FULLY APT EPISTEMIC PERFORMANCE, EPISTEMIC COGNITION, AND TRANSFER ACROSS ACADEMIC DISCIPLINES: PHDS, PEDAGOGY, AND PROCESSES

Victor M. Deekens

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the School of Education (Learning Sciences and Psychological Studies).

Chapel Hill
2019

Approved by:

Jeffrey A. Greene

Janice L. Anderson

Matthew L. Bernacki

Jason L. G. Braasch

R. Keith Sawyer

ABSTRACT

Victor M. Deekens: Fully Apt Epistemic Performance, Epistemic Cognition, and Transfer Across Academic Disciplines: PhDs, Pedagogy, and Processes
(Under the direction of Jeffrey A. Greene)

The ability to effectively determine trustworthy and accurate information is essential in the complex digital environment of the modern world. Apt epistemic performance, the competence required for individuals to reach epistemic aims (Barzilai & Chinn, 2018), is essential to successfully navigating today's complex information environment. However, research indicates that apt epistemic performance is both uncommon and discipline-specific. There is also renewed interest, from scholars and others, into questions about if and how knowledge, skills, and practices transfer from one academic discipline to another. Numerous theories have emerged to capture transfer. Notably, these include traditional cognitive theories and more recent conceptualizations such as the actor-oriented theory of transfer. More research is needed into whether and how apt epistemic performance transfers in order to better prepare students to navigate the complex digital world of today. I investigated the transfer of apt epistemic performance with data captured using think-aloud protocol. The data was gathered as experts from education (i.e., no-transfer), other social sciences (i.e., near-transfer), and natural sciences (i.e., far-transfer) reviewed four publications and answered questions about a complex problem in education. Specifically, nine participants, three representing each group (i.e., no-transfer/educators; near-transfer/other social scientists; and far-transfer/natural scientists), evaluated research evidence about the efficacy of flipped classroom pedagogy and made recommendations to a hypothetical colleague about whether or not to utilize a flipped classroom.

My findings indicated that experts, from both the near- and far-transfer groups, were able to positively transfer the apt epistemic performance they developed in their own disciplines to answer complex education questions. However, with respect to two topics, conclusions reached and source evaluation, there were distinct differences in the complexity demonstrated by in-domain (i.e., no-transfer/education) experts when compared to members of the transfer groups. Likewise, negative transfer, or the transfer of knowledge and skills that hinder thinking in a different domain, was demonstrated by members of the far-transfer group as they worked with qualitative data. I found value in the aspects of apt epistemic performance, as defined by Barzilai and Chinn (2018), and both actor-oriented and cognitive theories of transfer, as tools to understand the transfer of epistemic performance across domains. My findings have implications for the study of epistemic cognition and transfer.

ACKNOWLEDGEMENTS

I would like to sincerely thank my family and colleagues who were absolutely essential to the completion of this work. This list must begin with my family and, specifically, my wife, Jennifer, who unfailingly supported this endeavor by providing the perfect mix of motivation and confirmation to ensure that I would succeed. Thank you! Thank you also to my daughters, Kenzie and Emery, who sacrificed so much to ensure that I could meet this goal. I also thank my Mom and Dad, Carolyn and Mac, for providing continued encouragement throughout this process and eagerly asking for status updates along the way. I would also like to thank my brother, Carter, and sister-in-law, Tammy, for providing a listening ear and frequently providing encouragement, and even shelter, as I pushed to finish this project.

I owe my sincerest gratitude to my mentor and advisor, Dr. Jeff Greene, who gave so much of his time and energy over more than 12 years to ensure that I could complete this task. All of your students owe you a great debt, and, on behalf of them, I thank you for giving so much of yourself to ensure that we develop as researchers, teachers, and colleagues. I would also like thank my committee members who gave their time and effort to make this project better at every stage. I also extend my thanks to my fellow graduate students and colleagues who toiled through this process with me. This list begins with my Learning Sciences and Psychological Studies cohort: Dr. Kelly Barber-Lester, Dr. Kayley Lyons, and Kristin Bedell. Thank you. I also owe a deep gratitude to members of the CLICK Research Group who consistently impressed and challenged me throughout our time together. This group includes my co-coder for this project, Brian Cartiff, who spent countless hours helping bring this project to fruition and ensuring we

reached the best conclusions possible. The CLICK group also includes my friends and colleagues: Dr. Nikki Lobczowski, Dana Copeland, Seung Yu, Dalila Dragnic-Cindric, Bekah Duke, Bekah Freed, Robert Plumley, and Liz Allen. Each of them made a lasting impact on me through numerous hours of discussion, shared challenges, and scholarly victories. Many wishes for continued success to each of you. I also would like to thank the members of my original UNC cohort, Dr. Jane Robertson Evia and Dr. Lara Costa for their continued friendship.

I also would like to thank the members of my Army family who made this possible, but are unfortunately too numerous to name. In lieu of listing all of those who have helped me in this journey, I will specifically thank Colonel Everett S. P. Spain, PhD and Dr. Margie Carroll who, along with so many others, helped to make the completion of this project possible through their own efforts and sacrifices. Thank you.

TABLE OF CONTENTS

CHAPTER I: Introduction

Much of the current discussion in the United States centers around truth and falsehood in public discourse. This discourse occurs in a complex information environment consisting of a plethora of news and media sources of varying quality. Approximately 58% of surveyed U.S. adults indicated it is difficult to determine what is true in today's news environment (Gallup, 2018). For example, recent survey results from the Pew Research Center indicated that 68% of Americans consume at least some of their news from social media and that a majority, 57%, view this content as largely inaccurate (Pew, 2018). The now common use of the term "fake news" reflects ongoing concerns with the veracity of information. Educational scholars are also focused on questions about truth as reflected in the 2019 theme of the American Educational Research Association (AERA) Annual Meeting: "Leveraging education research in a "post-truth" era: Multimodal narratives to democratize evidence" (AERA, 2018, para. 1). The introduction to the conference seized upon the challenges of operating in an environment where "personal beliefs and emotions hold more sway than objective facts and evidence" (AERA, 2018, para. 4). Concerns like these have increased scholarly interest in epistemic cognition (EC) or how learners "acquire, understand, justify, change, and use knowledge in formal and informal contexts" (Greene, Sandoval, & Bråten, 2016, p. 1). Individuals' EC plays an important role in the determinations they make about truth and knowledge. EC, sometimes described as "beliefs about knowing and knowledge" (Kuhn, Cheney, & Weinstock, 2000, p. 309) or "cognition of or relating to knowledge" (Greene et al., 2016, p. 3), is an important predictor of intellectual performance both in and out of school (Greene et al., 2016; Strømsø & Kammerer, 2016) and

differences in students' EC relate to multiple educational outcomes and processes including academic achievement (Greene, Cartiff, & Duke, 2018) and self-regulated learning (Muis, 2007, 2008). In the current information environment, individuals' quest for knowledge requires the activation of EC to ensure that they learn verifiable information (Alexander, 2016). Thus, successful learners must demonstrate apt epistemic performance, or performance that achieves valuable learning goals through competence that is grounded in EC (Barzilai & Chinn, 2018).

Given EC's importance to knowledge acquisition, researchers have begun investigating practical considerations such as whether or not efficacious interventions can be designed to increase learners' EC skills (see Bråten, 2016). Equally important are questions about whether or not apt epistemic performance developed in one academic discipline or situation can be effectively utilized in, or transferred to, other academic disciplines. To my knowledge, only one group of researchers (Greene, Chinn, & Deekens, under review) has conducted research into the transfer of apt epistemic performance. Greene and colleagues investigated the transfer of apt epistemic performance by asking university professors from different academic disciplines to learn and make judgments about the replication crisis in psychology. Specifically, Greene and colleagues asked whether or not the ability to enact apt epistemic performance university professors were likely to have in their own academic disciplines effectively transferred as they worked to answer complex questions about the replicability crisis in psychology. To test this, they asked professors with expertise in psychology, other social sciences (e.g., anthropology), and natural sciences (e.g., chemistry) to participate. The other social scientists represented the near transfer group whereas the natural scientists represented far transfer. The preliminary findings from Greene and colleagues' study indicated that this research design is a promising

way to investigate whether and how apt epistemic performance transfers. Using a similar research design, I asked a new sample of professors to learn about a pedagogical practice and make judgments about its efficacy. In doing this, I expanded upon Greene and colleagues' work by furthering the investigation into whether and how EC and apt epistemic performance transfer. This question is particularly relevant as University professors should have knowledge about teaching, but it is unclear whether they can transfer the apt epistemic performance they enact in their own disciplines to questions of pedagogy.

University professors in the United States are generally perceived as experts in their field, and, given this expertise and their positions, they are crucial to the development of the nation's college graduates. Professors are selected for this important role only after rigorous screening by university hiring committees following years of study in a specific field during which they acquire and demonstrate expertise on relevant topics. This expertise enables apt epistemic performance in their chosen academic discipline. University professors are largely responsible for determining the pedagogical practices used in their courses. For example, the professor typically selects the structure of a course, the curriculum, and the measures of student performance and learning. In concert with these decisions, professors must understand and select from a diverse and constantly developing set of pedagogical tools including teaching techniques and technological advancements to determine the best ways to create, deliver, and access content as well as the most effective ways to connect students with each other, with the professor, and with other learning resources. Therefore, it is necessary to investigate whether apt epistemic performance established in a professor's academic discipline transfers to another aspect of their profession, teaching, to allow them to make the best pedagogical choices. Findings from this

investigation could inform whether and how to teach non-experts to transfer or engage in apt epistemic performance.

This project is built on historical and developing literature in both EC and transfer. In the following sections, I will briefly introduce both fields including the evolving definitions and models used to describe each. Compelling questions bearing on problems for researchers who investigate both EC and transfer include the definition and nature of knowledge and transfer, what develops as individuals learn, and to what extent both EC and transfer are situated in contexts.

Epistemic Cognition

First conceptualized by Perry (1968/1999), EC research is grounded in findings from the fields of psychology, philosophy, and education (Greene et al., 2016). Researchers have investigated how EC relates to multiple other educational variables including academic outcomes (e.g., Greene, Muis, & Pieschl, 2010), cognition and motivation (Hofer & Pintrich, 1997), and self-regulated learning (Greene, Yu, & Copeland, 2014). The proliferation of research into knowledge and ways of knowing has consequently led to the use of multiple terms to describe similar or overlapping concepts including epistemological beliefs (e.g., Schommer, 1990), epistemic beliefs (e.g., Kitchener, 2002), or personal epistemology (e.g., Hofer & Pintrich, 2002). With acknowledgement of the semantic and conceptual differences in these terms, following the advice of Greene, Sandoval, and Bråten (2016), I will use epistemic cognition throughout this proposal to refer to the construct and, when necessary, highlight and explain researchers' differing use of terminology.

Researchers have described EC using multiple models. A complete review of these models is beyond the scope of this paper and several influential models including justification models (e.g., Greene, Azevedo, & Torney-Purta, 2008; Bråten, Ferguson, Strømsø, & Anmarkrud, 2013) have been excluded purely for brevity's sake. Developmental models such as those created by Perry (1968/1999) and more recently by Kuhn, Cheney, and Weinstock (2000) have been particularly influential. Developmental models of EC portray a growing complexity of thought from naïve to more sophisticated ways of understanding the world. Instead of the term EC, Kuhn, Cheney, and Weinstock (2000) used the term epistemological understanding, which they defined as "beliefs about knowing and knowledge" (p. 309). In their model, they described development as learners increasing their ability to coordinate the subjective and objective dimensions of knowing. This development was depicted in a stage model where individuals in more mature stages of development exhibited more complex views about the nature of knowledge, the value of critical thinking, and the extent to which an external reality is understandable.

In contrast, researchers who proposed multi-dimensional models (e.g., Schommer, 1990) described EC as a system of independent beliefs about the nature of knowledge and knowing that play a key role in comprehension and learning but do not necessarily develop concurrently in a stage-like manner. After reviewing and comparing the extant EC literature, Hofer and Pintrich (1997) proposed the most influential multi-dimensional model. In their framework, Hofer and Pintrich described EC using four dimensions of belief. Two of these dimensions (i.e., simplicity of knowledge and certainty of knowledge) concern the nature of knowledge (i.e., what

knowledge is) and the other two (i.e., source of knowledge and justification of knowing) concern the nature of knowing or one's beliefs about how knowledge is gained.

