruent information and statistics found in the original experiment, the article also brought up other research not mentioned in the original study as well as referenced other works. (Sharon) For the most part, the two articles gave the same information without any grave inaccuracies. (Michelle)
	Macro-level category: Authority	
Venue	An evaluative decision based on participant's consideration of the body that published the information (e.g., publisher's reputation).	Now they do not specify whether they did online or offline research, but the collecting of this information was done by a group selected by the Canadian government, so I believe that the information presented is credible. (Jose)

Macro-level category: Accuracy		
Evidence	An evaluative decision based on participant's consideration of the sources of information provided as evidence. Participant may mention any of the following: 1. quality of evidence 2. type of evidence 3. sufficiency of evidence	1. Regarding the survey/quizzes, a strength is that it is an efficient, low-cost way of gathering knowledge from the participants, with the downside being that the accuracy of the memory of the participants may be inexact. (Sharon) 2. The information in the academic paper was also supported by graphs, tables and images as proof of their claim and to better the comprehension of the precise experiment. While there was an image in the media article no information could be gathered from it and was mostly just there for esthetic purposes. (Cameron) 3. Instead they [media article] use generalizations to convince their audience. Furthermore, the original research article has numerous graphs and figures which detailed the results of the study. However, the media only had an anecdotal image that did not illustrate the actually experiment performed. (Michelle)
Tone	An evaluative decision based on participant's consideration of the author's tone. Participant may mention any of the following, author's: 1. Diction 2. level of formality	1. In the original study, the language used is very developed and is not defined as readers are expected to know what is being said. Contrary to the scientific research article, the magazine article's vocabulary is much more simplistic and any larger words are defined so that the readers fully understand. (Sharon) 2. The academic paper was more complex, used more scientific language and was more precise overall. (Cameron)

Macro-level category: Purpose		
Argument	An evaluative decision based on participant's consideration of author or organization's argument structure (e.g., one sided or two sided argument).	This article briefly mentions that "A research team parsed 68 current cannabis studies last year on behalf of the Canadian government", but in this article he only talks about one, which seems kind of based against the legalization of cannabis, this article only mentions the cons of the drug but none of the benefits, he doesn't link these benefits nor does he talk about them inside his media article. (Jose)
Bias	An evaluative decision based on participant's consideration of author or organization's bias(es). Participant may mention any of the following: 1. source selection 2. labelling (e.g., expert, spokesperson, official)	1. Greene forgot to include the "Inconsistent evidence" portion of the text, thus twisting it." (Jose) 2. Another difference between the articles were their tone. The media article had a much more exciting tone, using sensational language such as "key role" or mentioning that one of the researchers shared a Nobel Prize. (Michelle)
Objectivity	An evaluative decision based on participant's consideration of source's objectivity. Participant may mention that the writing is either subjective and opinion-based or fact-driven and objective.	The original research had a more objective tone to it. Instead of using persuasive language, the researchers provided the data and explained their reasoning. (Michelle)
Purpose	An evaluative decision based on participant's consideration of source's purpose. Participant may identify author's/organization's purpose for creating content as any of the following*: • Persuade • Provoke • Inform or document • Entertain • Sell	Maybe he even wanted to try and divert voters from voting liberal in the election, but this is just crazy speculation so it most likely is not true. (Jose) *Participants did not identify the other types of purposes in their comparisons.

Macro-level category: Other		
Length	An evaluative decision based on participant's consideration of the content length.	The media article and the academic original article are quite different in how they were written. First off the academic article is 25 pages long while the media article is a couple of pages in length and barely takes five minutes to read. (Cameron)
Target audience	An evaluative decision based on participant's consideration of the audience the content was created for.	It is clear that with the level of language and scientific information that is used in the article that the regular readers of neuroscience news are most likely college and university students of that field or people who already graduated in neuroscience or a similar field. (Cameron) Therefore, the primary target for *Psychology Today* would be educated, early to middle aged women. (Michelle)
Writing quality	An evaluative decision based on participant's consideration of the contents' readability or syntax.	To me, this article is actually written quite well, it has correct grammar and presents it's ideas in an organized and top-down manner, and it represents the original article pretty well and doesn't twist anything its saying except for one thing that I found. (Jose) While it is certainly simpler and quicker to read, the media article maintains the main points of the study. (Cameron)

Appendix E

Justification Coding Scheme with Illustrative Examples

Micro-level category	Description	Example (Participant pseudonym)
	Macro-level category: Currency	
Date	An evaluative decision about a rank based on participant's consideration of various publication dates.	They also provided the original publication date. (Lillie) Furthermore, it is a more recent account of the event, which does not always guarantee reliability, but it may present a more accurate depiction of the current event. (Grace)
	Macro-level category: Relevance	
Alignment	An evaluative decision about a rank based on participant's consideration of webpage's alignment with their task definition. Participant may mention any of the following: 1. topic relevance (e.g., central or peripheral to task) 2. appropriateness of evidence for task (e.g., type of evidence)	1. The article has details on the topic but I feel it's too long and not focused enough on the topic. (Bessie) 2. It may be less effective in 'understanding concerns about medically-assisted death' because it lacks a breadth of perspectives. (Owen)
Type	An evaluative decision about a rank based on participant's consideration of the type or location of the article in source.	In the health section (Tyler) In addition to this, this article is also found within the health section, which made me think that the authors are more knowledgeable about their writing. (Kyle)
Uniqueness	An evaluative decision about a rank based on participant's consideration of how unique the information presented on the webpage is in relation to other sources supplied in activity.	It read as though this article was a summary of Article 1 [more reliable article]. (Nicole)

	Macro-level category: Authority	
Author expertise	An evaluative decision about a rank based on participant's consideration of author or organization's expertise related to the topic.	This article is written by one author who is a general reporter and doesn't have the same standards/qualifications (Katherine) The authors' title is health reporter, thus specifying they might have more knowledge about this subject (Hannah)
Venue	An evaluative decision based on participant's consideration of the body that published the information (e.g., publisher's reputation.	The venue of publication seems like a reliable source (Myrtle) Article 1 [the more reliable article] is a more well-known source. (Sam)
	Macro-level category: Accuracy	
Evidence	An evaluative decision about a rank based on participant's consideration of the sources of information provided as evidence. Participant may mention any of the following: 1. type of evidence 2. source of evidence 3. quality of evidence	1. Seems to have more scientific research to back up its claim. (Christine) 2. It lacks a reference page, so it doesn't credit the information it's using. (Elijah) 3. Less credible speakers with quotes. (Alexander)
Corroboration	An evaluative decision about a rank based on participant's consideration of whether information can be corroborated or not in other supplied sources.	Objectively verifiable statements. (Samuel)
Peer-review	An evaluative decision about a rank based on participant's consideration of the review process of information presented.	Furthermore, the text was seen by an editor if I look at the final note. (Kenneth) This article included corrections and an updated, thus it has been revised. (Hannah)
Tone	An evaluative decision based on participant's consideration of the author's tone. Participant may mention any of the following, author's: 1. diction 2. level of formality	1. Uses colourful adjectives which seems to attempt to evoke emotional responses from readers. (Samuel) 2. They used 'won't' in the title, which is not very formal. (Lillie)

Macro-level category: Purpose		
Argument	An evaluative decision based on participant's consideration of author or organization's argument structure (e.g., one sided or two sided argument).	It quotes people from both sides of the question. (Lucille) It also does not give multiple points of view on the situation. (Mia)
Bias	An evaluative decision based on participant's consideration of author or organization's bias(es).	There are several authors, which could reduce bias or introduce alternative perspectives. (Owen) Lacks public opinion → bias. (Alma)
Objectivity	An evaluative decision based on participant's consideration of source's objectivity. Participant may mention any of the following: 1. subjective and opinion-based writing 2. fact-driven and objective writing with facts, statistics, or research	1. Based on the title alone, the one from the Montreal Gazette makes it seem like they have a strong one-sided opinion and it seems like they really wrote the article for clicks. (Kayla) 2. Was more informative and neutral compared to the first article. (Abigail)
Purpose	An evaluative decision based on participant's consideration of source's purpose. Participant may identify author's/organization's purpose for creating content as any of the following*: • Persuade • Provoke • Inform or document • Entertain • Sell	It seems more likely they wrote the article for clicks. (Sam) It seems more like a piece of propaganda. Probably made to sway political ideas. (Julian) * Participants did not identify the other types of purposes in their comparisons.

Macro-level category: Other		
Interest	An evaluative decision about a rank based on learner's consideration of their interest in the content information provided in webpage and/or interest in reading the article itself.	It makes the issue more real and/or relatable (Alyssa) Just feels more authentic. (Wendy)
Length	An evaluative decision about a rank based on learner's consideration of the length of the webpages.	This article seemed to be short and not as detailed as the first one. (Abigail) The article was fairly short. (Lucas)
No justification	If a student wrote something in the space provided that was not indicative of any of the coding categories.	The other article is the one I chose because it was a better option than this, despite not being a 100% sure of my decision. (Sarah) In this article, we can see the concerns that medically-assisted death could bring in society. (Randy)
Writing quality	An evaluative decision about a rank based on learner's consideration of the readability of the webpage based on its syntax or spelling.	The wording in this article is easier to understand and follow along. (Ada) Typo within the second paragraph. (Taylor)

Bridging Text

Chapter 3 examined college students' epistemic thinking during authentic source evaluations. In Study 1, college students' epistemic metacognitive knowledge about source evaluations were collected in focus group interviews and their epistemic ideals were inferred from a comparison of two online sources. In Study 2, a larger group of college students' epistemic ideals were inferred from a source comparison task. Results revealed that college students examine a variety of features to determine a webpage's reliability, including its currency, relevance, authority, accuracy and purpose. Surprisingly, more students who relied on relevance criteria to justify their source evaluations accurately rated the news articles. These results were used to develop a source evaluation training based on the CRAAP test, a widely endorsed source evaluation tool.

The following chapter presents the third phase of a multiphase mixed methods research program aimed at implementing and evaluating an online source evaluation training. Due to the campus closures associated with COVID-19, the sample changed from students at the Quebec college to students at McGill University and adults from across Canada. In phase 3, participants were randomly assigned to one of two groups: the training group who watched a 35-min online video modeling CRAAP test use, or the control group who received no training. All participants completed an inquiry activity where they examined six online webpages, rank-ordered the webpages based on reliability, and justified their rank-ordering decisions. Participants also produced an essay using the sources provided. This study contributes an empirical evaluation of a widely cited source evaluation tool, the CRAAP test, and to better understanding adults' accuracy and reasoning during online source evaluations.

Chapter 4

Manuscript 2

Identifying CRAAP on the internet: Epistemic cognition during source evaluations

Denton, C. A., Muis, K. R., Armstrong, S., & Dubé, A. (in preparation). Identifying CRAAP on the internet: Epistemic cognition during source evaluations. *Contemporary Educational Psychology*.

Abstract

Individuals of all ages struggle to determine the reliability of information on the internet. To address this common issue, many educational institutions have endorsed the CRAAP test as an effective approach to support identification of unreliable information. The present study aimed to evaluate the efficacy of a source evaluation training based on the CRAAP test. Seventy-seven (*n* = 77) adults across Canada were recruited to evaluate six authentic webpages and then construct an argument on the topic. Half the sample received training to examine the currency, relevance, authority, accuracy, and purpose (i.e., CRAAP) of the webpages before completing the online activity. Results revealed that the training group provided a slightly more accurate rank-ordering of the webpages using currency and authority criteria, and the control group produced a slightly more complex essay on the topic. Despite the clear theoretical relationship between each CRAAP criteria and source evaluations, the findings suggest that the CRAAP test is not effective as a standalone evaluation tool. Suggestions for improving the tool are recommended.

Keywords: epistemic cognition, source evaluation, intervention, internet

The internet has become a valuable resource for quickly accessing information. With this access comes great responsibility to examine the quality of information before sharing it with others. People of all ages struggle to accurately evaluate sources (Gerjets et al., 2011; Mason et al., 2011; Stanford History Education Group, 2016; Subramaniam et al., 2015) and subsequently accept information from biased sources (Bråten et al., 2009; Halverson et al., 2010). Educators, researchers and non-profit organizations have developed instructional interventions, trainings and online activities to improve students' source evaluation skills (e.g., McGrew, 2020; The News Literacy Project, 2020). Although these resources were developed for grade school and university students, some are available to the public at no cost (The News Literacy Project, 2020; Stanford History Education Group, 2020). Given the importance of effective source evaluation skills, research is needed to understand how adults that may not be affiliated with an educational institution can be trained to better evaluate information on the internet.

The purpose of the present study was to examine the efficacy of a source evaluation training implemented in an online context with adults across Canada. The training aimed to improve their source evaluation accuracy and reasoning during an online inquiry task about mental illness. Without instruction, previous research has found that people rely on content- or design-focused criteria to justify their evaluation decisions, including comprehensibility and task relevance criteria (Braasch et al., 2013; Kąkol et al., 2017; Kiili et al., 2008). Based on previous work (Denton et al., in preparation), we found that college students applied a variety of content-focused and epistemic criteria to determine reliability. To account for adults' use of various criteria, we adopted the Apt-AIR framework (Barzilai & Chinn, 2018) and CRAAP test (Blakeslee, 2004; Meriam Library of CSU Chico, 2010) to develop the training. The CRAAP (currency, relevance, authority, accuracy, purpose) test has been included on numerous library

websites, with some informal attempts to extend the tool beyond its original checklist approach (Fielding, 2019). Although many interventions have incorporated aspects of the CRAAP test (e.g., establishing source's authority, Barzilai, Mor-Hagani, et al., 2020; Mason et al., 2014; Pretorius, 2018), the CRAAP test has not widely been empirically explored, especially through the lens of the Apt-AIR framework. The effectiveness of the training should be revealed by adults' justifications of their source evaluations and their source integration after the inquiry task.

Adapting the CRAAP Test

The CRAAP test is a popular framework for evaluating sources that can be found on library websites throughout North America (Batchelor, 2017; Smallwood, 2015). Commonly including about 25 questions or criteria, the CRAAP test can be used to assess the Currency, Relevance, Authority, Accuracy, and Purpose of a source (Blakeslee, 2010). The test has been adapted to include other aspects of online source evaluations, such as the presence of advertisements (Green, 2019) and the quality of linked webpages (Queen's University Library, 2020). Taken together, the CRAAP test reflects content-focused and epistemic criteria to assess a source's knowledge claims, author, and argument through vertical and lateral reading. See Table 1 for a description of each category with an illustrative example. Critiques of the CRAAP test have focused on the tool's reliance on information within the source being evaluated (i.e., vertical reading) rather than information from multiple sources (i.e., lateral reading, Breakstone et al., 2018). However, the relevance and accuracy criteria specifically support source integration, an intricate epistemic process where multiple perspectives are weighed and reconciled (Barzilai & Zohar, 2014, 2018). Further, the relevance criteria also encourage users to establish reliability based on the other criteria before making relevance-based judgments. To

emphasize the components that describe lateral reading practices, the online CRAAP training included modeling use of each category within a single webpage and across multiple webpages.

In their Apt-AIR framework, Barzilai and Chinn (2018) delineated the role of epistemic cognition in source evaluations. Epistemic cognition refers to thinking about the acquisition, justification, and use of knowledge (Greene et al., 2018). During source evaluations, epistemic cognition is used to examine the epistemic properties of knowledge claims (e.g., accuracy), sources (e.g., authority) and arguments (e.g., purpose, Barzilai and Zohar, 2014), as seen in the CRAAP test. Further, Barzilai and Chinn (2018) highlighted the cognitive components of epistemic ideals and reliable epistemic processes during source evaluations. Epistemic ideals describe the criteria used to justify accepting or rejecting a knowledge claim (Chinn et al., 2011, 2014). A variety of criteria are used to justify reliability assessments with content- or design-focused criteria used most frequently (Bates et al., 2006; Kąkol et al., 2017; Ulyshen et al., 2015). Although documented at lower rates, epistemic ideals have been spontaneously observed during source evaluations (Gerjets et al., 2011; Mason et al., 2011) with mixed findings about the role epistemic ideals play in accurate source evaluations (Mason et al., 2014, 2018). Under reliable epistemic processes, Barzilai and Chinn (2018) specifically described creative processes as procedures to construct an epistemic product such as a written argument that integrates multiple perspectives (Chinn et al., 2011, 2014). Accurate source evaluations have been associated with more complex written arguments (i.e., two-sided argument, Braasch et al., 2013; Mason et al., 2014). Based on the Apt-AIR framework, each category of the CRAAP test plays a role in source evaluation and integrations.

Table 1

Summary of CRAAP Test Criteria

Category	**Criteria Description**	**Illustrative Example**
Currency	Criteria that capture the publishing date, revision history, functionality of links and sensitivity of task to source currency.	A *reliable website* about a controversial socio-scientific topic has been published or revised recently enough to capture the present stances on the topic and successfully links to other source. An *unreliable website* may not present up-to-date stances on the topic or have broken links to other sources that may be used as evidence of claims.
Relevance	Criteria that capture the task relevance, target audience, appropriateness of information. Two questions also address the user's larger evaluation process: variety of sources gathered and comfort citing webpage in research paper.	A *reliable website* provides essential information on the topic that is appropriate to for users needed. This source is chosen by determining that it is the most *reliable* source for the user's needs after examining a variety of sources. An *unreliable website* may provide peripheral information on the topic or provide information that is either too formal or informal for the user's needs. A source is deemed *unreliable* if at least one other source examined provides more reliable information for the user's needs.
Authority	Criteria that capture indicators of qualifications on the topic, including the author and publisher's credentials and affiliations, their contact information and URL.	The author or publisher of a *reliable website* is trained or holds adequate experience on the topic and may be contacted for follow up on the content. The source's URL may indicate the webpage is part of a government (.gov) or educational institution (.edu). The author or publisher of an *unreliable website* may not be transparent, which prevents the user from following up on the content. The domain suffix may indicate that the website is a commercial site (.com) or network site (.net), but unreliable webpages can also be found within institutions (.edu) and non-profit organizations (.org).

Accuracy	Criteria that capture indicators of trustworthiness, such as evidence and support for claims, peer review, tone, and writing quality. One question also addresses the user's larger evaluation process: knowledge-based validation or corroboration.	A *reliable website* presents evidence and support for their claims about the topic. The user can verify information using their prior knowledge or by comparing the source's claims, evidence and arguments to other sources on the topic. Peer review may occur before the content is published. An *unreliable website* may not present evidence or support for their claims on the topic. The user may not be able to verify the information presented using their own knowledge or other sources. A peer review process may not occur before the content is published.
Purpose	Criteria that capture the author or publisher's motivation for publishing content, such as their purpose, objectivity, or biases.	A *reliable website* may provide a clear disclosure about ads on the webpage, present multiple perspectives with evidence and support, and minimize biases in writing and source selection. An *unreliable website* may not be transparent about ads revenue or have numerous unrelated ads, present subjective opinion without evidence or support for another perspective and use biased labelling and source selection to support claims.

Note. Adapted from Blakeslee (2004) and Meriam Library of CSU Chico (2010).

Barzilai and Chinn (2018) argued for a situated view of epistemic cognition where epistemic ideals and processes must be adapted to different domains, contexts, and tasks. To a small degree, the CRAAP test accounts for these differences through the relevance criteria, which provides task and context clues to calibrate epistemic cognition. Following Sosa (2015), Barzilai and Chinn (2018) asserted that apt epistemic cognition could be assessed by accurate judgments of information supported by competent use of epistemic ideals. Thus, an apt source evaluation is accurate and may be based on the appropriate use of content-focused criteria and epistemic ideals. Beyond task and context, researchers have identified that prior knowledge and attitude toward the topic have profound influences on apt source evaluations (Bråten et al., 2016; Maier & Ritcher, 2017): low prior knowledge (E. H. Jung et al., 2016; W. S. Jung et al., 2018) or differing beliefs about the topic (Taylor et al., 2015; Van Strien et al., 2014) often result in inaccurate reliability attributions. Within the CRAAP test, low prior knowledge may impact the user's ability to verify the webpage's accuracy using their knowledge (i.e., knowledge-based validation, Ritcher & Schmid, 2010), and their topic beliefs may influence their use of the accuracy and purpose criteria (e.g., accept belief-consistent information, Meppelink et al., 2019). Given the widespread endorsement of the CRAAP test by libraries, further research is needed to assess the strengths and weaknesses of this approach for online use.

Teaching Source Evaluation Skills

Within learning contexts, targeted interventions have demonstrated that students can be trained to use important epistemic ideals related to authority and accuracy during source evaluations (Britt & Aglinskas, 2002; McGrew et al., 2019; Pretorius, 2018). Like the CRAAP test, some interventions have used declarative knowledge or modeling to train students to assess the source's authoritativeness (e.g., author's expertise or purpose) and content accuracy (e.g.,

plausibility, corroboration, Mason et al., 2014; Wiley et al., 2009; Hämäläinen et al., 2020). For example, Mason and colleagues (2014) investigated the efficacy of SEEK (Source, Evaluation, Explanation, Knowledge) training by asking secondary school students to read sources from an offline digital environment that ranged in authoritativeness and viewpoint, rank the reliability of each website, and provide a justification for their ranking. Students also constructed a written argument in a transfer task. Beyond prior knowledge and argumentative reasoning skills, Mason and colleagues (2014) found that students who received training more accurately rank-ordered the most reliable websites and used SEEK criteria to justify their ranking decisions. Students also demonstrated better source integration in their written essays; specifically, individuals in the training group with higher prior knowledge produced more complex arguments. Whereas many interventions have been conducted with students in academic contexts, less is known about the efficacy of trainings with adults beyond university or graduate school. Researchers have found that adults are better able to assess the quality of information with age and experience (Eshet-Alkalai & Chajut, 2009, 2010); yet adults have also been shown to heavily rely on content- and design- focused criteria to assess reliability (Machackova & Smahel, 2018; Kąkol et al., 2017).

Much of the current understanding of source evaluations has been shaped by studies conducted in offline environments, as seen in Mason and colleagues' work (2011, 2014, 2018). Due to the situated nature of epistemic cognition (Chinn & Sandoval, 2018; Sandoval, 2017), findings in offline environments may not generalize to source evaluations on the internet. Drawing from the Apt-AIR framework, Barzilai, Mor-Hagani, and colleagues (2020) developed an offline learning environment with epistemic scaffolds to support students' source evaluations and integrations. Like Mason and colleagues (2014), the effectiveness of the training was determined through students' ratings of sources, their justifications, and their written essays

without scaffolds. After using the scaffolded environment, more students that received training provided at least one epistemic criterion to justify their source rating. The researchers classified the training's impact on students' source evaluations as a positive, small effect (Barzilai, Mor-Hagani, et al., 2020). The training had a significant impact on students' source interactions with more students that received training producing two-sided arguments and reconciling multiple perspectives. Despite increases in justification based on epistemic ideals, Mason and colleagues (2014) and Barzilai, Mor-Hagani and colleagues (2020) reported that use was low during source evaluations. In contrast, McGrew (2020) observed large improvements in students' justifications after training, documenting a shift from irrelevant justifications to epistemic ideals for nearly half of the participants (46%). Whereas education researchers using offline environments to train students and examine the effectiveness of their training have greatly contributed to understandings of source evaluations, findings from such studies have limited implications for evaluating sources on the internet. To date, few studies measuring epistemic cognition have been conducted in an authentic environment and none have explored adults' source evaluation behaviours.

The Present Study

The purpose of the present research was to examine the efficacy of a source evaluation training aimed at improving adults' accuracy and reasoning during online source evaluations. Adults' across Canada were selected because research has focused on students' limited use of appropriate epistemic ideals during source evaluations (Braasch et al., 2013; Halverson et al., 2010). The following research questions guided the study:

1. Do adults in an intervention group that received source evaluation training evaluate sources more accurately than those in a comparison group?

2. How do adults' epistemic ideals contribute to their overall source evaluations?
3. How do adults' epistemic ideals contribute to the complexity of their written argumentation?

In accordance with previous research (Mason et al., 2014; Wiley et al., 2009), we hypothesize that adults who receive targeted source evaluation training will more accurately rank-order authentic sources from the internet. Further, adults that rely on epistemic ideals will be more accurate than those who do not (Mason et al., 2014). Lastly, we hypothesize that adults who rely on epistemic ideals to justify their source evaluations will construct more complex arguments on the topic (Barzilai & Eshet-Alkalai, 2015; Braasch et al., 2013).

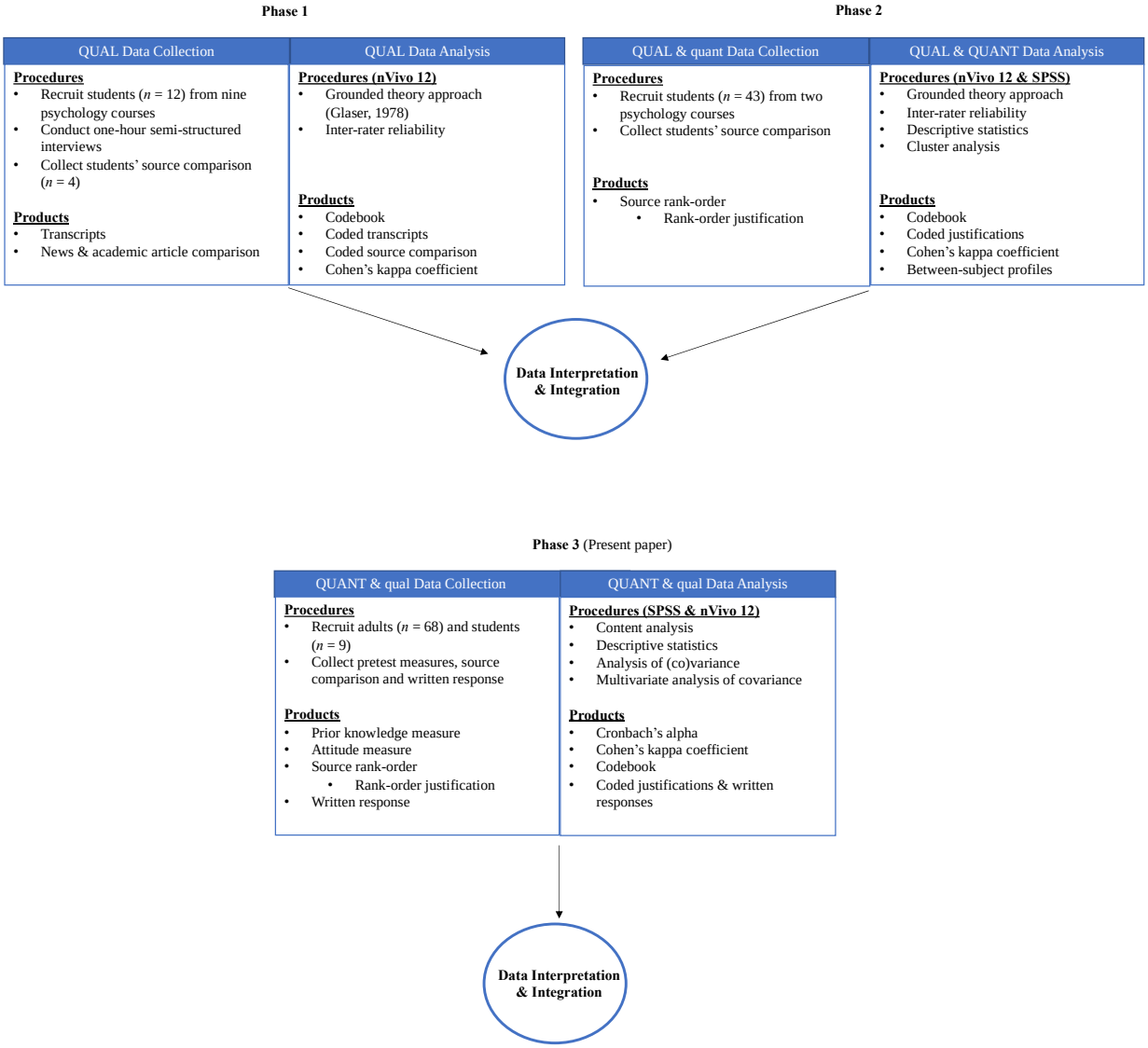
Methods

Study Design

Following Cresswell and Plano Clark's (2017) recommendations, we used a multiphase mixed methods design to develop an online training to improve adults' source evaluations. As depicted in Note. , we collected both qualitative and quantitative data to assess epistemic processes during source evaluations. Whereas Phase 1 and 2 were reported elsewhere (Denton et al., in preparation), this manuscript will focus on Phase 3. To represent how emphasis was placed on each method in Phase 3, we use "QUANT" to indicate the emphasis on quantitative methods, whereas our use of "qual" indicates lesser emphasis on qualitative methods (Cresswell & Plano Clark, 2011; Morse, 2003).

Figure 1

Schematic of Multiphase Mixed Methods Design



Note. Adapted from White and colleagues' (2019) Figure 1.

Participants

We conducted a priori power analysis using G*Power 3 (Faul et al., 2009), which revealed that a minimum of 82 participants would be needed to detect a medium effect size with power $(1-\beta)$ set at .80 and α set at .05. Two-hundred and two adults (n = 202) participated. Adults were recruited through Amazon's Mechanical Turk (n = 161) and through online advertisement at a large university (n = 41). After reviewing participants' responses, 93 participants were excluded from further analyses because they (a) did not complete the survey, (b) failed at least one attention check, or (c) were under 18 years of age. Of the remaining 109 participants, 41 had been randomly assigned to receive training and 68 to the control group. Following Tabachnick and Fidell's (2013) recommendations for unequal groups, we randomly selected 41 control participants with similar educational backgrounds for further analyses.

These participants (n = 82, 32 female, 1 non-binary) represented a wide range of age groups and educational backgrounds. Participants' ages ranged from eighteen to sixty-five years of age (M = 30.7 years, SD = 10.6). They reported completing secondary (31%), college (9%), undergraduate (45%), and master's (15%) level schooling. Participants self-identified as Caucasian (56%), Asian (29%), Black or African Canadian (6%), Latinx (4%), multiracial (3%), or Indigenous (2%). Most participants reported a political affiliation with the Liberal Party (50%), followed by the Conservative Party (18%), New Democratic Party (17%), among others (15%). All participants volunteered to take part in the research. The MTurk group was compensated $5 for their time and the university group was entered in a draw to win $100.

Procedure

The study was completed on the participant's personal device using an online survey platform. After providing electronic consent, participants completed topic knowledge and

attitude measures. Next, the training group received CRAAP test training while participants in the control group directly began the inquiry activity. In the inquiry activity, participants were instructed to take 30 minutes to read and examine each webpage about mental illness in preparation to rank-order the webpages. After examining the webpages, participants rank-ordered the webpages and explained why each webpage was assigned its rank. Next, participants wrote arguments about the relationship between mental illness and violent behaviour. Lastly, participants completed a demographics survey.

Online Training

Participants who were randomly assigned to the training group watched a 35-minute video that described the CRAAP criteria and modelled criteria use while examining sample webpages' reliability. The topic chosen for the sample activity was medically assisted death regulations in Québec, a common topic addressed in introductory psychology courses. The epistemic aim of the sample activity was understanding (Greene et al., 2014, 2018). To ensure treatment fidelity, participants' time to complete the study was recorded and they were asked to answer three questions about the video's content before moving on to the inquiry task. Participants that completed the study in less than 35 minutes or incorrectly answered one or more of the questions were excluded from analyses. Table 2 provides a description of how each category was addressed in the video.

Table 2

CRAAP Training Overview

Webpage #	Description	Criteria Assessed
1	An authentic source was selected from a search engine results page to model how participants could assess the webpage's: a. Timestamp or revision and copyright history b. design features c. link functionality d. currency for sample task	Currency, Relevance
2	A second authentic source was selected from a search engine results page to model how participants could assess the author's credentials and experience by looking for their biography. Since no biography was available for the author, a Google search was used to find their credentials, affiliations, and previous work. The venue's purpose was also discussed by highlighting the advertisements and searching for ownership and funding details.	Authority, Purpose
1, 2, 3	Using Snopes, a fact-checking website, evidence about the sample topic was search for to demonstrate that evidence can be assessed within and across sources. Snopes reports about medically assisted death was examined to demonstrate the websites' use. Returning to websites 1 and 2, the content's references were examined and indicators of a peer review processes were identified.	Accuracy
1 & 2, 4	Using Mediabiasfactcheck, a news media rating website, the publishers of webpages 1 and 2 were searched and their bias, ownership, and accuracy reports were examined. The contents of each webpage was also examined for indicators of bias in the text.	Authority, Accuracy, Purpose

Materials for the Inquiry Activity

Six authentic webpages about the relations between mental illness and violent behavior were selected and displayed in a Google search results format as the "top six results" using the search terms "mental illness + violence" (see Figure 2). The webpages held distinctive perspectives about this relationship and provided evidence from sources with varying levels of reliability (see Table 3 for webpage descriptions). To avoid interference related to webpage order, the search engine results page was randomly ordered. Sources listed on the search engine results page appeared with the (1) webpage title, (2) URL linked to the original webpage, and (3) brief

description of the content with keywords in bold. Participants were instructed to take 30 minutes to read and examine the webpages in any order, and to use external sources to help them assess each webpage's reliability. For the purpose of rank-ordering the articles based on their reliability, two independent raters evaluated each webpage using the CRAAP test. Their agreement, as measured by Cohen's kappa coefficient, was established at .81, with almost perfect agreement (Landis & Koch, 1977). Despite minor disagreements, both raters ranked the webpages in the same order. Each disagreement was discussed and resolved.

Figure 2

Search Engine Results Page for Inquiry Activity

Table 3

Webpage Descriptions for Inquiry Activity

Webpage	Title	**Author Description**	**Publication Type**	**Publication Venue**	Publication Year	Main Claims
1	Violent victimization more prevalent among people with mental health-related disabilities	Marta Burczycka, Analyst	Government website	Statistics Canada	2018	People who suffer from mental health-related disabilities are **more likely to be the victims of violence**.
2	Yes, the untreated seriously mentally ill *are m*ore violent than others	D. J. Jaffe, Executive Director of the Mental Illness Policy Org.	News website	National Review	2017	**Those with untreated mental illness are more prone to engaging in acts of violence**.
3	Media's damaging depictions of mental illness	Margarita Tartakovsky, Associate Editor	Educational website	PsychCentral	2019	The portrayal of mental illness in the **media contributes to the stigma against those that suffer from mental illness**, several myths are addressed.
4	Violence, mental illness and substance use	Canadian Mental Health Association, BC Division	Not-for-profit website	Here To Help	2013	A description of mental illnesses and their **relation to other behaviours including violence and substance use**.
5	People with mental illness twice as likely to be victims of violence: study	Katya Slepian, Reporter	News website	The Columbia Valley Pioneer	2018	**Those who experience mental illness are more likely experience violence** and demonstrating this through data can help reduce the stigma against mental illness.
6	How mental illness is misrepresented in the media	Kristin Fawcett, Contributor	News website	U.S. News & World Report	2015	**Media depictions of mental illness are often misleading** and contribute to the stigma around the subject.

Note. Differences influencing epistemic evaluations are in bold.

Measures of Epistemic Cognition

Overall Rank-ordering of Webpages

Participants were asked to rank-order the webpages from most to least reliable (Mason et al., 2014; Walvaran et al., 2013). Rankings were compared with the expert rank-order and scored based on their proximity to the correct ranking. For example, ranking the most reliable webpage third received a score of 2 because the assigned ranking was two ranks away from the experts' ranking. Scores were summed to serve as an indicator of participants' accuracy in rank-ordering the webpages with a lower score reflecting a more accurate ranking.

Justification for Webpage Rank-ordering

Participants were also asked to justify their rank-ordering of the two most reliable and two least reliable webpages in a written response (Braasch et al., 2013; Mason et al., 2014). Justifications were qualitatively analyzed by two independent raters and participants were given one point for each CRAAP criteria used to justify their rank-ordering. The coding scheme was tested using 14 participant responses (17% of sample), and the raters' agreement was established at .83, with near perfect agreement (Landis & Koch, 1977). Each disagreement was carefully examined and discussed to update the coding scheme. See Table 3 for selected criteria with illustrative examples and Appendix A for the full coding scheme. The first and third author coded the remainder of the justification responses. Participants' criteria use was summed for each macro-level category (i.e., currency, relevance, authority, accuracy and purpose) and a separate sum was calculated for their epistemic ideal use.

Table 3

Selected Justification Criteria with Illustrative Examples

Micro-level category	**Description**	**Illustrative example (Participant number)**
	Macro-level category: Currency	
Date	An evaluative decision about a rank based on participant's consideration of webpage's various publication dates. Participant may mention any of the following: • Original publication date • Revision date • Copyright date	"The time stamp and copyrights are up to date" (PN 227) "The currency of the article at first glance is from 2014, though on further inspection, it was last modified in October 2018 and contains relevant information from 2018." (PN 255)
Design	An evaluative decision about a rank based on participant's consideration of webpage's design or format features. Participant may mention any of the following: • Links to social media or other sharing capabilities • Pop up boxes • Videos/images/advertisements • Webpage "look" is older/newer • Format of webpage or its content	"Webpage opened with advertising all over it, pop ups at the bottom, etc." (PN 293) "The web page style hasn't been updated in a decade." (PN 232)
	Macro-level category: Relevance	
Alignment	An evaluative decision about a rank based on participant's consideration of webpage's alignment with their task definition. Participant may mention any of the following: • topic relevance (e.g., central or peripheral to task) • appropriateness of evidence for task (e.g., type of evidence) • current enough to complete task	"Being a national review, it might not pinpoint exactly what you are looking for on a local level" (PN 213) "It talks quite a bit about substance abuse so it's very relative to the topic" (PN 242)

Type	An evaluative decision about a rank based on participant's consideration of the type or location of the webpage. Participant may mention any of the following: • News article or news section • Journal article • Blog • Encyclopedia	"Columbia Valley Pioneer - A published news article / blog is considered a lesser credible source of information based on the fact that there can be underlying unknown factors behind the write-up." (PN 224) "Although the website has a quite in-depth analysis and has many reliable sources, I ranked it fifth on the list. This is because it is news article. In comparison to the other sites I would place it lower in reliability." (PN 277)
	Macro-level category: Authority	
Author expertise	An evaluative decision about a rank based on participant's consideration of author or organization's expertise related to the topic. Participant may mention any of the following: • author or organization's credentials (e.g., educational degree) • author or organization's affiliations • author or organization's years of experience • author or organization's funding status (e.g., what area are they funded for?)	"It's a government website which draws data from credible scientists who work for StatsCan." (PN 209) "It seems to me an important center and concerned with mental health and with a lot of experience." (PN 220)
Venue	An evaluative decision about a rank based on participant's consideration of the body that published the webpage. Participant may mention any of the following: • publisher's reputation (e.g., factual reporting, sourcing practices, fact checking system) • publisher's funding • publisher's ownership	"Because it was made by Statistics Canada which should be the more reliable since it has been made by the government." (PN 219)
	Macro-level category: Accuracy	
Evidence	An evaluative decision about a rank based on participant's consideration of the sources of information provided as evidence. Participant may mention any of the following: • quality of evidence • type of evidence • source of evidence • currency of evidence • sufficiency of evidence	"Although this article does have some evidence to back up its claims, they are relatively few." (PN 203) "The article's accuracy is questionable since the sources used for the examples used aren't scientific." (PN 255)

Corroboration	An evaluative decision about a rank based on participant's consideration of whether information can be corroborated or not in other sources.	"I don't know if it's reliable or not, less statistics and goes against another one [source]." (PN 292) "I have learnt similar information at school. I also learnt from home and my own experience." (PN 302)
	Macro-category: Purpose	
Argument	An evaluative decision based on participant's consideration of author or organization's argument structure (e.g., one sided or two sided argument).	"Seems to come from a very one sided source. Does not talk about positive part of mental illness victims." (PN 262) "Invalid reasoning and conclusion." (PN 300)
Bias	An evaluative decision based on participant's consideration of author or organization's bias(es). Participant may mention any of the following: 3. source selection 4. labelling (e.g., expert, spokesperson, official)	"The mediabias [rating from mediabiasfactcheck.com] was good IMO." (PN 290) "Their title attracts the wrong idea, even though are talking about how people with mental illness are not actually more violent the title of the article gives the opposite idea." (PN 294)

Source Integration in Essays

After completing the inquiry activity, participants were asked to use information from the webpages and their prior knowledge to construct an essay to address the question: Are individuals with mental illness more violent than individuals without mental illness? Participants were instructed to present and justify their position regarding the relationship between mental illness and violent behaviours. Two independent raters qualitatively analyzed participants' argument structure using 15 participant responses (15% of sample). The first and third authors' agreement, as measured by Cohen's kappa coefficient, was established at .76, with substantial agreement (Landis & Koch, 1977). The raters' discussed their disagreements and updated the coding scheme to code the remaining responses (see Appendix B). Following Chevrier and colleagues (2019), participants' written responses were given a score between 0 and 5, with 0 representing no argument and 5 representing a two-sided argument with source evaluation(s). Illustrative examples of each response level are presented in Table 4.