Despite their differences, the relatively early EC theorists, like Perry in the late 1960s or Schommer more than twenty years later in 1990, tended to describe EC as a consistent trait about a learner that applied universally across contexts. More recently, researchers (e.g., Hammer & Elby, 2003) have described knowledge and beliefs about knowledge as situated in contexts using a third type of EC model, epistemological resources models. Hammer and Elby challenged extant models that portrayed EC as a unitary construct and, instead, building on their own work in science education, proposed a more granular model. Grounded in a belief that multi-dimensional EC theorists (e.g., Hofer & Pintrich, 1997; Schommer, 1990) had more accurately captured EC than developmental theorists, Hammer and Elby built on multi-dimensional models and investigated the situated aspects of EC by asking questions about whether or not the skills and beliefs associated with understanding knowledge generalize or are instead specific to contexts. Fundamentally, they rejected previous theorists' assumptions that (1) learners have epistemological beliefs stored as declarative knowledge that are consciously accessed during learning, and (2) these beliefs about knowledge are trait-like (Mason, 2016). Instead, Hammer and Elby held that EC is different not just across academic disciplines but even within academic disciplines and across contexts. The specificity of EC in these models has raised questions about the applicability of epistemic resources-based models for answering questions about how to educate people to effectively employ EC.

The AIR (i.e., aims, ideals, and reliable processes) model (Chinn & Rinehart, 2016; Chinn, Rinehart, & Buckland, 2014) will provide the primary foundation for this project and was

characterized by Mason (2016) as a modern approach to describing EC. The AIR model does not fit neatly into one of the previously described categories of models (i.e., developmental, multi-dimensional, or epistemological resources models). Instead, Chinn and colleagues described EC primarily in terms of a learner's aims, either epistemic (e.g., achieving understanding) or non-epistemic (e.g., wasting time or relaxing), ideals (i.e., the standards learners use to determine if their aims have been met), and the reliable processes (i.e., cognitive and social processes and methods of inquiry by which knowledge is achieved). Their intent was to expand the definition of EC by building on the extant work in the field and incorporating philosophical considerations. Specifically, they built on Hofer and Pintrich's (1997) framework to produce a more encompassing model that they posited would more effectively address "the full range of psychologically important epistemic phenomena, concepts, and issues" (Chinn, Buckland, & Samarapungavan, 2011, p. 142). Overall, the AIR model added three novel contributions to extant models of EC: (a) the inclusion of motivation, or aims to use the language of Chinn and colleagues, which had traditionally been treated as distinct from EC; (b) the incorporation of a broader range of aims beyond simply the gathering of knowledge to include the acquisition of wisdom, gathering of evidence, a focus on justified beliefs, and the building of valuable models; and (c) a focus on the value of information to meet selected aims and the ideals and reliable processes used to achieve them (Chinn et al., 2014). Each of these distinctions will be valuable in this study. Specifically, Chinn and colleagues' expansion of EC to include descriptions of learners' aims (i.e., epistemic v. non-epistemic) and the reliable processes (e.g., source evaluation) they employ to attain knowledge directly connect to questions about whether and how experts' apt epistemic performance in one field transfers to another domain.

Barzilai and Zohar (2014; 2016) proposed another modern model of EC, the multi-faceted framework of epistemic thinking. In this model, the authors described EC as consisting of both cognitive and metacognitive aspects. Metacognition is a complex term that has been defined in multiple ways. Barzilai and Zohar adopted a broad characterization based on Flavell's original definition (1979) that typifies metacognition as "knowledge about cognition and regulation of cognitive activities" (Barzilai & Zohar, 2014, p. 16). Specifically, they described epistemic thinking as consisting of multiple facets that Barzilai and Chinn (2018) summarized as four aspects: (1) cognitive epistemic processes and strategies (i.e., thinking about the epistemic characteristics of information and sources); (2) epistemic metacognitive skills (i.e., planning, monitoring, and evaluating the nature of gained knowledge); (3) epistemic metacognitive knowledge (i.e., "knowledge, beliefs, ideas, and theories regarding the nature of knowledge and knowing"; Barzilai & Zohar, 2014, p. 20); and (4) epistemic metacognitive experiences (i.e., the feelings evoked by the knowledge building process). Grounded in this definition of EC, Barzilai and Zohar proposed that EC and metacognition overlap.

Citing the largely descriptive nature of the AIR framework, which created challenges to operationalizing the model for education, Barzilai and Chinn (2018) integrated it with facets of Barzilai and Zohar's (2014; 2016) multi-faceted framework to create the Apt-AIR framework. In this framework, they defined a critical term for this study, apt epistemic performance. Barzilai and Chinn argued that the goal of epistemic education should be to enhance apt epistemic performance and defined it using five key aspects. First, learners exhibiting apt epistemic performance perform reliable processes to achieve epistemic aims. Second, they adapt their epistemic processes to a range of situations. Third, they regulate and understand their epistemic

performance (i.e., they employ epistemic metacognitive knowledge). Additionally, they care about and enjoy their epistemic performance (i.e., they are motivated to succeed). Finally, they participate in epistemic performance with others. This fifth aspect is not solely focused on working in pairs or groups but also includes an understanding of the role of social configurations in knowledge making. For example, an individual demonstrating the fifth aspect of apt epistemic performance when reviewing scholarly work recognizes the value of the work of others in the field and strives to contribute to the knowledge being built by other scholars.

Researchers have used a variety of research methods to investigate EC including both qualitative and quantitative investigations of a wide variety of types of data such as student discourse (e.g., Herrenkohl & Cornelius, 2013), interviews (e.g., Weber & Mejia-Ramos, 2011), student question types (e.g., Portnoy & Rabinowitz, 2014), or think-aloud protocols (TAP; e.g., Greene & Yu, 2014; Shreiner, 2014). Likewise, researchers investigating epistemic-cognition-related questions have conducted research across a wide range of age groups and expertise levels from elementary students in science and history classes (e.g., Herrenkohl & Cornelius, 2013) to adult mathematicians (e.g., Weber & Mejia-Ramos, 2011) and other experts across a variety of academic disciplines. I modeled some components of this project after Shreiner's (2014) use of think-aloud protocols to investigate differences in reasoning between experts and novices as they employed historical knowledge to reason about a political issue. Shreiner selected political scientists to represent experts and high school students to represent novices. In her study, Shreiner compared differences in the EC employed by experts and novices as they explored a question in a single academic discipline, history. In contrast, I will focus on university professors with expertise in different academic disciplines (i.e., education, other social sciences, and

traditional sciences), as they wrestle with a complex question from a single discipline, education, in order to investigate if and how experts transfer their apt epistemic performance.

Transfer

Research about knowledge transfer can be traced to the beginning of the 20th century (Day & Goldstone, 2012) with key works attributed to Thorndike (1924; Thorndike & Woodworth, 1901). Transfer is at the core of all learning theories; questions about how effectively learners apply what they learn in one situation to novel problems are essential to understanding the efficacy of the education system (Bransford & Schwartz, 1999). Traditional cognitive views of transfer are grounded in the 1960s era psychological research on cognition when knowledge was represented by discrete symbols that could be reorganized to support analogical reasoning (Day & Goldstone, 2012). This conceptualization of cognition led to the widely utilized definition of transfer as "the recruitment of previously known structured symbolic representations in the service of understanding and making inferences about new, structurally similar cases" (Day & Goldstone, 2012, p. 154).

Researchers' interest in questions of transfer has seen a recent resurgence inspiring several new publications including a 3-year strand of articles in *The Journal of the Learning Sciences* (*JLS;* Lobato, 2006) and a special issue of *Educational Psychologist* focused on the "new conceptualizations" of transfer (Goldstone & Day, 2012a). Goldstone and Day (2012b) attributed this resurgence to multiple causes. First, modern theorists questioned the traditional cognitive conceptualization of transfer by challenging the view of knowledge as a stable mental entity that can be applied across situations without alteration. For example, Lave and Wenger (Lave, 1988; Lave & Wenger, 1991), in developments analogous to those proposed in EC by

epistemological resources model researchers (e.g., Hammer & Elby, 2002), contended that learning and knowledge are inherently situated in context. Researchers who have expressed a purely situative view of learning (e.g., Lave, 1988) have argued that there are serious concerns about whether transfer occurs at all. Specifically, situated researchers have questioned how effectively learners can transfer knowledge from one situation to a situation that is dissimilar (i.e., far transfer) but have left open the possibility of transfer among two similar situations (i.e., near transfer). The second driver of the resurgence in transfer research is a collective recognition among researchers that key principles (e.g., positive and negative feedback loops) appear naturally across multiple academic disciplines (e.g., economics, biology, and physics) leading researchers to question the most effective ways to ensure that learners recognize these similarities and apply knowledge across academic disciplines and contexts. Goldstone and Day's (2012b) final driving factor in the resurgence of transfer research is the creation of new methodologies for studying transfer that now can be combined with previous methodologies to address pressing research questions (e.g., asking explicit questions about whether or not learners consciously transfer knowledge). For example, new methodologies for studying transfer include implicit measures such as focusing on how learning about one topic prepares students to learn about another (i.e., the preparation for future learning model of transfer; Bransford & Schwartz, 1999) or measuring how students construe current situations to match previous ones.

This new interest in transfer has led to multiple additional conceptualizations of transfer and learning (Lobato, 2006). Lobato introduced several of these in the introductory article for the transfer strand in *JLS*. After reviewing the conceptualizations described by Lobato, I selected two as most relevant to this project. The first, preparation for future learning (Bransford &

Schwartz, 1999), is used to investigate the role that prior experience plays in learners' interpretation of new information. The second, the actor-oriented perspective (Lobato, 2012), counters the traditional cognitive approach to understanding transfer by framing transfer as the generalization of learning. Specifically, this perspective is differentiated from the traditional cognitive perspective along five dimensions including the nature of knowing and representation (i.e., EC), point of view, methods, goals, and what transfers (Lobato, 2006). For example, researchers investigating transfer from the traditional cognitive perspective focus on what transfers from an observer's point of view, however, those investigating it from an actor-oriented perspective focus on the learner's perspective. Likewise, there is a clear distinction between how researchers from each perspective describe what transfers. From the traditional cognitive view, clearly defined actions and strategies are what transfers. In contrast, researchers using an actor-oriented perspective take a more holistic view that includes a broad focus on the use of prior knowledge and skills from the learner's perspective. That is, researchers employing an actor-oriented perspective of transfer do not begin with an *a priori* solution that learners must achieve. Instead, researchers assume some level of transfer and look for instances of the use of prior knowledge, allowing for the capture of transfer that may be ignored from the traditional perspective.

These three perspectives on transfer (i.e., traditional, preparation for future learning, and actor-oriented) were each utilized for this project for a variety of reasons. For example, researchers utilizing a traditional view of transfer may argue that apt epistemic performance should transfer whereas researchers from either the preparation for future learning or actor-oriented perspectives may argue that transfer is either unlikely or, at least, qualitatively different

from transfer captured from the traditional cognitive perspective. These differing viewpoints enhanced my ability to investigate, recognize, and capture the transfer of apt epistemic performance as it occurred.

This Project

The goal of this project was to contribute to the broader literatures on EC and transfer while specifically answering questions about if and how apt epistemic performance transfers. I selected a single pedagogical practice, flipped classrooms, as an exemplar topic for participants to investigate. Flipped classrooms take many forms, but typically instructors utilizing a flipped classroom introduce students to content prior to class meetings, possibly using online tools (e.g., web presentations, podcasts, or recorded lectures), to maximize class time for teacher facilitated application of the content (Jensen, Kummer, & Godoy, 2015). Flipped classroom are not new, but interest in them is growing as exemplified by articles in the popular press (e.g., The New York Times; Fitzpatrick, 2012).

In order to investigate this, I asked University professors representing different academic disciplines to think-aloud while learning about flipped classrooms. The context for this project, university professors' pedagogical choices, provided a rich environment for the exploration of EC questions. Professors from education, other social sciences (e.g., history or psychology), and traditional sciences (e.g., chemistry or physics) possess expertise, and the related apt epistemic performance, in their respective fields, but, due to their role as university educators, should also be competent and knowledgeable about pedagogy. If apt epistemic performance transfers, it should do so for all university professors regardless of academic discipline. If this is the case, when they were asked to analyze research about pedagogical practices, I expected that all

participants would perform well. That is, they would engage deeply with the material and employ reliable processes to achieve epistemic aims. Simultaneously, they would adapt the reliable processes they use to understand research in their own disciplines to assess education research, allowing them to make a clear, reasoned recommendation about the utility of flipped classroom pedagogy. Additionally, they would regulate their epistemic performance. If apt epistemic performance only near transfers (i.e., apt epistemic performance transfers to an educational context from other social sciences but not from traditional hard sciences), I expected to see distinctly different and better apt epistemic performance from the other social science professors than from the traditional science experts as they reviewed the flipped classroom materials and answered questions.

Whereas, if apt epistemic performance does not transfer, I expected to see clear differences in both the quality and outcome of the apt epistemic performance exhibited by experts from education when compared with those with expertise in different academic disciplines as they analyzed education research. For example, if apt epistemic performance does not transfer, I predicted that experts from the traditional sciences may attempt to repeatedly apply a reliable process from their field (e.g., seeking evidence produced only in tightly controlled laboratory settings) to understand the evidence about flipped classrooms. This could result in negative transfer, or the application of knowledge learned in one context that is not beneficial in another context. In contrast, positive transfer occurs when information learned in one context is helpful in a new situation or on a novel learning task.

Following the example set by Greene and colleagues (under review), I conducted a thematic analysis using grounded theory (Glaser & Strauss, 1967) of think-aloud protocol data

gathered from professors at two universities in the United States as they reviewed research on the efficacy of flipped classrooms. I asked three professors from each of three academic backgrounds, education, other social sciences (e.g., history), and traditional hard sciences (e.g., chemistry and physics) to review a series of four provided articles about flipped classroom pedagogy before making a recommendation to a hypothetical third-year college professor at a similar university joining a department similar to theirs. In this scenario, the hypothetical new colleague asked for a specific recommendation about whether or not they should transition to a flipped classroom model or continue with more traditional methods. In coordination with a second coder, I conducted a thematic analysis of the transcripts of this data in order to answer these questions:

1. How do education experts differ from other experts in the EC they employ to evaluate education research?
2. To what degree and in what ways do experts from outside of education transfer their apt epistemic performance to the evaluation of education research?
3. If apt epistemic performance transfers, is this transfer generally positive (i.e., beneficial to achieve a complex understanding of the topic) or negative (i.e., a hindrance to achieving a complex understanding)?
4. What are the differences between apt epistemic performance when it is transferred between similar academic disciplines (i.e., near transfer; e.g., social sciences to education) when compared to transfer between less similar academic disciplines (i.e., far transfer; e.g., hard sciences to education)?

Given the significance of the pedagogical choices made by professors and the fact that the majority of professors in institutions of higher learning earned their doctorates and conduct research in areas other than education, questions about their apt epistemic performance, specifically apt epistemic performance, about the education research they use to make pedagogical decisions, are intriguing. Do ways of knowing and the understanding of the nature of knowledge that underpin a professor's expertise in one academic discipline translate to other academic disciplines? That is, do the skills and practices that allow, for example, history professors, to obtain expertise in history transfer effectively to educational topics as they make decisions about teaching? Can an increased understanding about how experts reason in academic disciplines other than their own reveal new insights about learning and the transfer of apt epistemic performance to assess research in general? Given the increasing interest in scholarship about EC (Greene et al., 2016) an understanding of if and how it transfers across academic disciplines is relevant for further research.

Enhancing the understanding of if and how EC transfers will contribute to literature on fostering EC. Understanding the EC and apt epistemic performance of experts in a given academic discipline will help education researchers grapple with what is necessary to move from "school performance to performance in the world" (Shreiner, 2014, p. 315) and will aid in the quest prescribed by Barzilai and Chinn (2018) and others to enhance education aimed at developing EC. Findings from this study have implications for researchers focused specifically on EC as well as those who focus on the transfer of knowledge and skills from one task to another. As predicted, results reflected initial findings from the ongoing Greene and colleagues' (under review) study that indicated that epistemic performance transferred in both positive and

negative ways. Likewise, findings from this study will further arguments that various perspectives on transfer provide a multitude of ways to capture apt epistemic performance and that further study is required. Simultaneously, this research can to developing theories of transfer by helping to refine what transfer is, if and how it occurs, and whether or not EC transfers.

CHAPTER II: Literature Review

In this review, I describe the extant literature on EC, transfer of learning, and flipped classroom pedagogy. First, I review EC research, briefly detailing the development of the field and a variety of different EC models before moving to a description of two modern models, the multi-faceted approach (Barzilai & Zohar, 2014) and the AIR model (Chinn & Rinehart, 2016). Next, I review Barzilai and Chinn's (2018) definition of apt epistemic performance, which they grounded in the theoretical underpinnings of their separate models of EC. Throughout this section, I introduce current terminology used to describe EC, describe the empirical research in support of each category of model, introduce ongoing questions in the field (e.g., how situated is EC and whether and how it is specific to academic disciplines), and build the case for the value of apt epistemic performance as a means to investigate the transfer of EC. Next, I review the development of transfer research from its conceptualization by Thorndike in the early 1900s to current models with an emphasis on questions that are currently being considered by researchers, a discussion of the recent resurgence in transfer-related research, and connections to this project. Next, I highlight the gap in current research on the transfer of EC to set the framework for the motivation and value of this study. Then, in support of the use of flipped classroom pedagogy as a context to investigate experts' EC, I review the current research on the use of flipped classrooms on college campuses and introduce the articles that participants were asked to review. Finally, I review my research questions.

Epistemic Cognition and Apt Epistemic Performance

Epistemic cognition researchers investigate how learners obtain, justify, alter, and utilize knowledge throughout their lives (Greene et al., 2016). EC is essential for learning and meta-analytic research has shown a correlation between EC and academic achievement (Greene et al., 2018). Over time, multiple researchers have proposed models of EC that vary along several facets including how their creators defined knowledge, how they described what develops as learners mature or advance from novice to expert, and whether and how they incorporated related educational concepts such as motivation, goals, and processes. Grounded in EC research, Barzilai and Chinn (2018) defined apt epistemic performance as the capability to parse truth from falsehood. In this project, I asked experts from different academic disciplines to demonstrate their apt epistemic performance as they learned about an education topic. University faculty are clearly experts in their own academic disciplines and possess knowledge about teaching and learning, however, it was not clear how deep their knowledge of pedagogical practice is and to what extent they review and critically consume content about teaching. An introduction to the different types of EC models coupled with a brief synopsis of the measurement methodology and empirical evidence that supports each model will help set the foundation for this project.

Developmental models of EC. Perry (1969/1999) is generally given credit for initially investigating EC as he sought to better understand how college students made meaning of their experiences. Using interviews conducted using a sample of predominantly white, elite, male college students, he described the changing aspects of what they knew and valued throughout their college years using a structured, stage-like model grounded in the cognitive developmental

theories of the time. In its simplest incarnation, Perry modeled the development of EC as a progression through nine positions that are generally grouped into four stages (Hofer & Pintrich, 1997). The first stage, dualism, is characterized by an absolutist view of the facts where questions have simple right or wrong answers that are determined and known by authority figures. Individuals progressing to Perry's second stage, multiplicity, recognize more uncertainty in the world than dualist thinkers, and, as they reach the end of this stage, they come to believe that because all knowledge is uncertain then all views are equally valid. Individuals in Perry's third stage, relativism, make an important transition that remains significant throughout the final stage as they recognize positionality and the role that individuals play in creating meaning in the world. Learners at this stage identify that knowledge is both contextual and relative, and they begin to understand their own role in selecting what they believe. In the final stage, commitment within relativism, individuals build on the relativism of the third stage increasing their focus on engaging with the world and committing to particular relativistic points of view.

Inspired by the work of Perry (1969/1999) and with an interest in improving perceived gaps in the initial model, including questions about the effect of the mostly white male sample on the results, and challenges with operationalizing the model for further study, researchers continued to build developmental models of EC (Hofer & Pintrich, 1997). For example, Belenky et al. (1986) sought specifically to investigate EC, which they called ways of knowing, in women, and expanded the pool of participants to include learners not enrolled in formal schooling. Today's EC researchers do not typically describe gender differences, but Belenky and colleagues' work to include women as well as their emphasis on the source of knowledge, in

contrast to Perry's focus on the nature of knowledge, helped to expand EC research (Hofer & Pintrich, 1997).

King and Kitchener (1994; 2004; Kitchener, 1983) proposed another developmental model of EC, the reflective judgment model, that depicted three stages of development (i.e., pre-reflective thinking, quasi-reflective thinking, and reflective thinking) grounded in research specifically focused on how college students dealt with ill-structured problems. They defined ill-structured problems as those that could not be completely defined or solved with absolute certainty. They developed and refined the reflective judgment model based on the analysis of interviews conducted while individuals dealt with these problems. Overall, they found differences in individuals' assumptions about knowledge, which they described using a developmental sequence.

Kuhn, Cheney, and Weinstock (2000) built on Hofer and Pintrich's (1997) criticism to restructure Kuhn's (1991) existing developmental model, wherein they sought to answer specific questions about what develops and why. This update yielded an influential developmental model of EC. Kuhn and colleagues identified the increasing ability of individuals to coordinate the subjective and objective components of knowledge as the core developmental task that leads to increasingly mature EC. According to Kuhn and colleagues, learners with immature EC see the objective dimension as dominant (i.e., the external world is directly knowable and personal assertions are copies of that external world). As individuals mature, subjectivity becomes dominant and the objective dimension becomes less influential, leading to a stage of development, similar to one described by Perry, where all ideas are considered equally valid. In more advanced stages of development, an individual coordinates both the objective and

subjective dimensions. To explain this progression, Kuhn and colleagues' updated model depicted an individual's increasing epistemological understanding in four levels beginning from the most basic, realist, and progressing through three more levels starting with absolutist then multiplist and finally evaluativist. In this model, Kuhn and colleagues described the increasing complexity of EC exhibited at each level as development along four subordinate components: 1) assertions 2) reality 3) knowledge and 4) critical thinking.

Measurement and empirical evidence: Developmental perspective of EC. EC researchers who utilized a developmental model to conduct research have traditionally relied upon interviews as their primary methodology (Mason, 2016). King and Kitchener (1994) utilized their reflective judgment interviews in both cross-sectional and longitudinal studies to assess each participant's level of development as described by their model (i.e., pre-reflective thinking, quasi-reflective thinking, and reflective thinking). They assessed participants' EC with an interview focused on four ill-structured problems with six follow-up questions created specifically to elicit assumptions about EC. Trained raters scored these transcripts for epistemic thought in three rounds. Overall, King and Kitchener (1994) found support for their model in results showing that more mature learners and learners who had more advanced education exhibited higher stages of reasoning. In reviewing the history of the reflective judgment model after more than 25 years of research, King and Kitchener (2004) reviewed both longitudinal and cross-sectional research they conducted using their reflective judgment interview methodology and reported strong support for their claims of a developmental sequence in EC. Simultaneously, they cited six additional longitudinal studies with durations ranging from 3 months to 4 years and

published between 1984 and 1990 by a variety of different researchers, each of which supported their claims of development in reflective judgment.

Developmental theorists, including Kuhn, Cheney, and Weinstock (2000), have also utilized written assessments to measure EC. For example, Kuhn and colleagues employed both a written assessment and a separate set of interviews grounded in their developmental framework to assess EC among seven groups of participants including children, teenagers, and adults. They distinguished four types of knowing judgments (i.e., domains): (1) of personal taste, (2) of beauty, (3) of value, and (4) of truth including separate categories about truth about the social world and truth about the physical world. Grounded in their core idea that EC development progresses as individuals coordinate the subjective and objective dimensions of knowing, they assessed each group's judgments in each domain. They used questionnaires in the first study to ask members of each group to make decisions about whether certain statements could be true and analyzed the patterns of their responses. For example, in order to assess participants' EC in the value judgments domain, Kuhn and colleagues juxtaposed three pairs of statements from two hypothetical individuals, Chris and Robin. Participants were shown one statement from each individual: (1) "Robin thinks lying is wrong" and (2) "Chris thinks lying is permissible in certain situations" (Kuhn et al., 2000, p. 317). Participants were asked to select from one of two responses about whether one or both individuals could be right: (1) "Only one right" or (2) "Both could have some rightness" (Kuhn et al., 2000, p. 316). If participants selected that both could have some rightness, then they were asked to determine whether or not one could be more right in this situation. Kuhn and colleagues characterized individuals as either absolutist, multiplist, or

evaluativist within particular domains (e.g., the value judgments domain) based on their responses to the three pairs of statements.

In order to clarify their findings, they conducted a second study using an interview technique to help delineate differences specifically between the absolutist and multiplist levels of their model. Overall, across both studies, they found support for their model, however, there was a great deal of variety in the patterns of development both within and across the knowing judgment domains. For example, they found what they considered an orderly pattern of overall EC development, but it was not consistent across the judgment domains. In a later publication, Kuhn and Weinstock (2002) provided an example of this inconsistency, describing the transition from absolutism to multiplism as occurring earlier in domains connected to personal judgments than in domains related to truth judgments.

In an effort that highlighted the continuing utilization of developmental models of EC, Barzilai and Weinstock (2015) recently utilized a written assessment based on Kuhn and colleagues' (Kuhn, 1991; Kuhn, Cheney, & Weinstock, 2000) model. The primary objective of the study was to create a more reliable and valid pen and paper-based tool in order to complement existing but resource intensive interview-based methods with measures that could be given to larger and more varied samples. They used exploratory and confirmatory factor analyses to investigate the data and found support for the developmental model of EC. They also found differences in development across academic disciplines, providing additional evidence for the ongoing debate about the extent to which EC is domain-general or domain-specific.