Table 4

Illustrative Examples of Essay Scores

Score	Illustrative example of essay response
0	To be honest, my background is not on the mental illness. However, I have seen many people having mental illness. They are more likely affected by their environment such as parents, school or work environment and people. And it could be resolved when their problematic environment treated.
1	I believe mentally ill individual can are likely to show aggressive behaviour (to different extent).
2	I believe that people who experience mental illness are not more violent compared to the rest of the population. Rather they people with mental illness are more vulnerable and prone to being a victim of violence. From the article from The Columbia Valley Pioneer, titled " People with mental illness twice as likely to be victims of violence: study", the article talked about how although some people who suffer from a mental illness may be violent, it outweighs the fact that these group of people are far more likely to be victims of violence. The article furthers talking about stats of people with mental illness who report the violence/assaults they experienced to the police, the number of people who report their assaults are only about 22%. I believe with this fact alone we can justify the misconception of mental illness associated with violence, these group of people feel as they have to defend themselves from the rest of the population to survive and have resorted to maybe a violent defensive mechanism. I believe people suffer from mental illness don't feel safe to discuss the violence they suffer as they will not get the right kind of help if they do so. I don't think people with mental illness present more violence than the rest of the population, I believe it is a stigma that people need to start stop believing. This will be one step into combating the misconception.
3	Individual with mental illness tends to be more violent than individuals without mental illness, because mental illness often come hand-in-hand with other issues, like binge drinking and heavy drug use, he added, which by themselves can make people more aggressive. Statistics Canada figures show 15 per cent of people with mental illness used drugs, compared to six per cent of those without. People living with mental health conditions may experience stigma, discrimination and social exclusion that significantly impacts on their lives. On the other side, mental illness individuals tend to be victimized more than other people. According to statistics Canada, individuals with mental health-related disabilities were less likely to report their victimization to the police (22% versus 31%). Among those who did, levels of satisfaction with police action were similar to those reported by victims with no such condition. So the reason why mental ill individual tends to be more violent than individuals without mental illness, is for multiple reason. They tend to have experienced more stuff, half (51%) of people with a mental health-related disability reported having experienced physical or sexual abuse as a child, and almost a quarter (23%) reported having experienced homelessness. Binge drinking (44%) and drug use (15%) were significantly more common among those with mental health-related disabilities. Mental ill people don't become violent just like this, it is a series of event that led to it.

4	Yes, I think mentally ill populations are more violent than other sectors of the population. I personally know family members and friends who show such violent tendencies and who have been diagnosed with some forms of mental illness. I also think that although many sufferers of mental illness become ill from exposures to forms of violence and victimization, and have been passive victims, I feel that their victimization begets violent tendencies for themselves. Its akin to the adage of the sins of the father being passed down and from one person or act onto another. I have family members and friends who due to victimization, have become more violent-prone. However, that's not to say that some sufferers of mental illness aren't docile and passive, but some sufferers do tend to be more aggressive and violent than populations that have not been diagnosed with such an illness. I also think that the severity of their violent tendencies and symptoms vary based on their specific diagnoses. For example, sufferers of PTSD and bipolar disorder may act out more acts of violence (to self and/or to others) than other relatively "normal" forms of mental illness. I feel this way because of statistics and certain news reporting, such as veterans of wars coming back home and kill themselves and/or their families. In general, I think that those who suffer from mental illness may not have such great control of their actions and thinking. Such distortions and lack of self-control therefore begets violent tendencies. I don't think this pertains to all mental illness sufferers, just some. So to sum up, yes individuals with mental illness can be more violent than those without mental illness, but I wouldn't tag this label to all sufferers of mental illness as a general statement.
5	In my opinion based on the articles read and prior understanding, individuals with mental illness are generally NOT more violent than individuals without mental illness. There is evidence that individuals with mental illness, however, experience violence more than individuals without mental illness, but this does not mean that they are violent themselves. The Statistics Canada article reported that 1 in 10 people with mental illness experience violence (i.e., are victimized vs. causing violence). The "Here to Help" article indicated that in a large US study, the risk of violence from a mentally ill person without a co-existing substance abuse problem was about the same as the general population over the next three years. Now in my response I did say "generally" since in some specific types of mental illness such as schizophrenia or psychosis, there is an increased risk of violence. There are complexities in understanding the link between violence and mental illness due to co-existing factors such as substance abuse, lower income, and marginalization. With these co-factors present, violence may be erroneously correlated to mental illness. If only 3% (estimated) of violent crime in Canada is attributed to mental illness, I stand by my opinion that in general individuals with mental illness are not more violent than those without mental illness.

Measures of Individual Differences

To mitigate the profound influence of individual differences on source evaluations (Bråten et al., 2016; Meppelink et al., 2019; Maier & Ritcher, 2017), participants completed measures of their prior knowledge about mental illness and attitudes toward those with mental illness. These measures were used to assess differences between the training and control groups.

Prior Knowledge Test

A knowledge test was developed based on common misconceptions about mental illness to capture participants' topic knowledge before completing the inquiry activity. The test consisted of 16 multiple choice items that covered central topics related to mental illness, including key terminology, statistics, and examples (CMHA, 2018a, 2018b). Correct answers were given a score of 1 and incorrect answers were given a score of 0. Scores were added to create a total sum (out of 16 possible) and were then converted to a percentage, which was used as an indicator of participant's prior knowledge about mental illness. Cronbach's alpha reliability score was unacceptable for this instrument with $\alpha = .46$. The broad range of questions included may account for the low score, which was solely used to assess differences between the groups' prior knowledge. Most participants had moderate prior knowledge about mental illness.

Attitudes about Individuals with Mental Illness

Participants' attitudes about mental illness were collected using an adapted version of Kobau and colleagues' (2010) instrument, which is designed to assess individual's attitudes about central topics related to mental illness on two subscales: negative stereotypes and recovery and outcomes. The validated instrument consists of 11 items to be rated on a 5-point Likert scale (1 = strongly disagree, 5 = strongly agree). Participants' responses were averaged for both subscales and used as indicators of their attitudes toward individuals with mental illness.

Cronbach's alpha reliability scores were poor for both subscales with $\alpha = .60$ for negative stereotypes and $\alpha = .56$ for recovery outcomes. In our sample, the scores were low due to different factor loadings. Instead of the two-factor structured validated by Kobau and colleagues (2010), we found an underlying four-factor structure. To assess differences between the groups, we retained the original structure. Participants in both groups primarily held neutral to positive attitudes toward individuals with mental illness.

Demographic Survey

Participants were asked to provide information about their educational background and demographic information including their age, sex, political affiliations and online habits.

Results

Preliminary Analyses

Prior to conducting the full analyses, we inspected each variable for skewness, kurtosis, and outliers. Following Tabachnick and Fidell's (2013) recommendations, we used acceptable ranges of |3| for skewness and |8| for kurtosis to assess the relative normality of the distributions. Analyses revealed that the distribution for participants' overall rank-ordering of webpages and their essay scores were normally distributed. However, all macro-level justifications and some micro-level essay scores were positively skewed and leptokurtic. Participants' scores for stating a claim, supporting their claim, and evaluating claims were negatively skewed. Given that these variables represented behavioral frequencies with meaningful zero points, no transformations were performed.

To examine univariate outliers, we converted each variable to a standardized *z*-score. We classified *z*-scores exceeding |3.3| as outliers (Tabachnick & Fidell, 2013). Analyses revealed univariate outliers for four macro-level justifications: currency ($n = 2$, $z = 3.55$ to 4.23),

relevance (n = 1, z = 3.56), authority (n = 1, z = 3.91), and accuracy (n = 1, z = 4.31). All cases were retained because the values were moderate and represented less than 2% of cases for each variable (Cohen et al., 2013). To check for multivariate outliers, Mahalanobis distances were calculated based on a $\chi 2$ distribution with 5 degrees of freedom and a critical cut-off point of 20.51 (α = .001; Meyers et al., 2016; Tabachnick & Fidell, 2013). Five participants were removed from the sample for a total of 77 participants (n = 38 in training group) included in subsequent analyses. Lastly, an inspection of a bivariate correlation matrix using a recommended critical cut-off point of .70 (Meyers et al., 2016) revealed one instance of multicollinearity between participants' claim and support for their claim in the essay (r = 1.00, p < .001). To mitigate any potential confounding effect, participants' support for their claim was removed from subsequent analyses.

To assess whether there were statistically detectable differences between participants in the two groups, we examined their individual characteristics. No statistically detectable differences emerged for prior knowledge about mental illness, $F(1,76) = .49$, $p = .486$; attitudes about negative stereotypes, $F(1,76) = .02$, $p = .897$; attitudes about recovery and outcomes, $F(1,76) = 1.97$, $p = .165$; education level, $F(1,76) = .14$, $p = .710$; political party affiliation, $F(1,76) = .51$, $p = .478$; or daily media use, $F(1,76) = .40$, $p = .531$. Given the study design, we also confirmed that the training group took longer (M = 73 mins, SD = 31) to complete the study than their control counterparts (M = 50 mins, SD = 25). Descriptive statistics for each variable are presented in Table 5.

Table 5

Means and Standard Deviations of Individual Characteristics

Variables	Training group (n = 38)		Control group (n = 39)	
	M	*SD*	*M*	*SD*
Prior knowledge about mental illness	68.6	16.5	66.4	11.2
Attitude (negative stereotypes)	2.5	0.7	2.6	0.5
Attitude (recovery and outcomes)	4.0	0.6	3.9	0.5
Highest level of education	2.4	1.1	2.5	1.1
Political affiliation	4.1	1.8	3.8	1.7
Daily media use	4.7	1.5	4.5	1.3

Source Evaluation

To assess participants' accuracy in evaluating the sources, an ANOVA was conducted with the overall scores of the webpage rank-ordering as the dependent variable. Results revealed no statistically detectable differences in overall rank-ordering, $F(1,75) = 3.58$, $p = .062$, or in rank-ordering of the most and least reliable webpages, $F(1,75) = 3.00$, $p = .088$. However, there was a non-significant trend such that the training group demonstrated slightly better accuracy in their overall ranking and their ranking of the most and least reliable webpages.

To assess the role evaluation criteria may have played in participants' ranking decisions, a MANOVA was conducted with the macro-level justification scores for the webpage rank-ordering as dependent variables. Results revealed a multivariate effect of condition, Pillai's Trace = .20, $F(5, 71) = 4.33$, $p = .008$, $\eta_p^2 = .195$. Univariate tests revealed statistically detectable differences in participants' use of the currency criteria, $F(1, 75) = 5.81$, $p = .018$, $\eta_p^2 = .072$, and the accuracy criteria, $F(1, 75) = 7.89$, $p = .006$, $\eta_p^2 = .095$. The training group used the currency criteria more often to justify their rank-ordering, whereas the control group employed the accuracy criteria. Although not statistically detectable, the training group justified their rank-ordering using authority more than their counterparts, and the control group used relevance and

purpose slightly more to justify their rank-ordering. Thus, the intervention group primarily relied on currency and authority as indicators of trustworthiness, which resulted in a slightly more accurate rank-ordering. The control group primarily relied on relevance, accuracy, and purpose.

To examine whether participants used more epistemic ideals when justifying the most or least reliable webpages, we conducted ANOVAs with epistemic ideals as the dependent variable. No differences were found between groups when justifying their ranking of the most reliable webpages, $F(1, 75) = 3.93$, $p = .051$, or the least reliable webpages, $F(1, 75) = .84$, $p = .363$. However, the control group used slightly more epistemic ideals, such as evidence and objectivity, to justify their rank-ordering of the most reliable webpages. Descriptive statistics for each variable are presented in Table 6.

Table 6

Means and Standard Errors of Criteria by Condition

Variables	Training group (n = 38)		Control group (n = 39)	
	M	*SE*	*M*	*SE*
Overall rank-ordering	13.5	0.5	14.7	0.4
Most & least reliable rank-ordering	8.9	0.5	10.0	0.5
Currency*	0.8	0.2	0.2	0.1
Relevance	0.7	0.2	1.0	0.2
Authority	1.2	0.3	1.1	0.2
Accuracy*	1.0	0.2	1.7	0.2
Purpose	1.3	0.3	1.6	0.3
Most reliable webpages (criteria)	1.7	0.3	2.5	0.2
Least reliable webpages (criteria)	1.9	0.3	2.3	0.2

Notes. Lower rank-ordering scores represent a more accurate rank-ordering
*$p < .05$

Source Integration

Whereas participants' justifications served as indicators of their ability to evaluate the trustworthiness of divergent information, their essay responses were used as an indicator of their ability to compare and reconcile information with varying degrees of trustworthiness. To

examine the complexity of source integration, an ANOVA was conducted with participants' total essay score as the dependent variable. Results revealed a statistically detectable difference between the groups' essay scores, $F(1,75) = 4.19$, $p = .044$, such that the control group constructed more complex essays than the training group.

To assess the contribution of participants' justifications to their integration ability, a MANCOVA was conducted with micro-level essay scores (claim, alternative perspective, evaluate claim, reconcile perspective) as the dependent variables and macro-level justification responses as covariates. Results revealed no statistically detectable difference between the groups' essay scores after controlling for their criteria use, Pillai's Trace = .13, $F(4, 67) = 2.48$, $p = .052$. However, the control group performed slightly better on three elements of the essay: providing a claim, alternative perspective, and reconciling perspectives. Descriptive statistics for each variable are presented in Table 7.

Table 7

Means and Standard Errors of Essay Scores by Condition

Variables	Training group (n = 38)		Control group (n = 39)	
	M	*SE*	*M*	*SE*
Total essay score	2.2	0.2	2.8	0.2
Claim	0.8	0.1	1.0	0.0
Alternative perspective	0.2	0.1	0.3	0.1
Evaluate claims	0.3	0.1	0.4	0.1
Reconcile perspectives	0.1	0.0	0.2	0.1

Discussion

The present study aimed to evaluate the efficacy of a source evaluation training for adults across Canada. We investigated their evaluations and their integrations of information from authentic online sources with varying levels of reliability. Through the lens of the Apt-AIR

framework, the CRAAP test encourages users to reflect on important epistemic ideals, such as author's expertise, adequacy of evidence, and publisher's motivations, and engage in reliable epistemic processes, including knowledge-based verification and corroboration. Although widely endorsed by educators (Batchelor, 2017), the CRAAP test has been criticized for being impractical and providing misleading evaluation results because this tool does not explicitly encourage students to leave the webpage they are evaluating (Breakstone et al., 2019; Duby, 2018; Fielding, 2019). By modeling use of the CRAAP test online, this study emphasized engagement in both vertical and lateral reading of sources—an extension of the original tool. Whereas the training group provided a slightly more accurate rank-ordering of the webpages using currency and authority criteria, the control group produced a slightly better source integration. Interestingly, after justifying their rankings, both groups accurately reported using the criteria gleaned from the MANOVA.

Based on these findings, participants gravitated to a few CRAAP categories, which they used to justify their assessments of each source. The control group's use of accuracy criteria suggests that they may have engaged in lateral reading to verify evidence. In contrast, the training group's use of currency and authority criteria may indicate that they relied on vertical reading, which is supported by their limited source integrations in the essay. The present study does not support the continued endorsement of the CRAAP test as a standalone tool, even after modeling its use. Although the CRAAP test cues users to engage in important epistemic processes, which require lateral reading, the majority of included questions (88%) do not require users to leave the webpage to be answered, as Breakstone and colleagues (2018) argued. Drawing from epistemic cognition research, the CRAAP test can be improved by shifting the

focus from checklist items to criteria and interacting processes aimed at evaluating claims, sources and arguments.

In the original CRAAP test (Meriam Library of CSU Chico, 2010), currency is used as an indicator of the source's timeliness. The associated criteria relate to the source's publication or revision date, up-to-dateness in relation to the task, and hyperlink functionality. Currency evaluations, in coordination with accuracy assessments, can be leveraged to assess the appropriateness of the source's claims based on the freshness of the data used as evidence (Kąkol et al., 2017) and the quality of the evidence linked to (Chiagouris et al., 2008). Here, these criteria are placed in relation to other sources. For example, when debunking the myth that the installation of 5G towers led to the 2020 pandemic, source's that have outdated evidence that contradicts current understandings of the topic or link to noncredible websites may receive low reliability evaluations. Assessing currency has been found to be a significant predictor of selecting high-quality information (Barnes et al., 2003), as seen in the training group's slightly better performance on the rank-ordering task.

The relevance criteria also provide useful indicators of the source's importance for the user's task. This category includes task relevance, target audience, and an overall reliability judgment. Relevance assessments contribute task and context cues with which epistemic ideals and processes can be adapted (e.g., type of webpage, Barzilai & Chinn, 2018). Relevance criteria are frequently used as indicators of reliability (Kiili et al., 2008). In the present study, the control group used relevance criteria in combination with accuracy and purpose criteria. Since relevance criteria can also be used as a final judgement of the source's overall reliability after accessing the other CRAAP criteria, the control group may have used relevance criteria as a tiebreaker for sources that had similar levels of accuracy or similar purposes.

Authority criteria are common features of source evaluation tools or trainings. In the CRAAP test, authority criteria include the author's transparency, their credentials and qualifications, and the URL. To establish reliability, the original CRAAP questions can be used as a starting point to investigate the author and venue's reputation. Beyond expertise, the author's previous work and notoriety can also be used to indicate their qualifications to write on a topic. Examination of an author or publishing venue's background can also provide insight about potential biases that emerge when examining the source's purpose.