Critiques of developmental EC research. Sandoval (2012) disputed the value of developmental EC theorie,s citing the continued failure of these models to explain empirical

findings. Specifically, he noted shifts in developmental psychology research including challenges to Piaget's work on which Perry and other developmental theorist grounded their initial theories of EC. Sandoval made a distinction among the developmental theories that I have not yet made, namely labeling some (i.e., Perry, 1969/1999; King and Kitchener, 1994) as classical developmental stage theories and others (e.g., Kuhn et al., 2000) as neo-Piagetian developmental theories. The key distinction between these two, according to Sandoval, is whether the qualitative changes that lead an individual to advance from one stage to the next are driven by the development of the underlying cognitive structure (i.e., Piagetian development) or a change in knowledge that is not as directly tied to changes in cognitive ability. Across both of these types of developmental theories of EC, Sandoval found a lack of empirical research to support the existence of either end of the developmental continuum claimed by developmental EC theorists. That is, researchers have typically not found evidence in support of either the most naïve stage with its pure focus on the objective (e.g., Perry's dualism; Kuhn, Cheney, & Weinstock's realist) or the most advanced stage of development (e.g., Perry's commitment within relativism; Kuhn Cheney, & Weinstock's evaluativist; King & Kitchener's reflective thinking) (Sandoval, 2012). For example, despite King and Kitchener's (2004) claims in their review of more than 25 years of research using their reflective judgment model that there was adequate support for a developmental trajectory, even they admitted that individuals could rarely be placed in a single stage.

Despite these challenges, developmental models of EC continue to be employed by researchers (e.g., Barzilai & Weinstein, 2015). Unquestionably, researchers who developed and investigated developmental models of EC, including Perry (1968/1999), laid the foundation for

the field. Additional EC theorists built on these models as they proposed qualitatively different ways to envision EC including as a set of interconnected beliefs that do not necessary develop in unison.

Multidimensional sets of beliefs models of EC. In an influential publication, Hofer and Pintrich (1997) reviewed the available developmental models described above and others (e.g., Baxter Magolda, 1987) before calling for the refinement of language in the field and improvement in the methodologies used to investigate EC. Simultaneously, they organized the existing models into categories to ease comparison and drive innovation in EC research. Overall, they categorized six models in terms of how each dealt, or failed to deal, with two core epistemological dimensions (i.e., the nature of knowledge and the nature of knowing) and two peripheral beliefs about learning, instruction, and intelligence (i.e., nature of learning and instruction and nature of intelligence). Ultimately, in their 1997 article, Hofer and Pintrich proposed their own model of EC that will be discussed below.

Hofer and Pintrich (1997) sought commonalities among extant EC models and proposed clear directions for the future language of the field but, prior to proposing their own model, the authors notably discussed only one dimensional model of EC, produced by Schommer (1990; 1993a; 1993b), that was qualitatively different from the developmental models discussed thus far. Inspired by the mixed results of experimental studies conducted to validate Perry's (1968/1999) model and perceived challenges with the assumptions of previous researchers who crafted unidimensional developmental paths of EC, Schommer instead modeled EC, which she called personal epistemology, as a system of independent beliefs about the nature of knowledge.

Notably, Kuhn and colleagues (2000) and King and Kitchener (1994) differentiated stages in their developmental models in ways that could be perceived as dimensional but they did not emphasize these dimensions as key characteristics of their models (Greene, Azevedo, & Torney-Purta, 2008). For example, as previously discussed, Kuhn and colleagues conducted empirical work to support their developmental model and specifically addressed changes in the patterns of EC exhibited by participants in categories. Empirical evidence collected in support of Kuhn and colleagues' model showed that as people developed across levels their views on each dimension changed.. However, their model is predominantly a developmental one that acknowledges the role different dimensions may play in development without emphasizing them.

In contrast to theorists who modeled EC using unitary developmental models, Schommer (1990) specifically described an individual's EC using four dimensions that do not necessarily develop in unison, but, instead, are developed and utilized independently. Each dimension is best conceptualized as a continuum beginning with what Schommer called a naïve view, which gives each factor its name, and progressing to a more advanced view. The first factor, simple knowledge, described beliefs about the complexity of knowledge ranging from individuals who believe that knowledge is simple and constructed of unambiguous and isolated facts to the more advanced view of knowledge as a series of interrelated concepts. Certain knowledge, the second factor, progressed from a description of individuals who see knowledge as absolute and unchanging to the more advanced view held by those who recognize knowledge as evolving over time. Individuals demonstrating a naïve characterization of the third factor, quick learning, report that people should be able to learn new information or it cannot be learned. According to Schommer, a more advanced learner instead recognizes that learning is a gradual process. The

final factor that Schommer found support for in her empirical investigations, fixed ability, is closely linked to implicit theories of intelligence (Dweck & Leggett, 1988), which would later become an influential psychological concept better known as mindset (Dweck, 2008). Schommer used fixed ability to describe a progression from a belief that intelligence is predetermined to a belief that it can be improved. Fundamentally, Schommer expanded researchers' existing theories of EC by suggesting that it may best be described as a system of independent dimensions that do not necessarily develop simultaneously. The simplicity of knowledge and certainty of knowledge factors were similar to constructs described in extant developmental models, however, her inclusion of additional factors and the insightful ways she employed empirical research, including the development of her own EC questionnaire and further investigation of the links between EC and classroom performance, helped to advance researchers' understanding of EC (Hofer & Pintrich, 1997).

After reviewing the existing literature including Schommer's (1990) model, Hofer and Pintrich (1997) proposed another multi-dimensional model that Chinn, Rinehart, and Buckland (2014) described as the most influential model of EC during the 15 years immediately following its publication. Hofer and Pintrich described EC using four dimensions specifically about knowledge and knowing. They recognized value in Schommer's description of beliefs about learning (e.g. Schommer's fixed ability and quick learning). Specifically, Hofer and Pintrich valued beliefs about learning to explain learner motivation and saw clear connections between beliefs about learning and beliefs about knowledge and knowing. However, they labeled beliefs about learning as not epistemic in nature and instead focused their model on knowledge and knowing. Hofer and Pintrich used two dimensions to describe beliefs about the nature of

knowledge: (1) beliefs about the certainty of knowledge and (2) beliefs about the simplicity of knowledge. The other two dimensions were used to describe beliefs about the nature of knowing: (1) source of knowledge and (2) justification for knowing. Each of these dimensions describes a continuum ranging from a naïve view to a more complex view. For example, beliefs about the certainty of knowledge range from the naïve belief that knowledge is absolute and unchanging to the more complex understanding that knowledge is evolving. The continuum of beliefs about the complexity of knowledge range from the naïve argument that knowledge is simply a gathering of isolated facts to the more complex understanding that knowledge is made up of interrelated concepts. According to Hofer and Pintrich, individuals who operate from a naïve conception of the source of knowledge understand it to be externally created and maintained by authority figures, in contrast, those who hold a more complex view see knowledge as something that is actively constructed through their own interactions. Finally, the fourth dimension, justification for knowing, ranges from the simple view that claims must be established through authority, observation, or simply feel to a more complex understanding that there are rules of inquiry, evaluation, and integration that must be used to justify knowledge claims.

Measurement and empirical evidence: Multi-dimensional perspective of EC.

Researchers who rely on multi-dimensional models to explain EC have most often utilized self-measures (Mason, 2016). Schommer's (1990) groundbreaking use of a Likert-type survey instrument to measure EC from a multi-dimensional perspective influenced other researchers. For example, Hofer (2000) used a revised and shorter version of Schommer's Epistemological Belief Questionnaire to investigate the domain-generality and specificity of EC beliefs. Multi-dimensional models of EC have been used in numerous contexts to evaluate the effects of EC on

multiple outcomes. For example, Muis (2007) reviewed multiple models of self-regulated learning (SRL) and summarized the role of EC as characterized by influential multi-dimensional models (i.e., Schommer, 1990; Hofer & Pintrich, 1997) in SRL processes. Greene, Muis, and Pieschl (2010) integrated multi-dimensional belief models of EC with SRL models to explain learning in computer-based learning environments. Bråten, Britt, Strømsø, and Rouet (2011) analyzed the role of learners' EC on a variety of learning outcomes and processes with a focus on explaining comprehension from multiple texts.

In Schommer's seminal 1990 article, she described two studies where she both developed and utilized her multi-dimensional set of beliefs model of EC. In study 1, she investigated the factors of her multi-dimensional framework using a self-report questionnaire and factor analysis. The questionnaire contained questions specifically designed to measure the existence of five hypothesized factors of EC. Her analysis supported the existence of the four individual factors of EC discussed thus far (i.e., fixed ability, simple knowledge, quick learning and certain knowledge). In study 2, she investigated the role of students' EC on their performance, monitoring, and conclusions as they conducted a learning task. Two groups of students were each assigned to read a passage. Each passage was presented as a textbook chapter that lacked a concluding final paragraph. For example, the students in one group read a psychology-related passage about four distinct theories of aggression with an underlying theme indicating that a complete theory would require the integration of aspects from each. Students were asked to imagine themselves as the author of the chapter tasked to write a concluding paragraph. In addition to producing the concluding paragraph, students completed a 10-item multiple-choice test to measure their comprehension of the material. Students' concluding paragraphs were coded

for simplicity and certainty on a dichotomous scale. Students who oversimplified the material received a rating of simple while those who elaborated and integrated the information were rated as complex. Likewise, students who highlighted uncertainty in their concluding paragraphs received a rating of uncertain while those who concluded that answers to the complex problems presented in the passages were known or would be known were rated as certain. Using regression analysis, Schommer found distinct effects of students' EC on their learning and comprehension including their processing of information and monitoring of their learning. For example, when asked to integrate a complex topic like the ones in this study, Schommer found that students' beliefs about quick learning (i.e., if they do not learn it quickly, they will not learn it all) affected their ability to integrate the information.

Critiques of multi-dimensional EC research. The reliance of multi-dimensional EC researchers on self-report measures and the related psychometric issues with reliability and validity have undermined the evidence used to support multi-dimensional models (Mason, 2016). Likewise, Mason highlighted additional issues with empirical evidence for multi-dimensional models of EC including challenges with explaining the developmental trajectory of the model (see Muis, Bendixen, & Haerle, 2006), low reliability of factors determined by factor analysis, and the use of homogeneous samples. Sandoval (2012) found little empirical evidence to support multi-dimensional sets of belief models (e.g., Schommer, 1990; Hofer & Pintrich, 1997) citing an overreliance on survey data and researchers' failure to consistently align their findings with the hypothesized model. Instead, Sandoval and other theorists (e.g., Hammer & Elby, 2003) argued that a more robust explanation of EC must be grounded in the situated cognitive

perspective championed by Lave (1988; Lave & Wenger, 1991) and others. This situated cognitive perspective is at the core of epistemological resources models of EC.

Epistemological resources models of EC. The dimensional models proposed by Schommer (1990), Hofer and Pintrich (1997) and others added to the existing developmental models of EC by proposing that epistemic thought may exist in multiple dimensions that do not necessarily develop from naïve to advanced ways of thinking in synchrony with each other as proposed in developmental model of EC. Other researchers proposed that EC is best defined using a context-sensitive definition focused upon a finer grain size (Hammer & Elby, 2002; 2003). From this perspective, EC is only interpretable in situ. Epistemological resources theorists concurred with researchers who proposed multi-dimensional set of belief models (e.g., Schommer, 1990; Hofer & Pintrich, 1997) that EC does not develop in unidimensional stages like those proposed by Perry (1969/1999) or King and Kitchener (1994). However, they rejected "unitarity" (Hammer & Elby, 2002, p.172), or the concept that an individual's EC is best described as a series of theories, traits, or beliefs, posited by multi-dimensional theorists including Schommer as well as Hofer and Pintrich. They found weaknesses with both types of prevailing EC models (i.e., developmental and multi-dimensional) where theorists described EC as a stable set of resources that are activated in all contexts.

In pursuit of refining existing models, epistemological resources model theorists (e.g., Hammer & Elby) proposed that EC is "made up from fine-grained, context-sensitive resources" (Hammer & Elby, 2003, p. 54). Epistemic resources are described as cognitive structures that may be activated in a variety of situations but are always situated in context. For example, Hammer, Elby, and colleagues have described several ways of understanding how someone

comes to know something (e.g., direct perception, transmission, or construction) and ways of understanding different forms of knowledge (e.g., hypothesis, rule, or prediction; Elby, Macrander, & Hammer, 2016). Notably, Hammer and Elby argued that EC varies not only across academic disciplines (e.g., math, science, history) but also across specific contexts within the academic discipline. For example, a student may believe that some components of history are set while others are subject to further fact finding and interpretation that may change over time. In later iterations, Elby and Hammer (2010) highlighted that epistemological resources are typically not apparent in isolation. Instead, the smallest grain-size where epistemological resources can be observed is within an epistemological frame. They defined an epistemological frame as "a locally coherent activation of a network of resources that may look like a stable belief or theory" (Elby & Hammer, 2010, p. 409).

Elby and Hammer (2010) argued that epistemological coherence exists only within a frame and is not global (i.e., it does not extend across contexts). To exemplify this, they described children's use of two different patterns of reasoning to explain their expectations of the physical world. First, many children hypothesize that an ice cube wrapped in a cloth would melt faster than one that is exposed to the air. Children sometimes explained this rationale based on a belief that cloth items (e.g., blankets or gloves) are inherently warm. In contrast, the same children understand that using a cloth item (e.g., an oven mitt) to protect one's hands when removing something from a hot oven could provide a layer of protection for the wearer thus providing not warmth but coolness. Elby and Hammer explained this discrepancy as an example of the context-dependency of thinking. In some contexts (i.e., melting ice cube), individuals

presume that softness provides warmth. Whereas, in other contexts (i.e., protecting hands from hot items), this same softness provides protection.

Elby and Hammer (2010) described individual's context-dependent thinking about knowledge and knowing as analogous to the fine-grained knowledge elements exemplified by the above example. Hammer and colleagues described an example of the fine-grained nature of an individual's EC in a previous paper (Hammer, Elby, Scherr, & Redish, 2005) when they interviewed a college student about his studying and tutoring. When preparing for his own test in physics, he focused on memorizing the text and studying each word of the homework solutions indicating a view that knowledge is held and dispensed by authority figures. However, when asked about his techniques for tutoring others, the same student specifically addressed encouraging students to build their understanding on their existing conceptions, indicating a view that knowledge is built on prior knowledge and not necessarily dispensed solely from authority figures. According to Hammer and Elby, this exemplified their conception of EC as defined by context and locally activated within an epistemological frame. According to Hammer and Elby, developmental and multi-dimensional EC theorists found evidence supporting their models of EC only because they were looking within individual epistemological frames.

Measurement and empirical evidence: Epistemological resources perspective of EC. Led by Hammer, Elby, and others, epistemological resource models have often been utilized to investigate learning in science contexts. Because of their claims that EC is purely situated in context, epistemological resource model theorists have typically not utilized traditional methods (e.g., self-report measures or questionnaires) to investigate EC (Mason, 2016). That is, researchers employing epistemological resource models are skeptical that responses to any task

conducted outside of an authentic context are indicative of a stable belief thus they deem measurement outside of a specific context as ineffectual (Sandoval, 2012). Instead, they used observations of teaching and learning or interview methods to measure EC in natural settings. These results are often presented as case studies like the one described by Elby and Hammer (2010) in which they analyzed exchanges between eighth-grade science students in a classroom and the types of interventions the teacher provided as students worked in groups to learn about rock formations. Elby and Hammer analyzed the speech of both the students and the teacher to highlight instances of students invoking several different epistemological resources in response to changes in the context. For example, as the students began the learning task they engaged with the provided reading and began writing down words they deemed important but did not understand. Elby and Hammer described the students' initial actions as demonstrating a view of the source of knowledge as propagated stuff or a transmission view (i.e., knowledge comes from authority). This view led the students to begin by simply mining information from an authoritative source. However, the teacher intervened and addressed them not at the level of content (i.e., correcting their use of terms or explaining rock formations) but at an epistemological level. The teacher encouraged them to reframe the task from gathering information from a source of authority (i.e., the reading) to instead focus on building on what they knew by using terms they understood. Elby and Hammer explained this shift as moving from a view of the source of knowledge as transmission to one of construction. Grounded in these types of case studies epistemological resources theorists have also addressed teacher development including ways that teachers can foster epistemological change (Elby & Hammer, 2010).

In addition to purely naturalistic studies, epistemological resources theorists have also varied the context to investigate the role of EC in learning. For example, Lising and Elby (2005) detailed a single college student's learning in an introductory physics course. They used video tapes of the student working in a group with peers while learning physics, recordings of her thinking aloud to answer physics problems, and interviews specifically designed to probe her EC. Notably, the video-taped group work occurred in a formal classroom setting and the follow-up problem solving and interviews occurred in a different, less formal context conducted by an individual who identified not as a physics instructor but as an education researcher. In the less formal setting, the student engaged in different kinds of reasoning (e.g., she used informal reasoning to solve problems) providing support for the role of context in determining which epistemological resources students use.

Critiques of multi-dimensional EC research. Barzilai and Chinn (2018) found value in epistemological resources models, however, they cited several challenges with these theories. First, epistemological resource theorists have not created normative arguments explaining when or why particular resources are more effective. Second, teachers are presumed to be able to recognize the best resources and framing and to lead students to activate or develop those resources in the right context. These weaknesses helped inform their respective models of EC (i.e., multi-faceted framework of epistemic thinking; Barzilai & Zohar, 2014; and the AIR model, Chinn & Rinehart, 2016) subsequently informing their definition of apt epistemic performance.

Domain generality or specificity. The types of models discussed to this point have varied in several key ways, such as, how each scholar or group of scholars described EC (e.g., as

a developmental construct or system of beliefs) and to what extent they conceptualized EC as broadly applicable across contexts (e.g., King & Kitchener, 2004) or uniquely influenced by context (e.g., Hammer & Elby, 2003). To this point, I have not specifically addressed another core question of EC research: whether EC is domain-general or domain-specific. That is, to what extent is EC different across academic disciplines. Fundamentally, the question of whether EC is domain-general or domain-specific is different from questions about the extent to which EC is situated in context. Specifically, questions about the domain specificity of EC are about the cognitions a learner has with respect to EC and the role that academic domains play in those cognitions. Whereas, questions about the situatedness of EC are about the effects of the environment on normative practices. In early work, researchers investigating EC from developmental or multi-dimensional sets of belief frameworks considered it to be domain-general. Muis and colleagues (2006) examined 19 empirical studies that addressed the domain specificity of EC and determined that there are both domain-general, domain-specific, and task-specific aspects. This view is now generally accepted (Mason, 2016).

The AIR model. The debate about how adequately developmental, multi-dimensional, or epistemological resources models capture EC coupled with questions about the extent to which EC is situated in context and whether or not it is domain-general or domain specific continue to interest researchers today. Among the modern models of EC that address these and other concerns is the AIR model proposed by Chinn and colleagues (Chinn et al., 2011; Chinn & Rinehart, 2016; Chinn et al., 2014). Seeking to expand on Hofer and Pintrich's (1997) framework, Chinn and colleagues incorporated both psychological and, building on the calls of others (e.g., Greene et al., 2008), philosophical constructs to develop a more complete model of

EC. Chinn and colleagues recognized the value of existing models, but they deemed improvements necessary due to challenges with the operationalization of existing models and research findings that indicated low predictive validity between EC factors and a variety of outcome variables (e.g., Schraw & Olafson, 2008) as well as psychometric problems with the instruments designed to measure EC (e.g., Debacker, Crowson, Beesley, Thoma, & Hestevold, 2008). Notably, Chinn and Rinehart highlighted specific reasons to incorporate more scholarship from philosophy into the existing educational models of EC. For example, the current study of epistemology in philosophy includes a focus on more than just knowledge, which has dominated the focus of educational EC models. Modern philosophical epistemology also includes investigations of theories, wisdom, understanding and evidence. Likewise, extant educational models of EC, like the ones discussed thus far in this paper, have not focused on the "methods and processes of inquiry" that "are central to the epistemic" (Chinn & Rinehart, 2016, p. 463). Chinn, Buckland, and Samarapungavan (2011) specifically identified several new developments in philosophical epistemology as informative for creating a new model. For example, philosophical epistemologists have shifted away from a concern with radical skepticism, or the idea that humans cannot attain knowledge, that was crucial to previous studies of epistemology and toward epistemological naturalism where the focus is on the processes that lead communities and individuals to form beliefs and generate knowledge. Additionally, philosophical epistemologists have begun investigating aims beyond simply knowledge and true belief (e.g., understanding). Grounded by these philosophical underpinnings, Chinn and Rinehart (2016) proposed four principles for the AIR model: "(1) EC is fundamentally social, (2) practices are central to EC, (3) EC is situated, and (4) EC is connected to ethical concerns" (p. 460).

Building on these new theoretical directions in philosophy, they broadened the definition of EC as the quest for "knowledge, understanding, useful models, and explanations" (Chinn et al., 2014, p. 19). The AIR model is complex (see Chinn & Rinehart, 2016) and not easily categorized into one of the aforementioned groups of EC theories (i.e., developmental, multi-dimensional, or epistemological resources). However, Chinn and colleagues built the AIR model based on a situated view of cognition (Sandoval, 2012) that acknowledged the role of context in knowledge and knowing. The current three component version of the AIR model (Chinn & Rinehart, 2016) was condensed from a five-component model originally described by Chinn, Buckland, and Samarapungavan in 2011.

The first of these components, epistemic aims and values, describes learners' goals as they engage with content. Epistemic aims are "goals related to developing some sort of representation of how the world is – of developing a cognitive 'take on the world'" (Chinn & Rinehart, 2016, p. 461). Some example epistemic aims are to gain knowledge or build an understanding or justification about a particular topic. Learners may begin a task with multiple aims some of which are epistemic (e.g., avoid a false belief) while others may not be epistemic in nature (e.g., to avoid putting forth effort). Learners exhibit the value component as they weigh their epistemic and non-epistemic aims and make choices about their continued behaviors. For example, when engaging with a complex topic a learner may value gaining a simple and cursory understanding over developing a robust and complex grasp of the topic. According to Chinn, decisions about the relative value of aims affect learner behavior and are directly tied to motivation. Continuing the above example, if learners value simply completing the task more than building a robust understanding, they are likely not motivated to continue engaging deeply

with the task. In reality, any learning opportunity involves multiple aims that are both epistemic and non-epistemic in nature. Chinn and colleagues' inclusion of aims and values is unique among EC models. This unique characteristic was particularly valuable for this project as there were opportunities for professors from different academic disciplines to establish different aims while wrestling with the flipped classroom problem.

Epistemic ideals, the second component of the AIR model, are the standards used to measure epistemic products (e.g., models and definitions; Chinn & Rinehart, 2016). For example, Chinn and Rinehart described scientists establishing an epistemic ideal that theories must fit with the preponderance of the evidence, not contradict evidence, generate new scholarship, and align with existing theories while being internally consistent. I predicted that professors from different academic disciplines may have different epistemic ideals based on their own training and that these ideals would manifest in the think-aloud data captured in this study.

Chinn and Rinehart's (2016) third component of EC is reliable processes. A reliable process is one that produces justifiable knowledge. For example, a process (e.g., repeated replications of a chemical reaction across multiple laboratories) that produces a quality model (e.g., the structure of a new chemical compound as judged by experts in the field) to describe a scientific phenomenon is a reliable process. Likewise, additional reliable processes may be defined by other fields of research. For example, text comprehension researchers (e.g., Braasch, Rouet, Vibert, & Britt, 2012) have inspired entire special issues (e.g., List & Alexander, 2017) about the reliable processes (e.g., integration, corroboration, and source evaluation) that allow learners to make sense of read material. Notably, reliable processes may be valid only in certain conditions. For example, visual perception is typically a reliable process for establishing facts

about nearby events, however, environmental factors (e.g., darkness or rain) may make it less reliable. Chinn and Rinehart recommended investigating what processes people perceive as reliable as part of ongoing inquiries in EC. In this project, I expected professors would utilize processes they deemed reliable from their own field when evaluating evidence presented in the articles regarding the utility of flipped classrooms. Evaluating their use of these processes helped me assess whether and how they transferred these processes, and, when they did, how effective they can be when employed to assess a pedagogical question.

Several unique attributes of the AIR model set it apart from extant models of EC. First, it is more inclusive than other models of EC because the role of epistemic aims, beliefs, and values are explicitly discussed. The incorporation of these concepts expanded upon the existing models in the field that generally only described individuals' beliefs about knowing and knowledge. Second, and perhaps most valuable for this project, Chinn and colleagues (2014) expanded on existing influential models of EC (e.g., Hofer & Pintrich, 1997) whose authors had only hypothesized about the role of individuals' EC on their learning processes and behaviors. Instead, Chinn and colleagues explicitly described reliable epistemic processes, ones that learners exhibiting apt epistemic performance in a particular academic discipline would effectively employ. Chinn et al. also clarified the connection between epistemic aims and behavior. The AIR model, more explicitly than other models, specifically addressed the role of reliable processes and normative behaviors in particular academic disciplines. For example, in history, the use of original source material and first-person accounts are important reliable processes. In other academic disciplines, such as math or chemistry, these types of processes are not as valuable. In general, Chinn and colleagues directly addressed how experts engage in EC

more than other theorists who tended to focus on beliefs instead of engagement. This description of EC in action makes the AIR model valuable for this project as I investigated whether and how EC transferred.