The accuracy criteria focus on the truthfulness of the source's content. Accuracy criteria include the presence of evidence and the ability to verify claims and arguments using other sources or prior knowledge. Indicators related to writing quality, style, and peer review process are also included. Evaluations of the source influence how claims and arguments are interpreted (Britt & Aglinskas, 2002; Wineburg, 1991). A shift from a criteria to a process focus would impact the CRAAP test's accuracy category greatest. Here, the trail of information must be followed, with subsequent source evaluations used to examine the quality of the evidence included, rather than solely checking that evidence is presented. For familiar topics, users may rely on their prior knowledge to verify information, and for unfamiliar topics they may rely on the source's content (Barzilai, Thomm, et al., 2020). In coordination with purpose evalautions, tone may be used as an indicator of potential biases (e.g., labelling bias).

Lastly, the purpose criteria reflect measures of the source's motivations for creating the content. Purpose criteria include categorizing the source's purpose (e.g., sell, inform, persuade), identifying their level of objectivity, and assessing biases. Funding is an indicator of a source's purpose. For example, the presence of advertisements signifies that the venue receives ad revenue, which may compromise the quality of information in favour of clicks. The author's

purpose may also be assessed by looking at the perspectives presented in their argument. Taken together, the CRAAP criteria address key aspects of online source evaluations. However, a shift from evaluation criteria to processes is needed to improve the effectiveness of the tool.

The findings from the present study and the suggestions for improving the CRAAP tool support McGrew and colleagues' recent work (McGrew, 2020; McGrew et al., 2019; Wineburg & McGrew, 2017) that explored the lateral reading behaviours of fact checkers. Wineburg and McGrew (2017) noted fact checkers' focus on embedding their evaluation within the larger context of the topic by reading laterally. The present study findings also support the relationship between source evaluations and integrations (Rouet & Britt, 2011). Unlike previous research (Barzilai, Mor-Hagani, et al, 2020; Mason et al., 2014), the control group produced more complex source integrations, often exhibiting multiple perspectives with some degree of claim evaluation or reconciliation. The control groups' use of purpose criteria, which emphasizes objectivity and balanced arguments, may have supported their ability to represent and reconcile multiple perspectives in their essays. Multiple documents research has extensively accounted for the processes that occur during source integrations, including the cognitive and metacognitive aspects of linking sources to content and each other (Salmerón et al., 2018). However, further research is needed to better understand which criteria support online source evaluations and integrations.

Limitations

This study is not without limitations which will aid future research on source evaluations. First, although we embedded indicators of treatment fidelity in the training group's design, study completion time and attention checks may not have ensured that participants watched the training video fully to apply the CRAAP criteria during the inquiry activity. Given the additional

time requirement, it is also possible that the training group may have become fatigued from the effort required to complete the tasks. To overcome this limitation, future online tutorials may benefit from an interactive design that allows participants to apply their skills through brief practice questions or a discovery approach, as seen in Pretorius' work (2018).

Second, we did not measure participants' baseline source evaluation and integration skills. Although we established that both groups were similar in age, and educational background, they may have had dissimilar experience with source evaluations. For example, the training group may have been familiar with some aspects of the CRAAP test and been prompted to use the currency and authority criteria after watching the video tutorial. By measuring participants' baseline skills, future research can better assess the impact of this training on participants' behaviours.

Lastly, the webpages selected for evaluation were published by organizations across North America, including the US, which influenced several participants' evaluations. Although only 10% of the sample justified their evaluations based on the publishing country, given the political climate at the time the study was conducted, this factor may have influenced Canadian participants. To minimize this impact, future research with authentic sources should assess the potential impact of political factors on evaluations of sources about socio-scientific issues.

Conclusions

Access to the internet has changed the way we interact with information and how we should teach important source evaluation skills. Based on the Apt-AIR framework, the CRAAP test includes assessments of authority and accuracy, which are related to desirable epistemic ideals and processes. This study makes methodological contributions to understanding source evaluations, providing online access to participants rather than a curated offline environment.

Further, this study provides evidence of the CRAAP test's shortcomings. This widely referenced evaluation tool requires explicit instruction to embed source evaluations within a larger context where multiple sources and perspectives are reconciled to establish reliability.

Acknowledgements

This research was supported by a grant from the Social Sciences and Humanities Research Council (SSHRC). The content is the sole responsibility of the authors and does not necessarily represent the official views of SSHRC.

References

Barnes, M. D., Penrod, C., Neiger, B. L., Merrill, R. M., Thackeray, R., Eggett, D. L., & Thomas, E. (2003). Measuring the relevance of evaluation criteria among health information seekers on the Internet. *Journal of Health Psychology*, *8*(1), 71–82. https://doi.org/10.1177/1359105303008001436

Barzilai, S., & Chinn, C. A. (2018). On the goals of epistemic education: Promoting apt epistemic performance. *Journal of the Learning Sciences*, *27*(3), 353–389. https://doi.org/10.1080/10508406.2017.1392968

Barzilai, S., & Eshet-Alkalai, Y. (2015). The role of epistemic perspectives in comprehension of multiple author viewpoints. *Learning and Instruction*, *36*, 86–103. https://doi.org/10.1016/j.learninstruc.2014.12.003

Barzilai, S., Mor-Hagani, S., Zohar, A. R., Shlomi-Elooz, T., & Ben-Yishai, R. (2020). Making sources visible: Promoting multiple document literacy with digital epistemic scaffolds. *Computers & Education*, *157*, 103980. https://doi.org/10.1016/j.compedu.2020.103980

Barzilai, S., Thomm, E., & Shlomi-Elooz, T. (2020). Dealing with disagreement: The roles of topic familiarity and disagreement explanation in evaluation of conflicting expert claims and sources. *Learning and Instruction*, *69*, 101367. https://doi.org/10.1016/j.learninstruc.2020.101367

Barzilai, S., & Zohar, A. (2014). Reconsidering personal epistemology as metacognition: A multifaceted approach to the analysis of epistemic thinking. *Educational Psychologist*, *49*(1), 13–35. https://doi.org/10.1080/00461520.2013.863265

Barzilai, S., Zohar, A. R., & Mor-Hagani, S. (2018). Promoting integration of multiple texts: A review of instructional approaches and practices. *Educational Psychology Review*, 1–27.

Batchelor, O. (2017). Getting out the truth: The role of libraries in the fight against fake news. *Reference Services Review*, 143-148.

Bates, B. R., Romina, S., Ahmed, R., & Hopson, D. (2006). The effect of source credibility on consumers' perceptions of the quality of health information on the Internet. *Medical Informatics and the Internet in Medicine*, *31*(1), 45–52. https://doi.org/10.1080/14639230600552601

Blakeslee, S. (2004). The CRAAP test. *Loex Quarterly*, *31*(3), 4.

Braasch, J. L. G., Bråten, I., Strømsø, H. I., Anmarkrud, Ø., & Ferguson, L. E. (2013). Promoting secondary school students' evaluation of source features of multiple documents. *Contemporary Educational Psychology*, *38*(3), 180–195. https://doi.org/10.1016/j.cedpsych.2013.03.003

Bråten, I., Strømsø, H. I., & Andreassen, R. (2016). Sourcing in professional education: Do text factors make any difference? *Reading and Writing*, *29*(8), 1599–1628.

Bråten, I., Strømsø, H. I., & Britt, M. A. (2009). Trust matters: Examining the role of source evaluation in students' construction of meaning within and across multiple texts. *Reading Research Quarterly*, *44*(1), 6–28.

Breakstone, J., McGrew, S., Smith, M., Ortega, T., & Wineburg, S. (2018). Why we need a new approach to teaching digital literacy. *Phi Delta Kappan*, *99*(6), 27–32.

Britt, M. A., & Aglinskas, C. (2002). Improving students' ability to identify and use source information. *Cognition and Instruction*, *20*(4), 485–522.

Chevrier, M., Muis., K. R., Pekrun, R., Sinatra, G., & Denton, C. A. (2019). *Thinking critically about socio-scientific issues: The roles of epistemic cognition and epistemic emotions*

[Unpublished manuscript]. Department of Educational and Counselling Psychology, McGill University.

Chiagouris, L., Long, M. M., & Plank, R. E. (2008). The consumption of online news: The relationship of attitudes toward the site and credibility. *Journal of Internet Commerce*, *7*(4), 528–549. https://doi.org/10.1080/15332860802507396

Chinn, C. A., Buckland, L. A., & Samarapungavan, A. (2011). Expanding the dimensions of epistemic cognition: Arguments from philosophy and psychology. *Educational Psychologist*, *46*(3), 141–167. https://doi.org/10.1080/00461520.2011.587722

Chinn, C. A., Rinehart, R. W., & Buckland, L. A. (2014). Epistemic cognition and evaluating information: Applying the AIR model of epistemic cognition. In D. N. Rapp & J. L.G. Braasch (Eds.), *Processing inaccurate information: Theoretical and applied perspectives from cognitive science and the educational sciences* (pp. 425–453). MIT Press.

Chinn, C., & Sandoval, W. A. (2018). Epistemic cognition and epistemic development. In F. Fischer, C. E. Hmelo-Silver, S. R. Goldman, & P. Resiman (Eds.), *International handbook of the learning sciences* (pp. 46–55). Routledge.

Canadian Mental Health Association (CMHA). (2018a). *Stigma and discrimination*. Retrieved from https://ontario.cmha.ca/documents/stigma-and-discrimination/

CMHA. (2018b). *Myths about mental illness*. Retrieved from: https://cmha.ca/documents/myths-about-mental-illness

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2013). *Applied multiple regression/correlation analysis for the behavioral sciences*. Routledge.

Cresswell, J. W., & Plano Clark, V. L. (2011). *Designing and conducting mixed methods research* (2nd ed.). Sage Publications.

Cresswell, J. W., & Plano Clark, V. L. (2017). *Designing and conducting mixed methods research* (3rd ed.). Sage Publications.

Denton, C. A., Muis, K.R., Armstrong, S., & Dubé, A. (in preparation). En-garde: Source evaluations in the digital age. *Contemporary Educational Psychology*.

Duby, H. R. (2018). What a load of CRAAP: Evaluating information in an era of "fake news". *Tennessee Libraries*, *68*(4).

Eshet-Alkalai, Y., & Chajut, E. (2009). Changes over time in digital literacy. *CyberPsychology & Behavior*, *12*(6), 713–715.

Eshet-Alkalai, Y., & Chajut, E. (2010). You can teach old dogs new tricks: The factors that affect changes over time in digital literacy. *Journal of Information Technology Education: Research*, *9*(1), 173–181.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Fielding, J. A. (2019). Rethinking CRAAP: Getting students thinking like fact-checkers in evaluating web sources. *College & Research Libraries News*, *80*(11), 620.

Gerjets, P., Kammerer, Y., & Werner, B. (2011). Measuring spontaneous and instructed evaluation processes during Web search: Integrating concurrent thinking-aloud protocols and eye-tracking data. *Learning and Instruction*, *21*(2), 220–231. https://doi.org/10.1016/j.learninstruc.2010.02.005

Green, J. (2019, Jan 22). Check yourself with lateral reading: Crash course navigating digital information #3 [Youtube video]. https://www.youtube.com/watch?v=GoQG6Tin-1E&t=157s.

Greene, J. A., Cartiff, B. M., & Duke, R. F. (2018). A meta-analytic review of the relationship between epistemic cognition and academic achievement. *Journal of Educational Psychology*. https://doi.org/10.1037/edu0000263

Greene, J. A., Copeland, D. Z., Deekens, V. M., & Seung, B. Y. (2018). Beyond knowledge: Examining digital literacy's role in the acquisition of understanding in science. *Computers & Education*, *117*, 141–159.

Greene, J. A., Yu, S. B., & Copeland, D. Z. (2014). Measuring critical components of digital literacy and their relationships with learning. *Computers & Education*, *76*, 55–69. https://doi.org/10.1016/j.compedu.2014.03.008

Halverson, K. L., Siegel, M. A., & Freyermuth, S. K. (2010). Non-science majors' critical evaluation of websites in a biotechnology course. *Journal of Science Education and Technology*, *19*(6), 612–620. https://doi.org/10.1007/s10956-010-9227-6

Hämäläinen, E. K., Kiili, C., Marttunen, M., Räikkönen, E., González-Ibáñez, R., & Leppänen, P. H. T. (2020). Promoting sixth graders' credibility evaluation of Web pages: An intervention study. *Computers in Human Behavior*, *110*, 106372. https://doi.org/10.1016/j.chb.2020.106372

Jung, E. H., Walsh-Childers, K., & Kim, H.-S. (2016). Factors influencing the perceived credibility of diet-nutrition information web sites. *Computers in Human Behavior*, *58*, 37–47. https://doi.org/10.1016/j.chb.2015.11.044

Jung, W. S., Chung, M.-Y., & Rhee, E. S. (2018). The effects of attractiveness and source expertise on online health sites. *Health Communication*, *33*(8), 962–971. https://doi.org/10.1080/10410236.2017.1323364

Kąkol, M., Nielek, R., & Wierzbicki, A. (2017). Understanding and predicting Web content credibility using the Content Credibility Corpus. *Information Processing & Management*, *53*(5), 1043–1061. https://doi.org/10.1016/j.ipm.2017.04.003

Kiili, C., Laurinen, L., & Marttunen, M. (2008). Students evaluating Internet sources: From versatile evaluators to uncritical readers. *Journal of Educational Computing Research*, *39*(1), 75–95. https://doi.org/10.2190/EC.39.1.e

Kobau, R., DiIorio, C., Chapman, D., Delvecchio, P., & SAMHSA/CDC Mental Illness Stigma Panel Members. (2010). Attitudes about mental illness and its treatment: Validation of a generic scale for public health surveillance of mental illness associated stigma. *Community Mental Health Journal*, *46*(2), 164–176. https://doi.org/10.1007/s10597-009-9191-x

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159. https://doi.org/10.2307/2529310

Machackova, H., & Smahel, D. (2018). The perceived importance of credibility cues for the assessment of the trustworthiness of online information by visitors of health-related websites: The role of individual factors. *Telematics and Informatics*, *35*(5), 1534–1541. https://doi.org/10.1016/j.tele.2018.03.021

Mason, L., Ariasi, N., & Boldrin, A. (2011). Epistemic beliefs in action: Spontaneous reflections about knowledge and knowing during online information searching and their influence on learning. *Learning and Instruction*, *21*(1), 137–151. https://doi.org/10.1016/j.learninstruc.2010.01.001

Mason, L., Junyent, A. A., & Tornatora, M. C. (2014). Epistemic evaluation and comprehension of web-source information on controversial science-related topics: Effects of a short-term

instructional intervention. *Computers & Education, 76,* 143–157. https://doi.org/10.1016/j.compedu.2014.03.016

Mason, L., Scrimin, S., Tornatora, M. C., Suitner, C., & Moè, A. (2018). Internet source evaluation: The role of implicit associations and psychophysiological self-regulation. *Computers & Education*, *119,* 59–75. https://doi.org/10.1016/j.compedu.2017.12.009

McGrew, S. (2020). Learning to evaluate: An intervention in civic online reasoning. *Computers & Education, 145,* 103711. https://doi.org/10.1016/j.compedu.2019.103711

McGrew, S., Smith, M., Breakstone, J., Ortega, T., & Wineburg, S. (2019). Improving university students' web savvy: An intervention study. *British Journal of Educational Psychology,* bjep.12279. https://doi.org/10.1111/bjep.12279

Meppelink, C. S., Smit, E. G., Fransen, M. L., & Diviani, N. (2019). "I was right about vaccination": Confirmation bias and health literacy in online health information seeking. *Journal of Health Communication*, *24*(2), 129–140. https://doi.org/10.1080/10810730.2019.1583701

Meriam Library of CSU Chico. (2010). Evaluating information - Applying the CRAAP test [PDF document]. *Meriam Library of CSU Chico*. https://library.csuchico.edu/sites/default/files/craap-test.pdf

Meyers, L. S., Gamst, G., & Guarino, A. J. (2016). *Applied multivariate research: Design and interpretation*. Sage Publications.

Morse, J. M. (2003). Principles of mixed methods and multimethod research design. In A. Tashakkori, & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 189–208). Sage Publications.

Pretorius, L. (2018). Experiential and self-discovery learning in digital literacy: Developing the discernment to evaluate source reliability. *College & Undergraduate Libraries*, *25*(4), 388–405. https://doi.org/10.1080/10691316.2018.1530626

Queen's University Library. (2020). CRAAP criteria. *Bachelor of health sciences toolbox.* https://guides.library.queensu.ca/BHSc/CRAAP

Richter, T., & Maier, J. (2017). Comprehension of multiple documents with conflicting information: A two-step model of validation. *Educational Psychologist*, *52*(3), 148–166. https://doi.org/10.1080/00461520.2017.1322968

Richter, T., & Schmid, S. (2010). Epistemological beliefs and epistemic strategies in self-regulated learning. *Metacognition and Learning*, *5*(1), 47–65.

Rouet, J.-F., & Britt, M. A. (2011). Relevance processes in multiple document comprehension. In M. T. McCrudden, J. P. Magliano, & G. Schraw (Eds.), *Text relevance and learning from text* (pp. 19–52). IAP Information Age Publishing.

Salmerón, L., Gil, L., & Bråten, I. (2018). Effects of reading real versus print-out versions of multiple documents on students' sourcing and integrated understanding. *Contemporary Educational Psychology*, *52*, 25–35. https://doi.org/10.1016/j.cedpsych.2017.12.002

Sandoval, W. A. (2017). Situating practices of epistemic cognition. In T. G. Amin & O. Levrini (Eds.), *Converging perspectives on conceptual change* (1st ed., pp. 253–260). Routledge. https://doi.org/10.4324/9781315467139-30

Smallwood, C. (2015). *The complete guide to using Google in libraries: Instruction, administration, and staff productivity* (Vol. 1). Rowman & Littlefield.

Stanford History Education Group. (2016). *Evaluating information: The cornerstone of online civic reasoning* [PDF document].

https://stacks.stanford.edu/file/druid:fv751yt5934/SHEG%20Evaluating%20Information%20Online.pdf

Stanford History Education Group. (2020). COR. *Civic Online Reasoning*. https://cor.stanford.edu

Subramaniam, M., Taylor, N. G., St. Jean, B., Follman, R., Kodama, C., & Casciotti, D. (2015). As simple as that?: Tween credibility assessment in a complex online world. *Journal of Documentation, 71*(3), 550–571. https://doi.org/10.1108/JD-03-2014-0049

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics: International edition*. Pearson.

Taylor, W. D., Johnson, G., Ault, M. K., Griffith, J. A., Rozzell, B., Connelly, S., Jensen, M. L., Dunbar, N. E., & Ness, A. M. (2015). Ideological group persuasion: A within-person study of how violence, interactivity, and credibility features influence online persuasion. *Computers in Human Behavior*, *51*, 448–460. https://doi.org/10.1016/j.chb.2015.04.071

The News Literacy Project. (2020). *Checkology*. Checkology. https://get.checkology.org

Ulyshen, T. Z., Koehler, M. J., & Gao, F. (2015). Understanding the connection between epistemic beliefs and Internet searching. *Journal of Educational Computing Research*, *53*(3), 345–383.

Van Strien, J. L., Brand-Gruwel, S., & Boshuizen, H. P. (2014). Dealing with conflicting information from multiple nonlinear texts: Effects of prior attitudes. *Computers in Human Behavior*, *32*, 101–111.