To date, due to its relatively new conceptualization, as well as complexities noted by Chinn and colleagues (2011), few researchers have conducted empirical research using the AIR model. Chinn and colleagues acknowledged the difficulties with measuring EC using the AIR model due to its complexity but have suggested the use of interviews as well as the analysis of learners engaged in an epistemic task as possible methods (Mason, 2016). Barzilai and Chinn (2018) specifically discussed epistemic education from a perspective grounded in the AIR model as well as a separate modern model of EC, the multi-faceted framework of epistemic thinking.

Multi-faceted framework of epistemic thinking. Barzilai and Zohar (2014; 2016) used metacognition, or thinking about thinking (Flavell, 1979), to define their model of EC, the multi-faceted framework of epistemic thinking. Barzilai and Zohar highlighted several factors that add to the complexity of their model including challenges among researchers in parsing the distinction between cognition and metacognition and the lack of a shared conceptual understanding of the intersections between EC, metacognition, self-regulation, and self-regulated learning. Unfortunately, differences in the use of terminology that persist in the field of EC research are also apparent in their model as they used epistemic thinking as the encompassing term and epistemic cognition as a subordinate element within their model. For clarity, throughout my brief review of the essential elements of this model (see Barzilai and Zohar, 2014; 2016) I will continue to use EC as the broader term. Barzilai and Zohar described EC as a multifaceted construct consisting of both cognitive and metacognitive aspects. Overall, they defined five

facets of EC that Barzilai and Chinn (2018) later summarized into four aspects when they described their own Apt-AIR framework where they defined apt epistemic performance. I will review the condensed four aspect model crafted by Barzilai and Chinn.

First, in the multifaceted framework of epistemic thinking, EC requires cognitive epistemic processes and strategies that are utilized to reason about specific sources, claims, and information. Thus, at the cognitive level, epistemic thinking requires making decisions about the epistemic status and properties of information and validating knowledge claims. These determinations are necessarily connected with achieving epistemic ends (e.g., true and justified beliefs). For example, a learner encountering information about climate change makes epistemic determinations about the validity of the information (e.g., Is this right?) and makes determinations about both the reliability of the source and how well an author's claims are supported. Barzilai and Zohar (2016) highlighted several other examples of cognitive epistemic strategies and processes including validating the plausibility of knowledge based on its consistency with other available information and understanding authors' viewpoints when reconciling disparate information. According to Barzilai and Zohar (2014), learners cannot attain metacognitive knowledge until they possess adequate cognitive level knowledge about the topic.

The remaining three aspects of the multi-faceted framework of epistemic thinking (i.e., epistemic metacognitive skills, epistemic metacognitive knowledge, and epistemic metacognitive experiences) all function at the metacognitive level (Barzilai & Zohar, 2016). Epistemic metacognitive skills include the planning, monitoring, and evaluation processes necessary to competently engage in epistemic strategies and processes at the cognitive level. For example,

learners employ epistemic metacognitive skills when they monitor and alter the types of sources they seek in an attempt to reach a true and justified belief.

Epistemic metacognitive knowledge includes an individual's theories, beliefs, and ideas about the nature of knowledge and knowing. Barzilai and Zohar (2016) incorporated both epistemic metacognitive knowledge about persons (i.e., information about an individual as a knower or information about others as knowers) and epistemic metacognitive knowledge about strategies and tasks (i.e., knowledge about how to learn that will result in knowledge; e.g., when and why to use a particular strategy) into this single facet. For example, if learners determine that evidence gleaned from peer-reviewed journals has more value than that gained from the first article returned by a Google search, they are creating epistemic metacognitive knowledge about a particular strategy.

Finally, epistemic metacognitive experiences include learners' affective and motivational responses to a cognitive task. For example, Barzilai and Zohar (2016) cited work featured in Kahneman's popular book *Thinking, Fast and Slow* (2011) that highlighted the role feelings of cognitive ease play in feelings of truth. That is, if people feel that they processed information easily they are more likely to believe that they gained a true and correct understanding.

According to Barzilai and Zohar (2016), these three components (i.e., epistemic metacognitive skills, epistemic metacognitive knowledge, and epistemic metacognitive experiences) interact with each other and with cognitive epistemic processes throughout a learning experience. For example, epistemic metacognitive skills may be informed by epistemic metacognitive knowledge while simultaneously contributing to the development of new epistemic metacognitive knowledge. Again, the multi-faceted framework of epistemic thinking is

complex and a complete explanation is beyond the scope of this project. I have introduced it primarily to provide a basic foundation for understanding the Apt-AIR framework and apt epistemic performance.

Apt epistemic performance. Barzilai and Chinn (2018) defined apt epistemic performance. The primary focus of their article was to synthesize current scholarship on epistemic education and help educators enhance their students' epistemic growth. They reviewed a range of recent efforts to enhance learners' EC, which were grounded both in the models of EC discussed thus far (e.g., epistemological resources models, Elby & Hammer, 2010) as well as those defined by different research agendas (e.g., Nature of Science, Abd-El-Khalick, 2013) and proposed their own framework, the Apt-AIR framework, to inform educators interested in enhancing student EC. In advocating for the value of apt epistemic performance as a construct they cited its connection to developing knowledge: "the goal of achieving apt epistemic performance and the goal of developing disciplinary knowledge are complementary and interacting" (Barzilai & Chinn, 2018, p. 363).

They defined five aspects of apt epistemic performance which was the primary measurable for this study. That is, at the core of this project are questions about whether and how experts' apt epistemic performance, as defined by Barzilai and Chinn (2018), transfers across academic disciplines. The first aspect, cognitive engagement in epistemic performance, describes the cognitive processes (i.e., "procedures, skills, strategies, and methods" Barzilai & Chinn, 2018, p. 367) employed during learning. Apt epistemic performance requires the selection of epistemic aims, the appropriate application of epistemic ideals to develop and evaluate epistemic products, and the utilization of reliable processes. Of particular note for this project, Barzilai and

Chinn highlighted that apt epistemic performance requires the use of domain-appropriate epistemic ideals (e.g., accuracy or coherence). The selection of domain-appropriate epistemic ideals is a prime example of the type of process that may prove important in this study as I assess whether and how participants transfer apt epistemic performance across disciplines.

The second aspect of apt epistemic performance is an ability to adapt to succeed in new situations. Barzilai and Chinn (2018) prescribed four factors that are necessary for adaptive epistemic performance. First, it requires recognition of the specific demands of a learning situation including awareness of specific task conditions where particular epistemic processes may be most beneficial. Second, in order to be adaptive a learner must be proficient with a wide variety of epistemic cognitive processes so that they can be utilized. Barzilai and Chinn cited the value of visual perception to achieve a true belief in certain situations (e.g., good eyesight and good lighting), but adaptive epistemic processes require that learners have additional cognitive processes available. The remaining two factors supporting adaptive epistemic performance (i.e., metacognitive knowledge about which epistemic processes and ideals to employ in a given situation and the ability to metacognitively regulate the use of the selected processes and ideals) serve as a segue to the third aspect of apt epistemic performance, regulating and understanding epistemic performance.

The third aspect of epistemic performance is the regulation and understanding of epistemic performance as supported by metacognitive knowledge and skills. This aspect, perhaps the most complex of Barzilai and Chinn's (2018) five, includes the use of metacognitive skills (e.g., epistemic planning, epistemic monitoring and evaluation, and epistemic control) to ensure

that a learner enacts the best epistemic processes. It also includes the selection of valuable epistemic aims and ideals to meet the desired learning goal.

Barzilai and Chinn's (2018) fourth aspect, caring and enjoying epistemic performance, describes the affective and motivational dispositions that drive a learner. Motivation and emotion are important precursors for learners to continue to seek epistemic aims. When motivation is lacking, it is unlikely that reliable epistemic processes will be utilized. Some examples of the motivation and affective dispositions include wonder, intellectual responsibility, and love of truth.

The fifth and final aspect of apt epistemic performance is participating with others. This requires that learners are able to achieve epistemic aims in a variety of social settings. This aspect goes beyond just participating with others directly. It also includes engagement in "individual processes that address social processes" (Barzilai & Chinn, 2018, p. 372) such as evaluating institutional knowledge producing processes (e.g., laboratory studies). Also, this aspect requires individuals to participate in the application of epistemic norms and criteria agreed upon within a particular field (e.g., peer review of scientific work). Exercising this aspect of apt epistemic performance requires a recognition that whereas learners have a capacity to conduct independent thinking and autonomous inquiry, the incorporation of knowledge built and maintained by a community is also essential to enhancing understanding.

Applying the Apt-AIR framework: Apt epistemic performance in action. Barzilai and Chinn's (2018) definition of apt epistemic performance is new and, thus, little empirical evidence exists. However, they provided a road map for its utilization by analyzing two classroom sessions of inquiry curriculum originally discussed in an article in *The Journal of the*

Learning Sciences (Puntambekar, Stylianou, & Goldstein, 2007). In the original study, Puntambekar and colleagues compared the performance of students in two different sixth grade classes. Both classes conducted a 10-week inquiry-based curriculum on simple machines that culminated in a design challenge. Students in one class developed significantly better understandings of the physics concepts and their connections than those in the other class. Puntambekar and colleagues attributed these differences to one teacher's ability to enhance her students' conceptual connections between both the science and engineering concepts and their own prior knowledge and experiences. Barzilai and Chinn reanalyzed this data through the Apt-AIR lens to investigate epistemic growth among the students. In this process, they assessed differences in each teachers' application of the five aspects of apt epistemic performance helping to explicate the differences in students' eventual conceptualization of the concepts. For example, when assessing how each of the two teachers addressed Aspect 1, Engaging in Epistemic Performance, Barzilai and Chinn noted that the teacher whose students formed more complete connections required her students to apply epistemic ideals. Specifically, she required that they evaluate the quality of questions they asked (e.g., how appropriate a question may be for meeting the final design goal) and to justify their answers to these questions using principles of physics. The less successful teacher did not require this and Barzilai and Chinn cited this focus on epistemic growth, whether the teacher consciously intended that as a goal or not, as one reason her students outperformed the other class.

Barzilai and Chinn also provided an example application of apt epistemic performance through two hypothetical learning activities (i.e., developing a scientific model and evaluating the trustworthiness of a website). Barzilai and Chinn then described the enactment of each of the

five aspects of apt epistemic performance (see Table 3, Barzilai & Chinn, 2018, p. 375-376) as a hypothetical learner engaged in these processes. For example, as learners evaluate the trustworthiness of a website, they enact Aspect 2, Adapting epistemic performance, when they use different evaluation criteria for different types of websites. Thus, they recognize that information obtained from a reputable news organization requires less scrutiny than information on an unvetted website. I applied these aspects and examples of their utilization to help understand whether and how apt epistemic performance transfers.

Transfer

The literature on transfer dates to the early 1900s when Thorndike (1924; Thorndike & Woodworth, 1901) first asked questions about how interventions designed to induce improvement in one mental "function" (e.g., spelling or attention) would influence performance in a different function. Thorndike crafted his research to counter the prevailing view at the beginning of the 20th century that learning challenging subjects (e.g., Latin) increased a learner's "formal discipline" or general ability to learn (Bransford & Schwartz, 1999) in a way analogous to building mental muscles. Thorndike's work showed that learners may perform well on a test to measure just learned material, but they may not transfer that knowledge to a novel situation, demonstrating that they had really learned very specific information in context. Thorndike described transfer with a focus on identical elements defined as shared, objective physical features of the environment or stimuli between two or more learning activities (Lobato, 2006). Seminal ideas from Thorndike's work, for example the role of overlapping features between learners' prior knowledge and the transfer required to solve a novel problem, continue to be a valuable part of transfer research (Day & Goldstone, 2012). In the influential publication, *How*

People Learn: Brain, Mind, Experience, and School (National Research Council, 2000), the authors defined transfer as, "the ability to extend what has been learned in one context to new contexts" (p. 51).

In order to further the discussion, I will define several key terms related to transfer research that played a role in this study. One distinction that is essential to this project is between near transfer and far transfer. Near transfer involves the utilization of knowledge or skills learned in one context in a separate but similar context. Far transfer requires that the learned skill or behavior be demonstrated in dissimilar contexts (Barnett & Ceci, 2002). This distinction occurs along a continuum such that the application of knowledge learned in biology to a different biology context is near transfer. Whereas, application of this same knowledge in history requires far transfer. Of note, researchers have not clearly defined points of distinction along the continuum between near transfer and far transfer. For instance, continuing the example from above, it is likely that applying a skill learned in biology to chemistry falls somewhere between the near transfer that occurs when that skill is used in a biology context and the far transfer required when it is employed in a history context.

Barnett and Ceci (2002) provided a framework to clarify the distinctions between near and far transfer. They described transfer using two factors. The first, content, described what transfers, and was designed to capture issues of spontaneity and specificity. Barnett and Ceci created three dimensions of content to help refine what transfers. The first, learned skill, described a continuum of specificity of what is transferred ranging from a specific procedure to a broader heuristic or principle. Performance change, the second dimension of content in Barnett and Ceci's taxonomy, described the nature of change in performance that is expected during

transfer. Barnett and Ceci suggested speed of performance, accuracy of performance, or novelty of approach as examples of measures of performance change. That is, for example, researchers could use the speed with which a learner completes a task in a novel situation as an example of transfer. Likewise, researchers could use the utilization of a novel approach to problem solving as a measure of transfer. This measure, novelty of approach, best fits this project. The third dimension of content, or what transfers, is the memory demands of the transfer task. At the lower end of the memory demand spectrum, a learner may be asked only to execute a previously learned task in a novel situation after being prompted and reminded. For example, a research study participant could be explicitly asked to apply an algorithm learned in class to solve a problem on a test. At the other end of the memory demand spectrum is a requirement that the learner independently recognize when a skill needs to be executed, recall how to do it, and execute it. In this study, experts were asked to transfer their apt epistemic performance from their own academic disciplines to understand an educational research question. Placing this task in each of Barnett and Ceci's three dimensions of the content factor from their taxonomy for far transfer (i.e., learned skill, performance change, and memory demands), I assess that participants were asked to apply a broad principle or heuristic (i.e., learned skill).

The second dimension of Barnett and Ceci's (2002) taxonomy is context, which they used to encompass when and where transfer occurs. Overall, they described five dimensions of context, however, only one, knowledge domain, is relevant for this study. Knowledge domain describes "the knowledge base to which the skill is to be applied" (p. 623). For example, knowledge domain describes whether a skill is applied in a history class or a chemistry class. Barnett and Ceci specifically noted that transfer from one science class to another would likely

constitute near transfer whereas transferring that same knowledge to an English class would constitute far transfer. This distinction of near versus far transfer is central to this project, and Barnett and Ceci's taxonomy provided the theoretical basis to describe the different groups (i.e., no transfer, near transfer, and far transfer) as distinct. However, they acknowledged that the breaks in the continuum that define near and far transfer are difficult to define. In this project, education researchers represent the no transfer group because the research being reviewed (i.e., flipped classroom research) is explicitly an education topic. Donald (1990) provided the basis to distinguish between the other social scientists (i.e., near transfer) and the traditional scientists (i.e., far transfer). She distinguished fields of study based on their criteria for truth (i.e., EC) noting differences between social scientists and traditional scientists. For example, according to Donald, professors from traditional sciences are likely to be logical empiricists and test information against theory. Social scientists may follow a similar process, but be less rigid in their application.

Another key distinction of transfer terminology is between negative transfer, the ineffective utilization of a previously learned behavior in a new situation, and positive transfer, the effective use of previously learned information (Bransford & Schwartz, 1994). For example, students who continued to utilize a successful heuristic learned to multiply whole numbers would be hindered by negative transfer if that heuristic proved ineffective for multiplying polynomials but they continued to use it. In this case, students would be better served by critically assessing the task and applying a different method. In contrast, positive transfer occurs when a learner employs previously learned information to successfully complete a new task.

Traditional or cognitive perspective on transfer. The most utilized, traditional view of transfer owes much to the cognitive views of psychology developed in the 1960s. The mainstream cognitive perspective on transfer encompasses several different strands of research including some that mirrored Thorndike's work (i.e., identical elements strand), which focused on similar features between the original learning environment and the transfer task. Researchers approaching transfer from this perspective focused specifically on similarities in either the task or the physical environment as learners attempted to transfer knowledge from one setting to another.

Another influential cognitive transfer theory is the structure-mapping strand (Gentner, 1983) where transfer is characterized as the mapping between the symbolic representations of knowledge created in the initial learning and transfer situations. In this perspective, psychological similarities (i.e., context free or cognitive similarities) are more important than similar features between the task or environment (Wagner, 2010). Both of these and other cognitive perspectives on transfer are connected through their depiction of knowledge as a system of discrete symbols where each symbol represents a separate meaningful concept that can be combined and altered according to a structured syntax. Grounded in this view of knowledge, transfer in the traditional cognitive sense is defined as the application of previously known knowledge to understanding new and similar instances (Day & Goldstone, 2012b). These and other cognitive perspectives on transfer are joined together under a single theoretical umbrella because of three key features (Lobato, 2012). First, each cognitive transfer theory requires an abstract representation of knowledge that can be effectively decontextualized to allow for transfer. This means that learners must be able to separate new knowledge from the exact

conditions and context in which it was learned for further use. Second, each is grounded on the idea that these decontextualized symbolic representations do not vary as they are applied to a novel concept through transfer (i.e., knowledge is stored in such a way that it can be applied without changing it). Finally, transfer only occurs if the representations learners construct of the two learning environments (i.e., initial and transfer situation) are identical, overlap, or can be clearly mapped together by an observer.

Challenges to the traditional perspective on transfer. Throughout the years, numerous researchers have questioned the value of continued focus on transfer research. Beginning as early as the late 1800s, there were challenges to existing conceptualizations of transfer research as defined by Thorndike (Lobato, 2006; 2012). For example, in the late 1800s, Höffding (1892) argued against a focus on task similarity between the initial learning task and the transfer task. Instead, mirroring concerns that would be raised almost 100 years later, he advocated for a focus on psychological similarities between the tasks. In spite of initial concerns, the next era of true challenge to transfer research began in the 1980s and 1990s following the increase in interest in situated cognition attributed to Lave (1988) and others. This interest in situated cognition inspired challenges to existing assumptions about knowledge, knowing, and learning that framed the traditional view of transfer, highlighting concerns about whether or not knowledge could truly be decontextualized. Theorists who defined the situated perspective (e.g., Lave, 1988) described cognition as inseparable from the context in which it was acquired and practiced. That is, whereas traditional cognitive transfer theorists were focused on decontextualizing knowledge and asking questions about whether and how these decontextualized components could be utilized in new ways, situated theorists argued that there was no value in decontextualized

knowledge. Lave's (1988) book *Cognition in Practice* focused specifically on the role of transfer from a situated perspective in math instruction and was an influential publication. Situated cognition theorists questioned the premise behind transfer research and whether it was possible. Separately, Carraher and Schliemann (2002) challenged the notion that transfer is a problem that must be explained. Instead, they proposed transfer as its own theory of learning. They encouraged learning theorists to abandon the quest to understand transfer as a direct utilization of prior knowledge to solve a novel problem and, instead, to focus on reframing how learning theorists cope with how experience and prior knowledge contribute to learning.

Despite these ongoing challenges, the basic concept of transfer as the application of prior knowledge to new problems remains at the root of all learning theories (Lobato, 2006). Interest in transfer research has continued in recent years with some additions and changes inspired by push back on the concept of transfer from multiple theorists (e.g., Carraher & Schliemann, 2002). Transfer was included in the president's address to the American Educational Research Association in 1999 as one of the six most basic areas for education research in the 21st century (Schoenfeld, 1999). Other indications of continued interest in transfer include The National Science Foundation's funding of two conferences specifically focused on transfer research in the early 2000s, a three-year strand in the *Journal of the Learning Sciences* focused on transfer that began in 2006 (Lobato, 2006), and a recent special issue of *Educational Psychologist* (Goldstone & Day, 2012a) devoted to new ideas about transfer. Likewise, transfer research in numerous contexts appears throughout *How People Learn II: Learners, Contexts, and Cultures* (National Academies of Science, Engineering, & Medicine, 2018) including, for example, focuses on transfer in workplace learning, discussion of the role of transfer in self-regulated learning, and

transfer as it occurs throughout the lifespan. In 2002, Barnett and Ceci argued succinctly that there are two reasons transfer had been a focus for researchers over the preceding 100 years. First, theoretically transfer is an important measure of models of learning. The efficacy of learning models is often determined by learners' ability to exhibit the learned behavior in new context. The second compelling reason that transfer has continued as a research focus is more practical. Students' ability to transfer is at the core of formal schooling as much of what is taught will not be directly required for future success. Instead, formal education is grounded in the idea that students will be able to apply both cognitive and metacognitive skills, as well as facts and reasoning ability they acquire in school to new situations. Thus, transfer, and the ability to measure it, remains essential to the formal education system.

Alternative Perspectives on Transfer. Today, multiple alternative approaches to transfer are prevalent in the research. Newer frameworks for investigating transfer (e.g., preparation for future learning; Bransford & Schwartz, 1999; actor-oriented perspectives; Lobato, 2012) make use of differing models of cognition, definitions of knowledge, and methods for capturing transfer than the traditional cognitive perspective (Lobato, 2006). An understanding of these additional frameworks of transfer informed this project, and provided additional grounding for coding the transfer of apt epistemic performance. Numerous additional models and definitions of transfer exist. A complete review is beyond the scope of this project, however, an introduction to two additional approaches as well as a brief review of some of the additional metaphors and definitions of transfer that researchers have proposed is valuable. Specifically, I focus on two of these perspectives a) the preparation for future learning and b) the actor-oriented transfer perspective.

The preparation for future learning perspective on transfer. The preparation for future learning perspective was carefully delineated from the traditional view of transfer by Bransford and Schwartz (1999). They did not take credit for inventing this perspective, instead they claimed that it existed in the literature but was not well defined. Specifically, they expressed concern that researchers in the field, including themselves, too frequently intermingled the cognitive view of transfer with a different conception, which they labeled preparation for future learning. They were instrumental in clearly differentiating the traditional view of transfer as "the ability to directly apply one's previous learning to a new setting or problem" (i.e., the traditional cognitive perspective; Bransford & Schwartz, 1999, p. 68) from a view of transfer as a measure of one's ability to "learn in knowledge-rich environments" (i.e., the preparation for future learning perspective; Bransford & Schwartz, 1999, p. 68).

Viewing transfer from a preparation for future learning perspective allowed them to capture evidence of transfer that was often missed when researchers employed the traditional view. For example, Bransford and Schwartz (1999) argued that existing cognitive theories of transfer could be applied effectively to measure transfer as experts work in a given academic discipline, but that these existing models were inadequate to measure the transfer that occurs as novices learn and advance toward expertise. To illustrate this point, Bransford and Schwartz described a study they conducted to investigate how students' general educational experiences shaped the way they grappled with a novel task. They asked fifth graders and college students to build a hypothetical recovery plan for endangered bald eagles that would be enacted at a state level. In general, all participants produced poor plans. The college students demonstrated clearly superior writing and spelling skills, but they still failed to produce a quality product that

addressed relevant issues. Viewed through the traditional cognitive lens, where transfer was defined as the application of knowledge gained in one context to another, this demonstrated a lack of transfer for both groups from their general education to this specific topic. However, viewed through the preparation for future learning perspective college students' plans were notably more advanced including addressing the interdependence of the eagles in ecosystems, an appreciation of the role of change and history in the plan, and incorporating the possibility that multiple solutions may be necessary. These differences clearly show the benefits of the additional experience and knowledge that were transferred to this novel situation but would not have been captured by traditional cognitive views of transfer. Likewise, they stated that the predominant cognitive theories of transfer focused exclusively on the direct application of prior knowledge while ignoring the role of transfer in applying existing knowledge to future learning.

The actor-oriented perspective on transfer. A third perspective on transfer is the actor-oriented perspective. From the actor-oriented perspective, transfer is not simply the application of knowledge from one task to another. Instead from an actor-oriented perspective the focus is on how learning is generalized and how prior activities and knowledge influence a learner's actions in a novel activity (Lobato, 2012). Lobato specifically highlighted the utility of the actor-oriented perspective as a response to Bransford and Schwartz's (1999) call to develop transfer perspectives that allow for the study of smaller changes in learning that lead novices to acquire expertise. The differences between the definition of transfer from the cognitive perspective and that ascribed to it from the actor-oriented perspective are nuanced and warrant a deeper look. Lobato described five dimensions where the two perspectives differ: "(a) the nature of knowing

and representing, (b) point of view, (c) what transfers, (d) methods and (e) goals" (Lobato, 2012, p. 234).