Walraven, A., Brand-Gruwel, S., & Boshuizen, H. P. A. (2013). Fostering students' evaluation behaviour while searching the internet. *Instructional Science*, *41*(1), 125–146. https://doi.org/10.1007/s11251-012-9221-x

White, A. M., DeCuir-Gunby, J. T., & Kim, S. (2019). A mixed methods exploration of the relationships between the racial identity, science identity, science self-efficacy, and science achievement of African American students at HBCUs. *Contemporary Educational Psychology*, *57*, 54–71. https://doi.org/10.1016/j.cedpsych.2018.11.006

Wiley, J., Goldman, S. R., Graesser, A. C., Sanchez, C. A., Ash, I. K., & Hemmerich, J. A. (2009). Source evaluation, comprehension, and learning in Internet science inquiry tasks. *American Educational Research Journal*, *46*(4), 1060–1106. https://doi.org/10.3102/0002831209333183

Wineburg, S., & McGrew, S. (2017). Lateral reading: Reading less and learning more when evaluating digital information [PDF document]. https://pdfs.semanticscholar.org/5b26/9628f4dc29b514dfcb0b5e429e49fc0dae6d.pdf?_ga=2.134731505.1090439893.1607365627-347836626.1607210879

Wineburg, S. S. (1991). On the reading of historical texts: Notes on the breach between school and academy. *American Educational Research Journal*, *28*(3), 495–519.

Appendices

Appendix A

Justification Criteria Coding Scheme with Illustrative Examples

Micro-level category	***Description***	***Illustrative example (Participant number)***
	Macro-level category: Currency	
Date	An evaluative decision about a rank based on participant's consideration of webpage's various publication dates. Participant may mention any of the following: • Original publication date • Revision date • Copyright date	"The time stamp and copyrights are up to date" (PN 227) "The currency of the article at first glance is from 2014, though on further inspection, it was last modified in October 2018 and contains relevant information from 2018." (PN 255)
Hyperlink	An evaluative decision about a rank based on learner's consideration of webpage's hyperlinks. Learner may mention: • Functionality of hyperlink	"Second hyperlink lead to an error page which raised a flag on reliability." (PN 269)
Design	An evaluative decision about a rank based on participant's consideration of webpage's design or format features. Participant may mention any of the following: • Links to social media or other sharing capabilities • Pop up boxes • Videos/images/advertisements • Format of webpage or its content	"Webpage opened with advertising all over it, pop ups at the bottom, etc." (PN 293) "The web page style hasn't been updated in a decade." (PN 232)

Macro-level category: Relevance		
Alignment	An evaluative decision about a rank based on participant's consideration of webpage's alignment with their task definition. Participant may mention any of the following: • topic relevance (e.g., central or peripheral to task) • appropriateness of evidence for task (e.g., type of evidence) • current enough to complete task	"Being a national review, it might not pinpoint exactly what you are looking for on a local level" (PN 213) "It talks quite a bit about substance abuse so it's very relative to the topic" (PN 242)
Uniqueness	An evaluative decision about a rank based on participant's consideration of how unique the information presented on the webpage is in relation to other sources.	"I believe the information in this article is unique." (PN 256)
Type	An evaluative decision about a rank based on participant's consideration of the type or location of the webpage. Participant may mention any of the following: • News article or news section • Journal article • Blog • Encyclopedia	"Columbia Valley Pioneer - A published news article / blog is considered a lesser credible source of information based on the fact that there can be underlying unknown factors behind the write-up." (PN 224) "Although the website has a quite in-depth analysis and has many reliable sources, I ranked it fifth on the list. This is because it is new article. In comparison to the other sites I would place it lower in reliability." (PN 277)

	Macro-level category: Authority	
Author expertise	An evaluative decision about a rank based on participant's consideration of author or organization's expertise related to the topic. Participant may mention any of the following: • author or organization's credentials (e.g., educational degree) • author or organization's affiliations • author or organization's years of experience • author or organization's funding status (e.g., what area are they funded for?)	"It's a government website which draws data from credible scientists who work for StatsCan." (PN 209) "It seems to me an important center and concerned with mental health and with a lot of experience." (PN 220)
Venue	An evaluative decision about a rank based on participant's consideration of the body that published the webpage. Participant may mention any of the following: • publisher's reputation (e.g., factual reporting, sourcing practices, fact checking system) • publisher's funding • publisher's ownership	"Because it was made by Statistics Canada which should be the more reliable since it has been made by the government." (PN 219)

	Macro-level category: Accuracy	
Evidence	An evaluative decision about a rank based on participant's consideration of the sources of information provided as evidence. Participant may mention any of the following: • quality of evidence • type of evidence • source of evidence • currency of evidence • sufficiency of evidence	"Although this article does have some evidence to back up its claims, they are relatively few." (PN 203) "The article's accuracy is questionable since the sources used for the examples used aren't scientific." (PN 255)
Corroboration	An evaluative decision about a rank based on participant's consideration of whether information can be corroborated or not in other sources.	"I don't know if it's reliable or not, less statistics and goes against another one [source]." (PN 292) "I have learnt similar information at school. I also learnt from home and my own experience." (PN 302)
Tone	An evaluative decision about a rank based on participant's consideration of the author's tone.	"It's cold and unemotional raw data without any incentive to skew results to prove a point." (PN 209) "I also found that it had slight bias and lacked a neutral tone." (PN 233)
Peer-review	An evaluative decision about a rank based on participant's consideration of the review process of information presented on the webpage.	"I believe information posted on a government website has likely been heavily vetted to ensure the government is not promoting false information." (PN 241) "The article was written by one person so it could be biased." (PN 294)

Macro-level category: Purpose		
Argument	An evaluative decision based on participant's consideration of author or organization's argument structure (e.g., one sided or two sided argument).	"Seems to come from a very one sided source. Does not talk about positive part of mental illness victims." (PN 262) "Invalid reasoning and conclusion." (PN 300)
Bias	An evaluative decision based on participant's consideration of author or organization's bias(es). Participant may mention any of the following: 1. source selection 2. labelling (e.g., expert, spokesperson, official)	"The mediabias [rating from mediabiasfactcheck.com] was good IMO." (PN 290) "Their title attracts the wrong idea, even though are talking about how people with mental illness are not actually more violent the title of the article gives the opposite idea." (PN 294)
Objectivity	An evaluative decision based on participant's consideration of source's objectivity. Participant may mention any of the following: 1. subjective and opinion-based writing 2. fact-driven and objective writing with facts, statistics, or research	"This is an opinion piece, whereas the other articles are objective reporting, based on studies and research, written from the third-person's perspective." (PN 203)
Purpose	An evaluative decision based on participant's consideration of source's purpose. Participant may identify author's/organization's purpose for creating content as any of the following*: • Persuade – author argues a position • Provoke – author uses tone or emotionally-driven language as a call to action for readers • Inform or document – provide details of context or procedure • Entertain - author wants readers to enjoy or be amused by content • Sell – author wants readers to buy products or services	"Because it uses an emotion provoking image (police officers with the yellow tape) to try and influence the way I perceive a topic, so that I am more willing to accept the information they present (due to being influenced emotionally)." (PN 297)

Appendix B

Essay Score Coding Scheme with Illustrative Examples

	1 point	**0 points**	**Illustrative examples (Participant number)**
Making a claim	Participant makes a claim or takes a position, regardless of the claim's accuracy.	Participant does not make a claim or they do not justify their claim.	"In my opinion, I think that individuals with mental illness are more violent than those without mental illness." (PN 202)
Presenting supportive grounds (One-sided argument)	Participant supports their position with valid arguments and evidence.	Participant does not support their position, or their arguments and evidence are invalid.	"Based on the presented data, it would seem that, in general, there is no direct connection between outward violence and mental illness. If anything, mentally ill are more susceptible to inward violence. That is to say their illness might cause them to be victimized more often than general population." (PN 258)
Presenting alternative perspectives (Two-sided argument)	Participant engages with an alternative perspective by identifying valid arguments and evidence in support of that perspective.	Participant only presents arguments that support one perspective. Participant may acknowledge an alternative perspective but does so without providing evidence to support that perspective.	"Now in my response I did say "generally" since in some specific types of mental illness such as schizophrenia or psychosis, there is an increased risk of violence. There are complexities in understanding the link between violence and mental illness due to co-existing factors such as substance abuse, lower income, and marginalization. With these co-factors present, violence may be erroneously correlated to mental illness." (PN 269)
Evaluating claims	Participant explains whether a claim is credible, limited or generalizable.	Participant accepts or dismisses a claim without evaluating it or explaining why it should be accepted or rejected.	"Since there are so many different kinds of mental illness and the degree of one person's mental illness can differ from someone else. One person might be stable with proper treatment and medication while others are not receiving treatment. So one person could have violent tendencies, so then media portrays all mental illness as being more violent. Instead each violent case should be examined individually because you can't stereotype one trait to an entire group of people." (PN 294)
Reconciling or integrating perspectives	Participant acknowledges the validity of arguments on both sides, which is consistent with their evaluation of claims.	Participant does not include a conclusion or creates a one-sided conclusion that does not include an alternative perspective.	"I agree that there are some people who suffer from mental illness that are extremely violent, that need to be confined, and can react unexpectedly. However, I believe that those are a small fraction of the mental health population. The media is quick to portray someone with mental health issues as violent as it adds to entertainment value. This perpetuates the stereotype. I do not

	believe that having a mental illness makes you a more violent person. However, I do believe and it was confirmed by the websites that people with mental illness are more likely to be victims of crime." (PN 281)

Chapter 5

General Discussion

Access to the internet has brought many opportunities and challenges for information consumers. With an abundance of unfiltered information, users are responsible for assessing the quality of the information they encounter—an increasingly daunting task when authors and publishers are obfuscated and technology changes rapidly (Ala-Mutka, 2011; Greene, 2016). A growing body of research has documented that difficulties with source evaluations occur at all ages, which has ignited efforts to improve students' evaluation skills (Braasch et al., 2013; McGrew et al., 2019; Zhang & Duke, 2011). These skills are important antecedents of the source integrations that are incorporated into many classroom activities at the heart of education (e.g., essays, Rouet & Britt, 2011). Formal education plays an important role in preparing young information consumers to perform in epistemically unfriendly environments, such as the internet (Barzilai and Chinn, 2018). Epistemic cognition, or thinking about the acquisition, justification and use of knowledge (Hofer, 2016), is a promising avenue to better understand how to support source evaluations (Greene & Yu, 2016) and integrations (Barzilai & Zohar, 2018). Unfortunately, educators and researchers have often used curated environments to train students on these important skills (Wiley et al., 2009), ones that mirror formal classroom activities. As a result, findings about EC have been somewhat inappropriately extended to online source evaluations and integrations (see Chapter 2). Further, recent theoretical work (Barzilai & Chinn, 2018) has delineated the relationship between EC and source evaluations. Despite the hypothesized utility of epistemic processes during source evaluations, some empirical investigations have reported low engagement in EC (Mason et al., 2014; Greene et al., 2018), whereas others have documented higher rates (Kąkol & Nielek, 2015; Kąkol et al., 2017). Given the situated nature of EC (Chinn & Sandoval, 2018), much remains unknown about the role and the importance of these skills for evaluating the quality of information on the internet.

The purpose of this thesis was to address this gap in the literature by examining indicators of epistemic metacognition and EC. By adopting the Apt-AIR framework (Barzilai & Chinn, 2018), these constructs were investigated using a mixed methods multiphase research design that included (1) a literature review covering relevant research in multiple fields, (2) focus group interviews, (3) source evaluation tasks, and (4) a source integration task. As a culmination of the present work, the popular and promulgated CRAAP test was evaluated to determine if it is indeed an effective way to evaluate information on the internet. Drawing from multiple fields, this thesis has contributed methodological suggestions for future research on source evaluations in an online context and supported that some findings from offline contexts hold true in both environments.

Contributions

First and foremost, the present thesis identified that research on source evaluations has lacked ecological validity (Chapter 2), and then source evaluations were explored during more authentic tasks (Chapters 3 and 4). Although there are similarities between source evaluations in online and offline environments, as Chinn and Sandoval (2018) argued, the details of the epistemic processes differ as a result of context. One key affordance of conducting online research is allowing participants to access additional information to evaluate the sources supplied. In contrast, participants establishing reliability in offline contexts must rely on the content of the supplied sources and their prior knowledge. Access to an online environment allows participants to trace the trail of information, as Cameron described (Chapter 3, Study 1). Online access also provides diverse features that can be used to assess reliability, such as advertisements (Krouwer et al., 2020), links to evidence with varying levels of reliability (Chiagouris et al., 2008), and user-generated content (Naab et al., 2020). Given these features,

Chapter 3 reexamined previous assumptions that epistemic ideals, as opposed to non-epistemic criteria, are superior during source evaluations. Overall, the results suggest that non-epistemic criteria provide important context clues that can support adaptive and accurate source evaluations (Chapters 3 & 4). In Muis and colleagues' (2018) theoretical work on epistemic emotions, facets of EC are organized within the four phases of self-regulated learning, with task and context factors influencing epistemic aims, ideals, and processes. The present thesis initiates further source evaluation research that takes these conditions into consideration.

Second, the efficacy of a training based on a popular source evaluation tool was examined with adults in Chapter 4. Although the CRAAP test includes important epistemic ideals and hints at epistemic processes, it cannot be recommended as a standalone evaluation tool, even with modeling. Chapter 4 contributes empirical evidence that the CRAAP test should be endorsed with caution by educational institutions. Like many checklists, the CRAAP test's emphasis on evaluation criteria, rather than epistemic processes, may be its greatest shortcoming. Chinn and colleagues (2020) pointed out that the controlled environment in which such checklist approaches are used are problematic. However, the present findings suggest that the narrow focus on improving epistemic ideals limits the effectiveness of these approaches even further. This thesis supports the efforts of McGrew and colleagues' work (McGrew, 2020; McGrew et al., 2019) training students on epistemic processes used by professional fact checkers.

Third, as Greene and colleagues (2008; Cartiff et al., 2020) called for, Chapter 4 explored EC and source evaluations in a novel population for education research: adults. Researchers have found that source evaluations improve with age and experience (Eshet-Alkali & Chajut, 2009, 2010), yet source evaluations are a lifelong skill that require updates as technologies change (Ala-Mutka, 2011). Thus, source evaluations, and the underlying EC that supports them, must

become adaptive. In the Apt-AIR framework, Barzilai and Chinn (2018) dedicated Aspect 2 to adaption, where they described selecting appropriate EC for different situations. The qualitative nature of this thesis adds illustrative examples of some epistemic ideals and processes used to assess reliability on the internet. In Chapter 3, participants identified many epistemic ideals that were observed in the evaluation tasks described in both manuscripts. Epistemic ideals were captured in focus group interviews (Chapter 3, Study 1) and justifications of reliability judgements (Chapters 3 & 4). Chapter 4 also captured creative epistemic processes in written arguments. The coding schemes and examples gathered in authentic tasks elaborate on Barzilai and Chinn's (2018) exemplar of evaluating a website and may be used in further research on the adaptive nature of EC during source evaluations.

Lastly, the present thesis approached this investigation from a pragmatist perspective, where the strengths of diverse frameworks and methodologies were drawn on to understand source evaluations (Johnson & Onwuegbuzie, 2004). By examining source evaluation in adults, the present thesis drew on the methods used by education researchers and the online contexts used in other disciplines to extend research on evaluation skills outside of academic contexts. As a result, the present thesis introduced more authentic materials and contexts for EC research to investigate source evaluations. Next, limitations which can be used to guide future research efforts will be discussed.

Limitations and Future Directions

The present thesis highlights the cognitive and, to a lesser extent, the metacognitive aspects of personal epistemology during source evaluations; however, epistemic motivation and epistemic emotions undoubtedly influenced the findings. The narrow empirical focus on epistemic ideals in the present thesis limited the ability to contextualize epistemic ideals within

the broader scope of the cognitive and metacognitive aspects of epistemic aims and reliable processes. Although Chapter 3 (Study 1) collected participants' epistemic metacognitive knowledge about strategies and tasks (Barzilai & Zohar, 2014, 2016), these processes were not captured in the other studies. During source evaluations, epistemic metacognitive knowledge serves the important purpose of monitoring EC as the epistemic properties of sources and evidence are examined and combined in integrations. Recent research has explored the role of epistemic metacognition in source evaluations and integrations, noting that participants who received cognitive and metacognitive support on these processes increased their performance on the tasks as well as their metacognitive knowledge about strategies and tasks (Barzilai et al., 2017, 2020). Training that supports epistemic thinking in online environments is a promising avenue to improve adaptive source evaluations. Chinn and colleagues (2020) recommended that such training can also support epistemic motivation when students examine topics that are important to them.

Source evaluations and integrations require concerted effort, and consequently, motivation to engage in these processes. Motivation has the potential to influence source evaluations in many ways, including the features that are attended to (Sillence et al., 2007). As such, in the present thesis the level of motivation to engage in EC may have influenced commitment to completing the tasks, which may have resulted in using different epistemic ideals than in a natural setting. Participants may have initially attempted to discern the most and least reliable source(s) and disengaged if they became confused by differing perspectives, as Josephine described in Chapter 3 (Study 1). To mitigate the impact of motivation on source evaluations, future research should consider more naturalistic settings, such as open-ended internet browsing or recruiting participants when they access websites as Machackova and

Smahel (2018) did. Such research may illuminate whether differences in motivation in favour of natural settings exist, or if experimental settings stimulate evaluations that do not occur during browsing.

The narrow focus on previous empirical work from the learning sciences in Chapter 2 (Part 2) also limited the research design and interpretation of the findings in Chapters 3 and 4. The present research included measures of important individual differences that were identified in Chapter 2 (Part 2); however, participants' familiarity with the supplied sources was not measured (Chapters 3 and 4). Whereas prior knowledge and topic familiarity go hand in hand, other fields have identified that source familiarity also plays a prominent role in source evaluations. Lucassen and Schraagen (2011, 2013) described previous experience with a source as a passive strategy that may be used alongside content-based evaluations or on its own during source evaluations. When using authentic sources, participants may have previous experiences with the selected publishing venues that influence their source evaluations. Future research should consider a broader review of source evaluation research that includes literature from information sciences and computer science to improve the design of research and the interpretation of source evaluation behaviours.

Controversial socio-scientific issues have been widely used as a context to explore EC. Access to the internet opens the possibility that users will encounter perspectives with which they do not agree. Like epistemic motivation, the present thesis did not explore participants' epistemic emotions during their source evaluations, which may have supported or constrained their EC (Muis et al., 2018). In naturalistic settings, the realities associated with algorithms and self-segregation online (Song et al., 2020) minimize the likelihood that users will experience cognitive dissonance, an important landmark described in many theories of personal

epistemology. Whereas the present thesis extended previous EC research by using an online environment, the underlying architecture of the internet as an information source was not explored in the adapted CRAAP training (Chapter 4). Fielding (2019) argued that training should emphasize source evaluation skills alongside knowledge about the nature and structure of various sources on the internet. To continue moving in the direction of more authentic activity, EC research has a clear path to developing reliable epistemic processes on the internet by incorporating training about these online features. Despite these limitations, the present thesis leveraged the strengths of multiple fields to extend EC research.

Conclusions

Thinking about the nature of knowledge and knowing is complex. Combined with the intricacies of online environments, there are many promising avenues to better understand source evaluations. The purpose of this thesis was to identify how theoretical conceptualizations of EC could guide research with adults in the underexplored and complex context of the internet. The consequences of inadequate source evaluations are long reaching—potentially impacting both the individual and society. Education is at the forefront of supporting adaptive source evaluations on the internet. To do so, research in this fraught environment is essential.

References

Ala-Mutka, K. (2011). *Mapping digital competence: Towards a conceptual understanding* {PDF document]. Institute for Prospective Technological Studies. ftp://s-jrcsvqpx102p.jrc.es/pub/EURdoc/JRC67075_TN.pdf

Barzilai, S., & Chinn, C. A. (2018). On the goals of epistemic education: Promoting apt epistemic performance. *Journal of the Learning Sciences*, *27*(3), 353–389. https://doi.org/10.1080/10508406.2017.1392968

Barzilai, S., & Ka'adan, I. (2017). Learning to integrate divergent information sources: The interplay of epistemic cognition and epistemic metacognition. *Metacognition and Learning*, *12*(2), 193–232.

Barzilai, S., Mor-Hagani, S., Zohar, A. R., Shlomi-Elooz, T., & Ben-Yishai, R. (2020). Making sources visible: Promoting multiple document literacy with digital epistemic scaffolds. *Computers & Education*, *157*, 103980. https://doi.org/10.1016/j.compedu.2020.103980

Barzilai, S., & Zohar, A. (2014). Reconsidering personal epistemology as metacognition: A multifaceted approach to the analysis of epistemic thinking. *Educational Psychologist*, *49*(1), 13–35. https://doi.org/10.1080/00461520.2013.863265

Barzilai, S., & Zohar, A. (2016). Epistemic (meta) cognition: Ways of thinking about knowledge and knowing. In J. A. Greene, W. A. Sandoval, & I. Bråten (Eds.), *Handbook of epistemic cognition* (pp. 409–424). Routledge.

Barzilai, S., Zohar, A. R., & Mor-Hagani, S. (2018). Promoting integration of multiple texts: A review of instructional approaches and practices. *Educational Psychology Review*, 1–27.

Braasch, J. L. G., Bråten, I., Strømsø, H. I., Anmarkrud, Ø., & Ferguson, L. E. (2013). Promoting secondary school students' evaluation of source features of multiple

documents. *Contemporary Educational Psychology*, *38*(3), 180–195. https://doi.org/10.1016/j.cedpsych.2013.03.003

Cartiff, B. M., Duke, R. F., & Greene, J. A. (2020). The effect of epistemic cognition interventions on academic achievement: A meta-analysis. *Journal of Educational Psychology*. https://doi.org/10.1037/edu0000490

Chiagouris, L., Long, M. M., & Plank, R. E. (2008). The consumption of online news: The relationship of attitudes toward the site and Ccedibility. *Journal of Internet Commerce*, *7*(4), 528–549. https://doi.org/10.1080/15332860802507396

Chinn, C. A., Barzilai, S., & Duncan, R. G. (2020). Education for a "post-truth" world: New directions for research and practice. *Educational Researcher*, 0013189X2094068. https://doi.org/10.3102/0013189X20940683

Chinn, C., & Sandoval, W. A. (2018). Epistemic cognition and epistemic development. In F. Fischer, C. E. Hmelo-Silver, S. R. Goldman, & P. Resiman (Eds.), *International handbook of the learning sciences* (pp. 46–55). Routledge.

Eshet-Alkalai, Y., & Chajut, E. (2009). Changes over time in digital literacy. *CyberPsychology & Behavior*, *12*(6), 713–715.

Eshet-Alkalai, Y., & Chajut, E. (2010). You can teach old dogs new tricks: The factors that affect changes over time in digital literacy. *Journal of Information Technology Education: Research*, *9*(1), 173–181.

Fielding, J. A. (2019). Rethinking CRAAP: Getting students thinking like fact-checkers in evaluating web sources. *College & Research Libraries News*, *80*(11), 620.

Greene, J. A. (2016). Interacting epistemic systems within and beyond the classroom. In J. A. Greene, W. A. Sandoval, & I. Bråten (Eds.), *Handbook of epistemic cognition* (pp. 265–278). Routledge.

Greene, J. A., Azevedo, R., & Torney-Purta, J. (2008). Modeling epistemic and ontological cognition: Philosophical perspectives and methodological directions. *Educational Psychologist*, *43*(3), 142–160. https://doi.org/10.1080/00461520802178458

Greene, J.A., Copeland, D. Z., Deekens, V. M., & Seung, B. Y. (2018). Beyond knowledge: Examining digital literacy's role in the acquisition of understanding in science. *Computers & Education*, *117,* 141–159.

Greene, J. A., & Yu, S. B. (2016). Educating critical thinkers: The role of epistemic cognition. *Policy Insights from the Behavioral and Brain Sciences*, *3*(1), 45–53. https://doi.org/10.1177/2372732215622223

Hofer, B. K. (2016). Epistemic cognition as a psychological construct: Advancements and challenges. In J. A. Greene, W. Sandoval, & I. Bråten (Eds.), *Handbook of epistemic cognition* (pp. 31–50). Routledge.

Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher*, *33*(7), 14–26. https://doi.org/10.3102/0013189X033007014

Kąkol, M., & Nielek, R. (2015). What affects web credibility perception? An analysis of textual justifications. *Computer Science*, *16*(3), 295-310. http://dx.doi.org/10.7494/csci.2015.16.3.295

Kąkol, M., Nielek, R., & Wierzbicki, A. (2017). Understanding and predicting Web content credibility using the Content Credibility Corpus. *Information Processing & Management*, *53*(5), 1043–1061. https://doi.org/10.1016/j.ipm.2017.04.003

Krouwer, S., Poels, K., & Paulussen, S. (2020). Moving towards transparency for native advertisements on news websites: A test of more detailed disclosures. *International Journal of Advertising*, *39*(1), 51–73. https://doi.org/10.1080/02650487.2019.1575107

Lucassen, T., & Schraagen, J. M. (2011). Factual accuracy and trust in information: The role of expertise. *Journal of the American Society for Information Science and Technology*, *62*(7), 1232–1242.

Lucassen, T., & Schraagen, J. M. (2013). The influence of source cues and topic familiarity on credibility evaluation. *Computers in Human Behavior*, *29*(4), 1387–1392. https://doi.org/10.1016/j.chb.2013.01.036

Machackova, H., & Smahel, D. (2018). The perceived importance of credibility cues for the assessment of the trustworthiness of online information by visitors of health-related websites: The role of individual factors. *Telematics and Informatics*, *35*(5), 1534–1541. https://doi.org/10.1016/j.tele.2018.03.021

Mason, L., Junyent, A. A., & Tornatora, M. C. (2014). Epistemic evaluation and comprehension of web-source information on controversial science-related topics: Effects of a short-term instructional intervention. *Computers & Education*, *76*, 143–157. https://doi.org/10.1016/j.compedu.2014.03.016

McGrew, S. (2020). Learning to evaluate: An intervention in civic online reasoning. *Computers & Education, 145*, 103711. https://doi.org/10.1016/j.compedu.2019.103711

McGrew, S., Smith, M., Breakstone, J., Ortega, T., & Wineburg, S. (2019). Improving university students' web savvy: An intervention study. *British Journal of Educational Psychology*, bjep.12279. https://doi.org/10.1111/bjep.12279

Muis, Chevrier, & Singh. (2018). The role of epistemic emotions in personal epistemology and self-regulated learning. *Educational Psychologist*, 1–20.

Naab, T. K., Heinbach, D., Ziegele, M., & Grasberger, M.-T. (2020). Comments and credibility: How critical user comments decrease perceived news article credibility. *Journalism Studies*, *21*(6), 783–801.

Rouet, J.-F., & Britt, M. A. (2011). Relevance processes in multiple document comprehension. In M. T. McCrudden, J. P. Magliano, & G. Schraw (Eds.), *Text relevance and learning from text* (p. 19–52). IAP Information Age Publishing.

Sillence, E., Briggs, P., Harris, P. R., & Fishwick, L. (2007). How do patients evaluate and make use of online health information? *Social Science & Medicine*, *64*(9), 1853–1862.

Song, H., Cho, J., & Benefield, G. A. (2020). The dynamics of message selection in online political discussion forums: Self-segregation or diverse exposure? *Communication Research*, *47*(1), 125–152.

Wiley, J., Goldman, S. R., Graesser, A. C., Sanchez, C. A., Ash, I. K., & Hemmerich, J. A. (2009). Source evaluation, comprehension, and learning in Internet science inquiry tasks. *American Educational Research Journal*, *46*(4), 1060–1106. https://doi.org/10.3102/0002831209333183

Zhang, S., & Duke, N. K. (2011). The impact of instruction in the WWWDOT framework on students' disposition and ability to evaluate web sites as sources of information. *The Elementary School Journal*, *112*(1), 132–154. https://doi.org/10.1086/660687

Bibliography

Ala-Mutka, K. (2011). *Mapping digital competence: Towards a conceptual understanding* [PDF document]. Institute for Prospective Technological Studies. ftp://s-jrcsvqpx102p.jrc.es/pub/EURdoc/JRC67075_TN.pdf

Alexander, P. A. (2016). The arguments for and the reasoning about epistemic cognition. In J. A. Greene, W. A. Sandoval, & I. Bråten (Eds.), *Handbook of epistemic cognition* (pp. 112–122). Routledge.

Alkali, Y. E., & Amichai-Hamburger, Y. (2004). Experiments in digital literacy. *CyberPsychology & Behavior*, *7*(4), 421–429.

Argelagós, E., & Pifarré, M. (2012). Improving information problem solving skills in secondary education through embedded instruction. *Computers in Human Behavior*, *28*(2), 515–526. https://doi.org/10.1016/j.chb.2011.10.024

Aribarg, A., & Schwartz, E. M. (2020). Native advertising in online news: Trade-offs among clicks, brand recognition, and website trustworthiness. *Journal of Marketing Research*, *57*(1), 20–34. https://doi.org/10.1177/0022243719879711

Australian Curriculum. (2020). *Digital technologies*. Australian Curriculum. https://www.australiancurriculum.edu.au/f-10-curriculum/technologies/digital-technologies/?year=12983&year=12984&year=12985&year=12986&year=12987&strand=Digital+Technologies+Knowledge+and+Understanding&strand=Digital+Technologies+Processes+and+Production+Skills&capability=ignore&capability=Literacy&capability=Numeracy&capability=Information+and+Communication+Technology+%28ICT%29+Capability&capability=Critical+and+Creative+Thinking&capability=Personal+and+Social+Capability&capability=Ethical+Understanding&capability=Intercultural+Understanding&priority

=ignore&priority=Aboriginal+and+Torres+Strait+Islander+Histories+and+Cultures&priority=Asia+and+Australia's+Engagement+with+Asia&priority=Sustainability&elaborations=true&elaborations=false&scotterms=false&isFirstPageLoad=false

Barnes, M. D., Penrod, C., Neiger, B. L., Merrill, R. M., Thackeray, R., Eggett, D. L., & Thomas, E. (2003). Measuring the relevance of evaluation criteria among health information seekers on the Internet. *Journal of Health Psychology*, *8*(1), 71–82. https://doi.org/10.1177/1359105303008001436

Barzilai, S., & Chinn, C. A. (2018). On the goals of epistemic education: Promoting apt epistemic performance. *Journal of the Learning Sciences*, *27*(3), 353–389. https://doi.org/10.1080/10508406.2017.1392968

Barzilai, S., & Eshet-Alkalai, Y. (2015). The role of epistemic perspectives in comprehension of multiple author viewpoints. *Learning and Instruction*, *36*, 86–103.

Barzilai, S., & Ka'adan, I. (2017). Learning to integrate divergent information sources: The interplay of epistemic cognition and epistemic metacognition. *Metacognition and Learning, 12*(2), 193–232.

Barzilai, S., Mor-Hagani, S., Zohar, A. R., Shlomi-Elooz, T., & Ben-Yishai, R. (2020). Making sources visible: Promoting multiple document literacy with digital epistemic scaffolds. *Computers & Education, 157*, 103980. https://doi.org/10.1016/j.compedu.2020.103980

Barzilai, S., Thomm, E., & Shlomi-Elooz, T. (2020). Dealing with disagreement: The roles of topic familiarity and disagreement explanation in evaluation of conflicting expert claims and sources. *Learning and Instruction, 69*, 101367. https://doi.org/10.1016/j.learninstruc.2020.101367

Barzilai, S., Tzadok, E., & Eshet-Alkalai, Y. (2015). Sourcing while reading divergent expert accounts: Pathways from views of knowing to written argumentation. *Instructional Science, 43*(6), 737–766.

Barzilai, S., & Weinstock, M. (2015). Measuring epistemic thinking within and across topics: A scenario-based approach. *Contemporary Educational Psychology, 42,* 141–158. https://doi.org/10.1016/j.cedpsych.2015.06.006

Barzilai, S., & Zohar, A. (2012). Epistemic thinking in action: Evaluating and integrating online sources. *Cognition and Instruction, 30*(1), 39–85. https://doi.org/10.1080/07370008.2011.636495

Barzilai, S., & Zohar, A. (2014). Reconsidering personal epistemology as metacognition: A multifaceted approach to the analysis of epistemic thinking. *Educational Psychologist, 49*(1), 13–35. https://doi.org/10.1080/00461520.2013.863265

Barzilai, S., & Zohar, A. (2016). Epistemic (meta) cognition: Ways of thinking about knowledge and knowing. In J. A. Greene, W. A. Sandoval, & I. Bråten (Eds.), *Handbook of epistemic cognition* (pp. 409–424). Routledge.

Barzilai, S., Zohar, A. R., & Mor-Hagani, S. (2018). Promoting integration of multiple texts: A review of instructional approaches and practices. *Educational Psychology Review*, 1–27.

Batchelor, O. (2017). Getting out the truth: The role of libraries in the fight against fake news. *Reference Services Review,* 143-148.

Bates, B. R., Romina, S., Ahmed, R., & Hopson, D. (2006). The effect of source credibility on consumers' perceptions of the quality of health information on the Internet. *Medical Informatics and the Internet in Medicine, 31*(1), 45–52. https://doi.org/10.1080/14639230600552601

Baxter Magolda, M. B. (2004). Evolution of a constructivist conceptualization of epistemological reflection. *Educational Psychologist, 39*(1), 31–42.

Belenky, M. F., Clinchy, B. M., Goldberger, N. R., Tarule, J. M., & others. (1986). *Women's ways of knowing: The development of self, voice, and mind* (Vol. 15). Basic books.

Bendixen, L. D. (2002). A process model of epistemic belief change. In B. K. Hofer & P. R. Pintrich (Eds.), *Personal epistemology: The psychology of beliefs about knowledge and knowing* (pp. 191–208). Lawrence Erlbaum Associates Publishers.

Bendixen, L. D., & Rule, D. C. (2004). An integrative approach to personal epistemology: A guiding model. *Educational Psychologist, 39*(1), 69–80.

Blakeslee, S. (2004). The CRAAP test. *Loex Quarterly, 31*(3), 4.

Braasch, J. L. G., Bråten, I., Strømsø, H. I., Anmarkrud, Ø., & Ferguson, L. E. (2013). Promoting secondary school students' evaluation of source features of multiple documents. *Contemporary Educational Psychology, 38*(3), 180–195. https://doi.org/10.1016/j.cedpsych.2013.03.003

Brand, A., Allen, L., Altman, M., Hlava, M., & Scott, J. (2015). Beyond authorship: attribution, contribution, collaboration, and credit. *Learned Publishing*, *28*(2), 151-155.

Brante, E. W., & Strømsø, H. I. (2018). Sourcing in text comprehension: A review of interventions targeting sourcing skills. *Educational Psychology Review, 30*(3), 773–799.

Bråten, I., Brante, E. W., & Strømsø, H. I. (2019). Teaching sourcing in upper secondary school: A comprehensive sourcing intervention with follow-up data. *Reading Research Quarterly, 54*(4), 481–505.

Bråten, I., Ferguson, L. E., Anmarkrud, Ø., & Strømsø, H. I. (2013). Prediction of learning and comprehension when adolescents read multiple texts: The roles of word-level processing, strategic approach, and reading motivation. *Reading and Writing, 26*(3), 321–348.

Bråten, I., Strømsø, H. I., & Andreassen, R. (2016). Sourcing in professional education: Do text factors make any difference? *Reading and Writing, 29*(8), 1599–1628.

Bråten, I., Strømsø, H. I., & Britt, M. A. (2009). Trust matters: Examining the role of source evaluation in students' construction of meaning within and across multiple texts. *Reading Research Quarterly, 44*(1), 6–28.

Bråten, I., Strømsø, H. I., & Samuelstuen, M. S. (2005). The relationship between internet-specific epistemological beliefs and learning within internet technologies. *Journal of Educational Computing Research, 33*(2), 141–171. https://doi.org/10.2190/E763-X0LN-6NMF-CB86

Breakstone, J., McGrew, S., Smith, M., Ortega, T., & Wineburg, S. (2018). Why we need a new approach to teaching digital literacy. *Phi Delta Kappan, 99*(6), 27–32.

Britt, M. A., & Aglinskas, C. (2002). Improving students' ability to identify and use source information. *Cognition and Instruction, 20*(4), 485–522.

Britt, M. A., Rouet, J.-F., & Braasch, J. L. (2013). Documents as entities. In A. Britt, S. Goldman, & J.-F. Rouet (Eds.), *Reading–from words to multiple texts* (pp. 160-179). Routledge.

Bromme, R., Kienhues, D., & Porsch, T. (2010). Who knows what and who can we believe? Epistemological beliefs are beliefs about knowledge (mostly) to be attained from others. In L. D. Bendixen & F. C. Feucht (Eds.), *Personal epistemology in the classroom: Theory, research, and implications for practice* (pp. 163-194). Cambridge University Press.

Camiling, M. K. S. (2019). Ehealth literacy of high school students in the Philippines. *IAFOR Journal of Education, 7*(2), 69–87. https://doi.org/10.22492/ije.7.2.04

Canadian Mental Health Association (CMHA). (2018a). *Stigma and discrimination*. https://ontario.cmha.ca/documents/stigma-and-discrimination/

CMHA. (2018b). *Myths about mental illness*. https://cmha.ca/documents/myths-about-mental-illness

Cartiff, B. M., Duke, R. F., & Greene, J. A. (2020). The effect of epistemic cognition interventions on academic achievement: A meta-analysis. *Journal of Educational Psychology*. https://doi.org/10.1037/edu0000490

Chandler, M. J., Hallett, D., & Sokol, B. W. (2002). Competing claims about competing knowledge claims. In B. K. Hofer & P. R. Pintrich (Eds.), *Personal epistemology: The psychology of beliefs about knowledge and knowing* (pp. 145–168). Lawrence Erlbaum Associates Publishers.

Chandler, M., Boyes, M., & Ball, L. (1990). Relativism and stations of epistemic doubt. *Journal of Experimental Child Psychology, 50*(3), 370–395.

Cheng, Y.-H., & Tsai, C.-C. (2020). Online research behaviors of engineering graduate students in Taiwan. *Journal of Educational Technology & Society, 20*(1), 169–179.

Chevrier, M., Muis., K. R., Pekrun, R., Sinatra, G., & Denton, C. A. (2019). *Thinking critically about socio-scientific issues: The roles of epistemic cognition and epistemic emotions* [Unpublished manuscript]. Department of Educational and Counselling Psychology, McGill University.

Chiagouris, L., Long, M. M., & Plank, R. E. (2008). The consumption of online news: The relationship of attitudes toward the site and Credibility. *Journal of Internet Commerce, 7*(4), 528–549. https://doi.org/10.1080/15332860802507396

Chinn, C. A., Barzilai, S., & Duncan, R. G. (2020). Education for a "post-truth" world: New directions for research and practice. *Educational Researcher,* 0013189X2094068. https://doi.org/10.3102/0013189X20940683

Chinn, C. A., Buckland, L. A., & Samarapungavan, A. (2011). Expanding the dimensions of epistemic cognition: Arguments from philosophy and psychology. *Educational Psychologist, 46*(3), 141–167. https://doi.org/10.1080/00461520.2011.587722

Chinn, C. A., & Rinehart, R. W. (2016). Epistemic cognition and philosophy: Developing a new framework for epistemic cognition. In J. A. Greene, W. A. Sandoval, & I. Bråten (Eds.), *Handbook of epistemic cognition* (pp. 460–478). Routledge.

Chinn, C. A., Rinehart, R. W., & Buckland, L. A. (2014). Epistemic cognition and evaluating information: Applying the AIR model of epistemic cognition. In D. N. Rapp & J. L.G. Braasch (Eds.), *Processing inaccurate information: Theoretical and applied perspectives from cognitive science and the educational sciences* (pp. 425–453). MIT Press.

Chinn, C. A., & Sandoval, W. A. (2018). Epistemic cognition and epistemic development. In F. Fischer, C. E. Hmelo-Silver, S. R. Goldman, & P. Resiman (Eds.), *International handbook of the learning sciences* (pp. 46–55). Routledge.

Cho, B.-Y., Woodward, L., & Li, D. (2018). Epistemic processing when adolescents read online: A verbal protocol analysis of more and less successful online readers. *Reading Research Quarterly, 53*(2), 197–221.

Clark, J. C., Rawson, C., Hobbs, L., Oughtred, C., Hayes, K., Kelly, L., & Higgins, J. (2017). A pedagogy for epistemic agency in the judgment of accuracy and reliability. *Qwerty-Open and Interdisciplinary Journal of Technology, Culture and Education, 11*(2), 27–47.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2013). *Applied multiple regression/correlation analysis for the behavioral sciences*. Routledge.

Cresswell, J. W., & Plano Clark, V. L. (2011). *Designing and conducting mixed methods research* (2nd ed.). Sage Publications.

Cresswell, J. W., & Plano Clark, V. L. (2017). *Designing and conducting mixed methods research* (3rd ed.). Sage Publications.

Cugelman, B., Thelwall, M., & Dawes, P. (2009). The dimensions of web site credibility and their relation to active trust and behavioural impact. *Communications of the Association for Information Systems, 24*. https://doi.org/10.17705/1CAIS.02426

Cunningham, A., & Johnson, F. (2016). Exploring trust in online health information: A study of user experiences of patients.co.uk. *Health Information & Libraries Journal, 33*(4), 323–328. https://doi.org/10.1111/hir.12163

De La Paz, S., & Felton, M. K. (2010). Reading and writing from multiple source documents in history: Effects of strategy instruction with low to average high school writers. *Contemporary Educational Psychology, 35*(3), 174–192. https://doi.org/10.1016/j.cedpsych.2010.03.001

DeBacker, T. K., Crowson, H. M., Beesley, A. D., Thoma, S. J., & Hestevold, N. L. (2008). The Challenge of measuring epistemic beliefs: An analysis of three self-report instruments. *The Journal of Experimental Education, 76*(3), 281–312. https://doi.org/10.3200/JEXE.76.3.281-314

Denton, C. A., Muis, K.R., Armstrong, S., & Dubé, A. (in preparation). En-garde: Source evaluations in the digital age. *Contemporary Educational Psychology.*

Dochterman, M. A., & Stamp, G. H. (2010). The determination of Web credibility: A thematic analysis of Web user's judgments. *Qualitative Research Reports in Communication, 11*(1), 37–43. https://doi.org/10.1080/17459430903514791

Duby, H. R. (2018). What a load of CRAAP: Evaluating information in an era of "fake news". *Tennessee Libraries*, 68(4).

Eastin, M. S., Yang, M.-S., & Nathanson, A. I. (2006). Children of the net: An empirical exploration into the evaluation of Internet content. *Journal of Broadcasting & Electronic Media, 50*(2), 211–230. https://doi.org/10.1207/s15506878jobem5002_3

Elby, A., & Hammer, D. (2010). Epistemological resources and framing: A cognitive framework for helping teachers interpret and respond to their students' epistemologies. I In L. D. Bendixen & F. C. Feucht (Eds.), *Personal epistemology in the classroom: Theory, research, and implications for practice* (Vol. 4, pp. 409–434). Cambridge University Press.

Eshet-Alkalai, Y., & Chajut, E. (2009). Changes over time in digital literacy. *CyberPsychology & Behavior*, 12(6), 713–715.

Eshet-Alkalai, Y., & Chajut, E. (2010). You can teach old dogs new tricks: The factors that affect changes over time in digital literacy. *Journal of Information Technology Education: Research, 9*(1), 173–181.

Eshet, Y. (2012). Thinking in the digital era: A revised model for digital literacy. *Issues in Informing Science and Information Technology, 9*(2), 267–276.

Eysenbach, G. (2002). How do consumers search for and appraise health information on the world wide web? Qualitative study using focus groups, usability tests, and in-depth interviews. *BMJ, 324*(7337), 573–577. https://doi.org/10.1136/bmj.324.7337.573

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Ferguson, L. E., Bråten, I., & Strømsø, H. I. (2012). Epistemic cognition when students read multiple documents containing conflicting scientific evidence: A think-aloud study. *Learning and Instruction, 22*(2), 103–120.

Fielding, J. A. (2019). Rethinking CRAAP: Getting students thinking like fact-checkers in evaluating web sources. *College & Research Libraries News, 80*(11), 620.

Flanagin, A. J., & Metzger, M. J. (2007). The role of site features, user attributes, and information verification behaviors on the perceived credibility of web-based information. *New Media & Society, 9*(2), 319–342. https://doi.org/10.1177/1461444807075015

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American Psychologist, 34*(10), 906.

Flavell, J. H., Miller, P. H., & Miller, S. A. (2002). *Cognitive development* (4th ed., Vol. 338). Prentice-Hall Englewood Cliffs, NJ.

Gerjets, P., Kammerer, Y., & Werner, B. (2011). Measuring spontaneous and instructed evaluation processes during Web search: Integrating concurrent thinking-aloud protocols and eye-tracking data. *Learning and Instruction, 21*(2), 220–231. https://doi.org/10.1016/j.learninstruc.2010.02.005

Glaser, B.G. (1978). *Theoretical sensitivity: Advances in the methodology of grounded theory*. The Sociology Press.

Glaser, Barney G, & Strauss, A. L. (2017). *Discovery of grounded theory: Strategies for qualitative research*. Routledge.

Green, J. (2019, Jan 22). *Check yourself with lateral reading: Crash course navigating digital information #3* [Youtube video]. https://www.youtube.com/watch?v=GoQG6Tin-1E&t=157s.

Greene, J. A. (2016). Interacting epistemic systems within and beyond the classroom. In J. A. Greene, W. A. Sandoval, & I. Bråten (Eds.), *Handbook of epistemic cognition* (pp. 265–278). Routledge.

Greene, J. A., Azevedo, R., & Torney-Purta, J. (2008). Modeling epistemic and ontological cognition: Philosophical perspectives and methodological directions. *Educational Psychologist, 43*(3), 142–160. https://doi.org/10.1080/00461520802178458

Greene, J. A., Cartiff, B. M., & Duke, R. F. (2018). A meta-analytic review of the relationship between epistemic cognition and academic achievement. *Journal of Educational Psychology, 110*(8),1084-1111. https://doi.org/10.1037/edu0000263

Greene, J. A., Copeland, D. Z., Deekens, V. M., & Seung, B. Y. (2018). Beyond knowledge: Examining digital literacy's role in the acquisition of understanding in science. *Computers & Education, 117*, 141–159.

Greene, J. A., Sandoval, W. A., & Bråten, I. (2016). *Handbook of epistemic cognition*. Routledge.

Greene, J. A., & Yu, S. B. (2016). Educating critical thinkers: The role of epistemic cognition. *Policy Insights from the Behavioral and Brain Sciences, 3*(1), 45–53. https://doi.org/10.1177/2372732215622223

Greene, J. A., Yu, S. B., & Copeland, D. Z. (2014). Measuring critical components of digital literacy and their relationships with learning. *Computers & Education, 76*, 55–69. https://doi.org/10.1016/j.compedu.2014.03.008

Greeno, J. G. (2015). Commentary: Some prospects for connecting concepts and methods of individual cognition and of situativity. *Educational Psychologist, 50*(3), 248–251. https://doi.org/10.1080/00461520.2015.1077708

Greer, J. D. (2003). Evaluating the credibility of online information: A test of source and advertising influence. *Mass Communication and Society, 6*(1), 11–28. https://doi.org/10.1207/S15327825MCS0601_3

Gui, M., & Argentin, G. (2011). Digital skills of internet natives: Different forms of digital literacy in a random sample of northern Italian high school students. *New Media & Society, 13*(6), 963–980.

Halverson, K. L., Siegel, M. A., & Freyermuth, S. K. (2010). Non-science majors' critical evaluation of websites in a biotechnology course. *Journal of Science Education and Technology, 19*(6), 612–620. https://doi.org/10.1007/s10956-010-9227-6

Hämäläinen, E. K., Kiili, C., Marttunen, M., Räikkönen, E., González-Ibáñez, R., & Leppänen, P. H. T. (2020). Promoting sixth graders' credibility evaluation of Web pages: An intervention study. *Computers in Human Behavior, 110*, 106372. https://doi.org/10.1016/j.chb.2020.106372

Hammer, D., & Elby, A. (2002). On the form of a personal epistemology. In B. K. Hofer and P. R. Pintrich (Eds.), *Personal epistemology: The psychology of beliefs about knowledge and knowing* (pp. 169–190). Lawrence Erlbaum Associates Publishers.

Harris, P. R., Sillence, E., & Briggs, P. (2009). The effect of credibility-related design cues on responses to a Web-based message about the breast cancer risks from alcohol: Randomized controlled trial. *Journal of Medical Internet Research, 11*(3), e37. https://doi.org/10.2196/jmir.1097

Hart, W., Albarracín, D., Eagly, A. H., Brechan, I., Lindberg, M. J., & Merrill, L. (2009). Feeling validated versus being correct: A meta-analysis of selective exposure to information. *Psychological Bulletin, 135*(4), 555.

Hirsh, S. G. (1999). Children's relevance criteria and information seeking on electronic resources. *Journal of the American Society for Information Science, 50*(14), 1265–1283.

Hofer, B. K. (2000). Dimensionality and disciplinary differences in personal epistemology. *Contemporary Educational Psychology, 25*(4), 378–405.

Hofer, B. K. (2001). Personal epistemology research: Implications for learning and teaching. *Educational Psychology Review, 13*(4), 353–383.

Hofer, B. K. (2004). Epistemological understanding as a metacognitive process: Thinking aloud during online searching. *Educational Psychologist, 39*(1), 43–55.

Hofer, B. K. (2016). Epistemic cognition as a psychological construct: Advancements and challenges. In J. A. Greene, W. A. Sandoval, & I. Bråten (Eds.), *Handbook of epistemic cognition* (pp. 31–50). Routledge.

Hofer, B. K., & Pintrich, P. R. (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research, 67*(1), 88–140.

Hootsuite & We Are Social (2020). *Digital 2020*. Hootsuite. https://hootsuite.com/resources/digital-2020

Iding, M., & Klemm, E. B. (2005). Pre-service teachers critically evaluate scientific information on the World Wide Web. *Computers in the Schools, 22*(1–2), 7–18. https://doi.org/10.1300/J025v22n01_02

Iordanou, K., Muis, K. R., & Kendeou, P. (2018). Epistemic perspective and online epistemic processing of evidence: Developmental and domain differences. *The Journal of Experimental Education, 84*(4), 531-551. https://doi.org/10.1080/00220973.2018.1482857

Jenkins, H. (2009). *Confronting the challenges of participatory culture: Media education for the 21st century*. MIT Press.

Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher, 33*(7), 14–26. https://doi.org/10.3102/0013189X033007014

Jung, E. H., Walsh-Childers, K., & Kim, H.-S. (2016). Factors influencing the perceived credibility of diet-nutrition information web sites. *Computers in Human Behavior, 58*, 37–47. https://doi.org/10.1016/j.chb.2015.11.044

Jung, W. S., Chung, M.-Y., & Rhee, E. S. (2018). The Effects of Attractiveness and Source Expertise on Online Health Sites. *Health Communication, 33*(8), 962–971. https://doi.org/10.1080/10410236.2017.1323364

Kąkol, M., & Nielek, R. (2015). What affects web credibility perception? An analysis of textual justifications. *Computer Science, 16*(3), 295-310. http://dx.doi.org/10.7494/csci.2015.16.3.295

Kąkol, M., Nielek, R., & Wierzbicki, A. (2017). Understanding and predicting Web content credibility using the Content Credibility Corpus. *Information Processing & Management, 53*(5), 1043–1061. https://doi.org/10.1016/j.ipm.2017.04.003

Kalichman, S. C., Cherry, C., Cain, D., Weinhardt, L. S., Benotsch, E., Pope, H., & Kalichman, M. (2006). Health information on the Internet and people living with HIV/AIDS:

Information evaluation and coping styles. *Health Psychology, 25*(2), 205–210. https://doi.org/10.1037/0278-6133.25.2.205

Kammerer, Y., Amann, D. G., & Gerjets, P. (2015). When adults without university education search the Internet for health information: The roles of Internet-specific epistemic beliefs and a source evaluation intervention. *Computers in Human Behavior, 48*, 297–309. https://doi.org/10.1016/j.chb.2015.01.045

Kammerer, Y., Bråten, I., Gerjets, P., & Strømsø, H. I. (2013). The role of Internet-specific epistemic beliefs in laypersons' source evaluations and decisions during Web search on a medical issue. *Computers in Human Behavior, 29*(3), 1193–1203. https://doi.org/10.1016/j.chb.2012.10.012

Kammerer, Y., Kalbfell, E., & Gerjets, P. (2016). Is this information source commercially biased? How contradictions between web pages stimulate the consideration of source information. *Discourse Processes, 53*(5–6), 430–456. https://doi.org/10.1080/0163853X.2016.1169968

Kienhues, D., Bromme, R., & Stahl, E. (2008). Changing epistemological beliefs: The unexpected impact of a short-term intervention. *British Journal of Educational Psychology, 78*(4), 545–565.

Kienhues, D., Ferguson, L., & Stahl, E. (2016). Diverging information and epistemic change. In J. A. Greene, W. A. Sandoval, & I. Bråten (Eds.), *Handbook of epistemic cognition* (pp. 318–330). Routledge.

Kiili, C., Laurinen, L., & Marttunen, M. (2008). Students evaluating Internet sources: From versatile evaluators to uncritical readers. *Journal of Educational Computing Research, 39*(1), 75–95. https://doi.org/10.2190/EC.39.1.e

Kim, M. (2015). Partisans and controversial news online: Comparing perceptions of bias and credibility in news content from blogs and mainstream media. *Mass Communication and Society, 18*(1), 17–36. https://doi.org/10.1080/15205436.2013.877486

King, P. M., & Kitchener, K. S. (1994). *Developing reflective judgment: Understanding and promoting intellectual growth and critical thinking in adolescents and adults*. Jossey-Bass.

Kiousis, S., & Dimitrova, D. V. (2006). Differential impact of Web site content: Exploring the influence of source (public relations versus news), modality, and participation on college students' perceptions. *Public Relations Review*, *32*(2), 177–179. https://doi.org/10.1016/j.pubrev.2006.02.013

Kitchener, R. F. (2011). Personal epistemology and philosophical epistemology: The view of a philosopher. In J. Elen, E. Stahl, R. Bromme, & G. Clarebout (Eds.), *Links between beliefs and cognitive flexibility* (pp. 79–103). Springer.

Kitchner, K. S. (1983). Cognition, metacognition, and epistemic cognition. *Human Development, 26*(4), 222–232.

Knight, S., Rienties, B., Littleton, K., Mitsui, M., Tempelaar, D., & Shah, C. (2017). The relationship of (perceived) epistemic cognition to interaction with resources on the internet. *Computers in Human Behavior, 73*, 507–518. https://doi.org/10.1016/j.chb.2017.04.014

Kobau, R., DiIorio, C., Chapman, D., Delvecchio, P., & SAMHSA/CDC Mental Illness Stigma Panel Members. (2010). Attitudes about mental illness and its treatment: Validation of a generic scale for public health surveillance of mental illness associated stigma. *Community Mental Health Journal, 46*(2), 164–176. https://doi.org/10.1007/s10597-009-9191-x

Kouzy, R., Abi Jaoude, J., Kraitem, A., El Alam, M. B., Karam, B., Adib, E., Zarka, J., Traboulsi, C., Akl, E., & Baddour, K. (2020). Coronavirus goes viral: Quantifying the COVID-19 misinformation epidemic on Twitter. *Cureus*. https://doi.org/10.7759/cureus.7255

Krouwer, S., Poels, K., & Paulussen, S. (2020). Moving towards transparency for native advertisements on news websites: A test of more detailed disclosures. *International Journal of Advertising, 39*(1), 51–73. https://doi.org/10.1080/02650487.2019.1575107

Kuhn, D. (1991). *The skills of argument*. Cambridge University Press.

Kuhn, D. (1999). A developmental model of critical thinking. *Educational Researcher, 28*(2), 16–46.

Kuhn, D., & Weinstock, M. (2002). What is epistemological thinking and why does it matter? In B. K. Hofer & P. R. Pintrich (Eds.), *Personal epistemology: The psychology of beliefs about knowledge and knowing* (pp. 121–144). Lawrence Erlbaum Associates Publishers.

Kuhn, D., Cheney, R., & Weinstock, M. (2000). The development of epistemological understanding. *Cognitive Development, 15*, 309–328. https://doi.org/10.1016/S0885-2014(00)00030-7

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159. https://doi.org/10.2307/2529310

Leu, D. J., Kinzer, C. K., Coiro, J., Castek, J., & Henry, L. A. (2017). New literacies: A dual-level theory of the changing nature of literacy, instruction, and assessment. *Journal of Education, 197*(2), 1–18.

Liao, Q. V., & Fu, W.-T. (2014). Age differences in credibility judgments of online health information. *ACM Transactions on Computer-Human Interaction, 21*(1), 1–23. https://doi.org/10.1145/2534410

Lombardi, D., Seyranian, V., & Sinatra, G. M. (2014). Source effects and plausibility judgments when reading About climate change. *Discourse Processes, 51*(1–2), 75–92. https://doi.org/10.1080/0163853X.2013.855049

Louca, L., Elby, A., Hammer, D., & Kagey, T. (2004). Epistemological resources: Applying a new epistemological framework to science instruction. *Educational Psychologist, 39*(1), 57–68. https://doi.org/10.1207/s15326985ep3901_6

Lucassen, T., & Schraagen, J. M. (2011). Factual accuracy and trust in information: The role of expertise. *Journal of the American Society for Information Science and Technology, 62*(7), 1232–1242.

Lucassen, T., & Schraagen, J. M. (2013). The influence of source cues and topic familiarity on credibility evaluation. *Computers in Human Behavior, 29*(4), 1387–1392. https://doi.org/10.1016/j.chb.2013.01.036

Luttrell, W. (2010). *Qualitative educational research: Readings in reflexive methodology and transformative practice*. Routledge.

Machackova, H., & Smahel, D. (2018). The perceived importance of credibility cues for the assessment of the trustworthiness of online information by visitors of health-related websites: The role of individual factors. *Telematics and Informatics, 35*(5), 1534–1541. https://doi.org/10.1016/j.tele.2018.03.021

Madden, A. D., Ford, N., Gorrell, G., Eaglestone, B., & Holdridge, P. (2012). Metacognition and web credibility. *The Electronic Library, 30*(5), 671–689. https://doi.org/10.1108/02640471211275710

Maier, J., & Richter, T. (2013). How nonexperts understand conflicting information on social science issues. *Journal of Media Psychology, 25*(1), 14-26. https://doi.org/10.1027/1864-1105/a000078

Maier, J., & Richter, T. (2016). Effects of text-belief consistency and reading task on the strategic validation of multiple texts. *European Journal of Psychology of Education, 31*(4), 479–497.

Mason, L., Ariasi, N., & Boldrin, A. (2011). Epistemic beliefs in action: Spontaneous reflections about knowledge and knowing during online information searching and their influence on learning. *Learning and Instruction, 21*(1), 137–151. https://doi.org/10.1016/j.learninstruc.2010.01.001

Mason, L., Boldrin, A., & Ariasi, N. (2010). Epistemic metacognition in context: Evaluating and learning online information. *Metacognition and Learning, 5*(1), 67–90. https://doi.org/10.1007/s11409-009-9048-2

Mason, L., Junyent, A. A., & Tornatora, M. C. (2014). Epistemic evaluation and comprehension of web-source information on controversial science-related topics: Effects of a short-term instructional intervention. *Computers & Education, 76*, 143–157. https://doi.org/10.1016/j.compedu.2014.03.016

Mason, L., Scrimin, S., Tornatora, M. C., Suitner, C., & Moè, A. (2018). Internet source evaluation: The role of implicit associations and psychophysiological self-regulation. *Computers & Education, 119*, 59–75. https://doi.org/10.1016/j.compedu.2017.12.009

McGrew, S. (2020). Learning to evaluate: An intervention in civic online reasoning. *Computers & Education, 145*, 103711. https://doi.org/10.1016/j.compedu.2019.103711

McGrew, S., Smith, M., Breakstone, J., Ortega, T., & Wineburg, S. (2019). Improving university students' web savvy: An intervention study. *British Journal of Educational Psychology, 89*(3), 485-500. https://doi.org/10.1111/bjep.12279

Menchen-Trevino, E., & Hargittai, E. (2011). Young adults' credibility assessment of Wikipedia. *Information, Communication & Society, 14*(1), 24–51.

Meppelink, C. S., Smit, E. G., Fransen, M. L., & Diviani, N. (2019). "I was right about vaccination": Confirmation bias and health literacy in online health information seeking. *Journal of Health Communication, 24*(2), 129–140. https://doi.org/10.1080/10810730.2019.1583701

Meriam Library of CSU Chico. (2010). *Evaluating information - Applying the CRAAP test* [PDF document]. Meriam Library of CSU Chico. https://library.csuchico.edu/sites/default/files/craap-test.pdf

Merk, S., Rosman, T., Muis, K. R., Kelava, A., & Bohl, T. (2018). Topic specific epistemic beliefs: Extending the Theory of Integrated Domains in Personal Epistemology. *Learning and Instruction, 56,* 84–97.

Merk, S., Rosman, T., Rueß, J., Syring, M., & Schneider, J. (2017). Pre-service teachers' perceived value of general pedagogical knowledge for practice: Relations with epistemic beliefs and source beliefs. *Plos One, 12*(9), e0184971.

Merk, S., Schneider, J., Bohl, T., Kelava, A., & Syring, M. (2017). Teacher students' epistemic beliefs about general pedagogical knowledge: Topic-, source-and context specificity. *Journal for Educational Research Online/Journal Für Bildungsforschung Online, 9*(1), 169–189.

Meyer, H. K., Marchionni, D., & Thorson, E. (2010). The journalist behind the news: Credibility of straight, collaborative, opinionated, and blogged "news." *American Behavioral Scientist, 54*(2), 100–119. https://doi.org/10.1177/0002764210376313

Meyers, L. S., Gamst, G., & Guarino, A. J. (2016). *Applied multivariate research: Design and interpretation*. Sage Publications.

Ministère de l'Enseignement supérieur (MEES). (2018). *Digital action plan for education and higher education* [PDF document]. http://www.education.gouv.qc.ca/fileadmin/site_web/documents/ministere/PAN_Plan_action_VA.pdf

Morgan, D. L. (1988). *Focus groups as qualitative research* (Vol. 16). Sage Publications.

Morse, J. M. (2003). Principles of mixed methods and multimethod research design. In A. Tashakkori, & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 189–208). Sage Publications.

Muis, K. R. (2007). The role of epistemic beliefs in self-regulated learning. *Educational Psychologist, 42*(3), 173–190. https://doi.org/10.1080/00461520701416306

Muis, K. R., Bendixen, L. D., & Haerle, F. C. (2006). Domain-generality and domain-specificity in personal epistemology research: Philosophical and empirical reflections in the development of a theoretical framework. *Educational Psychology Review, 18*(1), 3–54.

Muis, K. R., Chevrier, M., & Singh, C. (2018). The role of epistemic emotions in personal epistemology and self-regulated learning. *Educational Psychologist*, 1–20.

Muis, K. R., & Duffy, M. C. (2013). Epistemic climate and epistemic change: Instruction designed to change students' beliefs and learning strategies and improve achievement. *Journal of Educational Psychology, 105*(1), 213.

Muis, K. R., Etoubashi, N., & Denton, C. A. (2020). The catcher in the lie: The role of emotions and epistemic judgments in changing students' misconceptions and attitudes in a post-truth era. *Contemporary Educational Psychology, 62*, 101898. https://doi.org/10.1016/j.cedpsych.2020.101898

Muis, K. R., Pekrun, R., Sinatra, G. M., Azevedo, R., Trevors, G., Meier, E., & Heddy, B. C. (2015). The curious case of climate change: Testing a theoretical model of epistemic beliefs, epistemic emotions, and complex learning. *Learning and Instruction, 39*, 168–183.

Muis, K. R., & Singh, C. (2018). The three faces of epistemic thinking in self-regulated learning. In D. H. Schunk & J. A. Greene (Eds.), *Handbook of self-regulated of learning and performance* (pp. 434–456). Routledge/Taylor & Francis Group.

Muis, K. R., Trevors, G., & Chevrier, M. (2016). Epistemic climate for epistemic change. In J. A. Greene, W. A. Sandoval, & I. Bråten (Eds.), *Handbook of epistemic cognition* (pp. 331–359). Routledge.

Murphy, P. K., Alexander, P. A., Greene, J. A., & Edwards, M. N. (2007). Epistemological threads in the fabric of conceptual change research. In S. Vosniadou, A. Baltas, & X. Vamvakoussi (Eds.), *Advances in learning and instruction series. Reframing the conceptual change approach in learning and instruction* (p. 105–122). Elsevier Science.

Murphy, P., Alexander, P., & Muis, K. R. (2012). Knowledge and knowing: The journey from philosophy and psychology to human learning. *Educational Psychology Handbook, 1*, 189–226.

Naab, T. K., Heinbach, D., Ziegele, M., & Grasberger, M.-T. (2020). Comments and credibility: How critical user comments decrease perceived news article credibility. *Journalism Studies, 21*(6), 783–801.

Nokes, J. D., Dole, J. A., & Hacker, D. J. (2007). Teaching high school students to use heuristics while reading historical texts. *Journal of Educational Psychology, 99*(3), 492–504. https://doi.org/10.1037/0022-0663.99.3.492

Nunkoosing, K. (2005). The problems with interviews. *Qualitative Health Research, 15*(5), 698–706.

Perfetti, C. A., Rouet, J.-F., & Britt, M. A. (1999). Toward a theory of documents representation. In H. van Oostendorp & S. R. Goldman (Eds.), *The construction of mental representations during reading* (p. 99–122). Lawrence Erlbaum Associates Publishers

Perry, W. G. (1970). *Forms of intellectual and ethical development in the college years: A scheme*. Holt, Rinehart and Winston.

Press, S. J., & Wilson, S. (1978). Choosing between logistic regression and discriminant analysis. *Journal of the American Statistical Association, 73*(364), 699–705.

Pretorius, L. (2018). Experiential and self-discovery learning in digital literacy: Developing the discernment to evaluate source reliability. *College & Undergraduate Libraries, 25*(4), 388–405. https://doi.org/10.1080/10691316.2018.1530626

Queen's University Library. (2020). *CRAAP criteria*. Bachelor of health sciences toolbox. https://guides.library.queensu.ca/BHSc/CRAAP

Rains, S. A., & Karmikel, C. D. (2009). Health information-seeking and perceptions of website credibility: Examining Web-use orientation, message characteristics, and structural features of websites. *Computers in Human Behavior, 25*(2), 544–553. https://doi.org/10.1016/j.chb.2008.11.005

Ricco, R., Schuyten Pierce, S., & Medinilla, C. (2010). Epistemic beliefs and achievement motivation in early adolescence. *The Journal of Early Adolescence, 30*(2), 305–340.

Richter, T., & Maier, J. (2017). Comprehension of multiple documents with conflicting information: A two-step model of validation. *Educational Psychologist, 52*(3), 148–166. https://doi.org/10.1080/00461520.2017.1322968

Richter, T., & Schmid, S. (2010). Epistemological beliefs and epistemic strategies in self-regulated learning. *Metacognition and Learning, 5*(1), 47–65.

Robertson-Lang, L., Major, S., & Hemming, H. (2011). An exploration of search patterns and credibility Issues among older adults seeking online health information. *Canadian Journal on Aging / La Revue Canadienne Du Vieillissement, 30*(4), 631–645. https://doi.org/10.1017/S071498081100050X

Rouet, J.-F., & Britt, M. A. (2011). In M. T. McCrudden, J. P. Magliano, & G. Schraw (Eds.), *Text relevance and learning from text* (p. 19–52). IAP Information Age Publishing.

Rouet, J.-F., Britt, M. A., & Durik, A. M. (2017). RESOLV: Readers' representation of reading contexts and tasks. *Educational Psychologist, 52*(3), 200–215. https://doi.org/10.1080/00461520.2017.1329015

Royce, J. (1978). Three ways of knowing and the scientific world view. *Methodology and Science, 11*, 146–164.

Royce, J. R. (1959). The search for meaning. *American Scientist, 47*(4), 515–535.

Rule, D. C., & Bendixen, L. D. (2010). The integrative model of personal epistemology development: Theoretical underpinnings and implications for education. In L. Bendixen & F. C. Feucht (Eds.), *Personal epistemology in the classroom: Theory, research, and implications for practice* (pp.94-123). Cambridge University Press.

Ryu, S., & Sandoval, W. A. (2012). Improvements to elementary children's epistemic understanding from sustained argumentation. *Science Education, 96*(3), 488–526.

Sabourin, J. L., Shores, L. R., Mott, B. W., & Lester, J. C. (2013). Understanding and predicting student self-regulated learning strategies in game-based learning environments. *International Journal of Artificial Intelligence in Education, 23*(1–4), 94–114.

Salmerón, L., Gil, L., & Bråten, I. (2018). Effects of reading real versus print-out versions of multiple documents on students' sourcing and integrated understanding. *Contemporary Educational Psychology, 52*, 25–35. https://doi.org/10.1016/j.cedpsych.2017.12.002

Sandoval, W. A. (2005). Understanding students' practical epistemologies and their influence on learning through inquiry. *Science Education, 89*(4), 634–656.

Sandoval, W. A. (2009). In defense of clarity in the study of personal epistemology. *The Journal of the Learning Sciences, 18*(1), 150–161. https://doi.org/10.1080/10508400802581700

Sandoval, W. A. (2012). Situating epistemological development. *International Society of the Learning Sciences*, 347 – 354.

Sandoval, W. A. (2014). Science education's need for a theory of epistemological development. *Science Education, 98*(3), 383–387.

Sandoval, W. A. (2017). Situating practices of epistemic cognition. In T. G. Amin & O. Levrini (Eds.), *Converging perspectives on conceptual change* (1st ed., pp. 253–260). Routledge. https://doi.org/10.4324/9781315467139-30

Sandoval, W. A., & Çam, A. (2011). Elementary children's judgments of the epistemic status of sources of justification. *Science Education, 95*(3), 383–408.

Sandoval, W. A., & Redman, E. H. (2015). The contextual nature of scientists' views of theories, experimentation, and their coordination. *Science & Education, 24*(9–10), 1079–1102.

Sandoval, W. A., Greene, J. A., & Bråten, I. (2016). Understanding and promoting thinking about knowledge: Origins, issues, and future directions of research on epistemic cognition.

Review of Research in Education, 40(1), 457–496. https://doi.org/10.3102/0091732X16669319

Schacter, J., Chung, G. K., & Dorr, A. (1998). Children's Internet searching on complex problems: Performance and process analyses. *Journal of the American Society for Information Science, 49*(9), 840–849.

Schommer-Aikins, M. (2004). Explaining the epistemological belief system: Introducing the embedded systemic model and coordinated research approach. *Educational Psychologist, 39*(1), 19–29.

Schommer, M. (1990). Effects of beliefs about the nature of knowledge on comprehension. *Journal of Educational Psychology, 82*(3), 498.

Sillence, E., Briggs, P., Harris, P. R., & Fishwick, L. (2007). How do patients evaluate and make use of online health information? *Social Science & Medicine, 64*(9), 1853–1862.

Sinatra, G. M. (2016). Thoughts on knowledge about thinking about knowledge. In J. A. Greene, W. A. Sandoval, & I. Bråten (Eds.), *Handbook of epistemic cognition* (pp. 479–491). Routledge.

Sinatra, G. M., & Chinn, C. A. (2012). Thinking and reasoning in science: Promoting epistemic conceptual change. In K. R. Harris, S. Graham, T. Urdan, A. G. Bus, S. Major, & H. L. Swanson (Eds.), *APA educational psychology handbook: Application to learning and teaching* (Vol. 3, pp. 257–282). American Psychological Association.

Smallwood, C. (2015). *The complete guide to using Google in libraries: Instruction, administration, and staff productivity* (Vol. 1). Rowman & Littlefield.

Song, H., Cho, J., & Benefield, G. A. (2020). The dynamics of message selection in online political discussion forums: Self-segregation or diverse exposure? *Communication Research, 47*(1), 125–152.

Sosa, E. (2015). *Judgment and agency*. Oxford University Press, USA.

Stadtler, M., Scharrer, L., Macedo-Rouet, M., Rouet, J.-F., & Bromme, R. (2016). Improving vocational students' consideration of source information when deciding about science controversies. *Reading and Writing, 29*(4), 705–729. https://doi.org/10.1007/s11145-016-9623-2

Stanford History Education Group. (2016). *Evaluating information: The cornerstone of online civic reasoning* [PDF document]. https://stacks.stanford.edu/file/druid:fv751yt5934/SHEG%20Evaluating%20Information%20Online.pdf

Stanford History Education Group. (2020). *COR*. Civic Online Reasoning. https://cor.stanford.edu

Starling, M. S., Cheshire, C., Deardorff, J., & Nuru-Jeter, A. (2018). Late adolescent user experiences with online sexual health resources: A qualitative study. *American Journal of Sexuality Education, 13*(4), 487–503. https://doi.org/10.1080/15546128.2018.1505572

Strømsø, H. I., & Bråten, I. (2010). The role of personal epistemology in the self-regulation of internet-based learning. *Metacognition and Learning, 5*(1), 91–111. https://doi.org/10.1007/s11409-009-9043-7

Strømsø, H. I., & Kammerer, Y. (2016). Epistemic cognition and reading for understanding in the internet age. In J. A. Greene, W. A. Sandoval, & I. Bråten (Eds.), *Handbook of epistemic cognition* (pp. 230–246). Routledge.

Subramaniam, M., Taylor, N. G., St. Jean, B., Follman, R., Kodama, C., & Casciotti, D. (2015). As simple as that?: Tween credibility assessment in a complex online world. *Journal of Documentation, 71*(3), 550–571. https://doi.org/10.1108/JD-03-2014-0049

Swasy, A., Tandoc, E., Bhandari, M., & Davis, R. (2015). Traditional reporting more credible than citizen news. *Newspaper Research Journal, 36*(2), 225–236. https://doi.org/10.1177/0739532915587298

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics: International edition*. Pearson.

Tandoc, E. C., & Thomas, R. J. (2017). Readers value objectivity over transparency. *Newspaper Research Journal, 38*(1), 32–45. https://doi.org/10.1177/0739532917698446

Taylor, W. D., Johnson, G., Ault, M. K., Griffith, J. A., Rozzell, B., Connelly, S., Jensen, M. L., Dunbar, N. E., & Ness, A. M. (2015). Ideological group persuasion: A within-person study of how violence, interactivity, and credibility features influence online persuasion. *Computers in Human Behavior, 51*, 448–460. https://doi.org/10.1016/j.chb.2015.04.071

The News Literacy Project. (2020). *Checkology.* Checkology. https://get.checkology.org

Thon, F. M., & Jucks, R. (2017). Believing in expertise: How authors' credentials and language use influence the credibility of online health information. *Health Communication, 32*(7), 828–836. https://doi.org/10.1080/10410236.2016.1172296

Treise, D., Walsh-Childers, K., Weigold, M. F., & Friedman, M. (2003). Cultivating the science Internet audience: Impact of brand and domain on source credibility for science information. *Science Communication, 24*(3), 309–332. https://doi.org/10.1177/1075547002250298

Trevors, G. J., Muis, K. R., Pekrun, R., Sinatra, G. M., & Muijselaar, M. M. (2017). Exploring the relations between epistemic beliefs, emotions, and learning from texts. *Contemporary Educational Psychology, 48,* 116–132. https://doi.org/10.1016/j.cedpsych.2016.10.001

Tsai, C.-C. (2004). Beyond cognitive and metacognitive tools: The use of the Internet as an 'epistemological' tool for instruction. *British Journal of Educational Technology, 35*(5), 525–536.

Ulyshen, T. Z., Koehler, M. J., & Gao, F. (2015). Understanding the connection between epistemic beliefs and Internet searching. *Journal of Educational Computing Research, 53*(3), 345–383.

Van Strien, J. L., Brand-Gruwel, S., & Boshuizen, H. P. (2014). Dealing with conflicting information from multiple nonlinear texts: Effects of prior attitudes. *Computers in Human Behavior, 32,* 101–111.

Walraven, A., Brand-Gruwel, S., & Boshuizen, H. P. A. (2013). Fostering students' evaluation behaviour while searching the internet. *Instructional Science, 41*(1), 125–146. https://doi.org/10.1007/s11251-012-9221-x

White, A. M., DeCuir-Gunby, J. T., & Kim, S. (2019). A mixed methods exploration of the relationships between the racial identity, science identity, science self-efficacy, and science achievement of African American students at HBCUs. *Contemporary Educational Psychology, 57,* 54–71. https://doi.org/10.1016/j.cedpsych.2018.11.006

Wiblom, J. (2019). Developing students' critical reasoning about online health information: A capabilities approach. *Res Sci Educ,* 24.

Wiley, J., Goldman, S. R., Graesser, A. C., Sanchez, C. A., Ash, I. K., & Hemmerich, J. A. (2009). Source evaluation, comprehension, and learning in Internet science inquiry tasks. *American*

Educational Research Journal, 46(4), 1060–1106. https://doi.org/10.3102/0002831209333183

Wineburg, S. S. (1991). On the reading of historical texts: Notes on the breach between school and academy. *American Educational Research Journal, 28*(3), 495–519.

Wineburg, S. S., & McGrew, S. (2017). *Lateral reading: Reading less and learning more when evaluating digital information* [PDF document]. https://pdfs.semanticscholar.org/5b26/9628f4dc29b514dfcb0b5e429e49fc0dae6d.pdf?_ga=2.134731505.1090439893.1607365627-347836626.1607210879

Wopereis, I. G., & van Merriënboer, J. J. (2011). Evaluating text-based information on the World Wide Web. *Learning and Instruction, 21*(2), 232–237.

Zhang, S., & Duke, N. K. (2011). The impact of instruction in the WWWDOT framework on students' disposition and ability to evaluate web sites as sources of information. *The Elementary School Journal, 112*(1), 132–154. https://doi.org/10.1086/660687