First, with respect to the nature of knowing and representing, researchers studying transfer from the cognitive perspective defined knowledge using the symbolic representation metaphor. Cognitive perspective researchers (e.g., Gentner, 1983) acknowledged that the symbolic representations learners build of the world are not exact replicas of the world and are, in fact, subject to learner's goals and prior knowledge. However, a review of empirical transfer research (Wagner, 2010) found that researchers using both the identical elements and structure-mapping strands of the cognitive perspective of transfer research treated these representations as if they were direct interpretations of the world (Lobato, 2012). In contrast, from the actor-oriented perspective knowing is "a product of interpretive engagement with the experiential world, through an interaction of prior learning experiences, tasks and artifactual affordances, discursive interplay with others, and artifactual goals" (Lobato, 2012, p. 234). Thus, investigating transfer from this perspective does not devalue the existing and ongoing research into transfer from more traditional cognitive approaches. Instead, it adds a new perspective providing for additional ways to capture transfer.

The second key distinction Lobato described between the actor-oriented perspective and the cognitive perspective centers around whether transfer is measured from an observer's point of view or an actor's point of view. From the traditional cognitive perspective, transfer is typically determined based on how effectively the learner performs a particular principle or strategy that was determined *a priori* by researchers and can be observed. From the actor's point of view, the efficacy of transfer is determined by whether or not prior experience shapes the

learner's actions in the transfer situation regardless of how correctly the novel task is performed. Using the actor's point of view allows researchers to capture the effects of prior learning in new situations that could be disregarded by the traditional cognitive approach.

Researchers employing the actor-oriented perspective also differ from those who employ a more traditional cognitive view of transfer when they define what transfers (Lobato, 2012). This distinction is at the crux of the difference between the two perspectives and is relevant for this study. From a traditional cognitive perspective, what transfers are well-defined strategies and actions. In contrast, from the actor-oriented perspective, transfer is more holistic. Lobato cited an example from Thompson (2011) to demonstrate the difference in the two perspectives. In the example, a sixth-grade student was asked to solve a word problem that required an understanding of the relationship between speed, time, and distance traveled (i.e., how long will it take a character to travel 200 centimeters at 25cm/sec?). The student solved this problem by drawing a series of 25 cm line segments until she reached 200 centimeters and then counted the segments to determine the answer. In a follow-up question, the same student was asked to complete a related task with a different unknown variable (e.g., at what speed must a different character move to travel 400 centimeters in 9 seconds). The student struggled with the second problem, first attempting to create line segments to represent speeds then attempting a guess-and-test strategy to resolve the problem to no avail. From the traditional cognitive perspective, where what transfers is characterized as strategies and actions, this student did not effectively transfer prior knowledge because she attempted to use different strategies (i.e., a line segment to represent speed then guess-and-test) but did not solve the problem. However, from the actor-oriented perspective, where researchers focus on a more holistic view of what transfers, the students'

attempts to employ her knowledge about the relationship between speed, time, and distant traveled did constitute transfer as they demonstrated the value of her prior knowledge and showed progress toward understanding the concepts.

The fourth distinction between the predominant cognitive view of transfer and the actor-oriented perspective is in the methods used to investigate transfer (Lobato, 2012). Researchers investigating transfer using traditional cognitive views have typically followed a particular methodology. Students are taught something in one learning situation and then asked to duplicate some part of the newly learned skill on a new task or in a new environment. For example, students in an experimental group may be taught how to solve a particular type of multiplication problem using a simple equation then asked to apply that knowledge to a word problem. Then researchers compare the performance of students in the experimental group to the performance of students in a control group who did not receive the initial instruction. If the experimental group performs better than the control group, then this is cited as evidence of transfer. If they do not outperform the control group, then no transfer is observed. This method is defined purely from an observer's perspective. In this way, researchers predetermined the best solution to the problem and judged the presence or absence of transfer purely on whether or not the experimental group demonstrated this solution. In contrast, researchers employing the actor-oriented perspective of transfer typically rely upon qualitative methodologies and inductive coding rather than on an *a priori* list of correct solutions (Lobato, 2012). Coders employing an actor-oriented perspective typically analyze data to infer learners' ways of thinking and comprehending, in addition to how they make meaning of the novel task. Researchers search for conceptual connections between previously known material and learners' performance on a

transfer task. This is necessary because the reasoning and meaning making exhibited by learners during transfer cannot always be anticipated. This methodology was informative for this study.

The final difference that Lobato (2012) highlighted between the actor-oriented perspective on transfer and the traditional cognitive view was researchers' goals. Often, the primary goals of mainstream transfer research conducted through a traditional cognitive lens are to capture whether or not transfer occurs, which kinds of knowledge transfers, and the pedagogical methods that encourage transfer. Researchers employing the actor-oriented perspective assume that learners transfer some of their existing knowledge to new tasks. Led by this assumption, they instead focus on understanding the nature of the connections learners make between what they know and a novel transfer situation. That is, they are interested in the nature of how they generalize their knowledge to new situations regardless of the accuracy of whether or not they obtain an agreed upon solution. Continuing the example from the previous paragraph, researchers approaching transfer from the actor-oriented perspective would be less interested in whether students in the experimental group performed better on the word problem than the control group. Instead, they would be interested in qualitative differences in the ways that learners approach the problem allowing more space to capture evidence of transfer.

Transfer summary. Through over 100 years of scholarly interest in transfer, multiple theories have been created to describe and investigate it. Major questions remain throughout the field including such complex and foundational ones as whether transfer exists, how it is best defined, what transfers, and whether far transfer is even possible. Despite these core, unanswered questions a robust and continuing field of study exists around transfer. Here, I have briefly introduced three influential theories of transfer each of which has its own merits. Following the

example set by Greene and colleagues (under review), I did not select a single theory but instead utilized each in my investigation to further my understanding of the data and enhance my ability to capture transfer.

Transfer and Apt Epistemic Performance

After a comprehensive review of the literature, I found only one project (Greene et al., under review) that asked specific questions about whether and how EC and apt epistemic performance transfer. This is not surprising given the complexity of both fields and the challenges presented by the growing beliefs across fields that transfer, EC, and knowledge in general are discipline-specific. Despite the challenges with studying these topics, there are clear indications of the value of understanding EC due to its direct links to academic performance and the role apt epistemic performance plays in learning in the complex digital environment that learners face today (Greene et al., 2018). Likewise, despite ongoing challenges to transfer research such as questions about whether far transfer is even possible (e.g., Lave, 1988) and proposals from some researchers (e.g., Carraher & Schliemann, 2002) that transfer should be its own learning theory, transfer research remains a critical component of learning theory.

I utilized Barzilai and Chinn's (2018) aspects of apt epistemic performance and their provided applications of the model as guides for analyzing my data. Findings from Greene and colleagues (under review) were instructive for my analysis. For example, Greene and colleagues found evidence of positive transfer when participants employed reliable epistemic processes from their own academic disciplines to assess the replicability crisis in psychology.

Flipped Classrooms

Flipped classroom pedagogy is new and interesting but questions remain about whether "the concept is truly revolutionary or simply 'a new fad'" (Yilmaz, 2017, p. 251). Simply, instructors employing flipped classroom pedagogy utilize a mix of technology driven (e.g., online video) and traditional content (e.g., textbooks) to pull the content attainment portion of a course, which was traditionally done in class (e.g., through lecture), forward to before class. Then, instructors use class time to facilitate activities and discussion (Jensen et al., 2015). Researchers have found evidence supporting positive outcomes for students in flipped classrooms. Hao (2016) reviewed several of these studies and found evidence for higher learner satisfaction (Baepler, Walker, & Driessen, 2014), higher scores on academic assessments (Forsey, Low, & Glance, 2013; Wilson, 2013), higher course evaluations, increased levels of cooperation, and better attitudes toward both the course and the instructor (Wilson). Despite these positive findings, other researchers have questioned the value of flipped classrooms, reporting findings that indicate gains associated with flipping the classroom may be better explained simply as the result of an increase in active learning that could be achieved in other, more efficient ways (e.g., additional teaching materials or peer instruction; Jensen et al., 2015). Researchers have also investigated the contributions of other constructs to the efficacy of flipped classrooms, including e-learning readiness (i.e., preparedness to learn online including for example time management skills, goal directed behaviors, high self-efficacy for learning with technology; Hao, 2016; Yilmaz, 2017), and raised questions about the effects of socio-economic status on access to the necessary technology to support flipped classrooms including the possible reduction of the benefits for students who lack access to the materials because of financial

constraints (Jensen et al., 2015). Researchers have investigated the role of flipped classrooms in undergraduate classrooms across the spectrum of academic disciplines including social sciences (e.g., sociology; Forsey et al., 2013) and math, (e.g., statistics; Wilson, 2013).

Summary

Prior to the creation of the two modern models of EC that ground this project (i.e., the AIR model and the multifaceted framework of epistemic thinking), scholars created numerous models to describe EC (e.g., developmental or multi-dimensional set of beliefs) that continue to be utilized today. These scholars disagreed on several key concepts, notably: (1) whether EC develops as a unitary concept or as a disparate set of beliefs about knowledge and knowing; (2) the extent to which EC is domain-general or domain-specific; and (3) whether, and to what extent, EC is situated in context or stored in a user's memory. The AIR model and multifaceted framework of epistemic thinking are modern models of EC grounded on previous work in EC but also incorporating multiple outside perspectives (e.g., philosophical epistemology) and constructs that were downplayed or ignored by previous models (e.g., motivation or metacognition). Building on these two models, Barzilai and Chinn (2018) defined five aspects of apt epistemic performance that are the key elements of this study.

Likewise, following more than 100 years of scholarly interest in transfer, multiple models describing it have been created. For this project, I selected and reviewed three (i.e., traditional cognitive perspective, actor-oriented transfer perspective, and the preparation for future learning perspective) perspectives that guided my search for the transfer of the five aspects of apt epistemic performance as experts investigated a complex problem. I analyzed the data for evidence of the transfer of apt epistemic performance from each of these three perspectives.

An enhanced understanding of whether and how EC transfers will be valuable for scholars and practitioners interested in learning, epistemic education, EC, and transfer. Given the expanding role of flipped classrooms as a pedagogical option for college professors across academic majors and the nascent state of research into the efficacy of this method, flipped classroom research provides a ripe opportunity to investigate the EC and transfer of apt epistemic performance of University professors. These questions are particularly interesting because it is expected but not guaranteed that professors have some knowledge of pedagogy. Regardless of their understanding of pedagogical practice they are required to teach and this investigation of whether or not they are able to employ the apt epistemic performance developed in their own disciplines to understand a teaching related problem is valuable for both scholarly and practical purposes. I investigated the following research questions.

Research Questions

1. How do education experts differ from other experts in the EC they employ to evaluate education research?
2. To what degree and in what ways do experts from outside of education transfer their apt epistemic performance to the evaluation of education research?
3. If apt epistemic performance transfers, is this transfer generally positive (i.e., beneficial to achieve a complex understanding of the topic) or negative (i.e., a hindrance to achieving a complex understanding)?
4. What are the differences between apt epistemic performance when it is transferred between similar academic disciplines (i.e., near transfer; e.g., social sciences to

education) when compared to transfer between less similar academic disciplines (i.e., far transfer; e.g., hard sciences to education)?

CHAPTER III: Methods

This study was based on think-aloud protocol data gathered from nine university professors representing expertise from three different types of academic disciplines (i.e., education, other social sciences, and natural sciences). I used inductive coding based in grounded theory to look for themes in the ways which members of these groups transferred the apt epistemic performance they demonstrated in their own fields to answer questions about an education topic. The reported findings may benefit researchers investigating EC as well as transfer.

Participants

A total of nine University professors were recruited from two colleges on the east coast. These sites were chosen based on convenience and both allowed access to faculty. Both institutions were necessary because this project required access to professors from a variety of academic disciplines including education, however, because the smaller of the institutions did not offer an education major only a small number of faculty members hold doctorates in education or similar fields. Participants were recruited to ensure there was at least one faculty member from each institution in each group (i.e., no transfer, near transfer, and far transfer) to alleviate concerns that differences between groups could be caused solely by institutional affiliation. Participants were recruited via e-mail using an IRB approved message. First, I selected academic departments at both institutions where faculty members' expertise was likely to fit one of the required groups (e.g., physics departments for the far-transfer group). Then I screened each faculty member's publicly available profile on their respective institution's

website to ensure they were available (e.g., not listed as on sabbatical) and that they did not publicize expertise in flipped classroom pedagogy. I e-mailed every faculty member in each department, and the first to respond were included as participants with one exception: one participant was specifically recruited to fill a gap in one of the groups. Many faculty members were contacted before enough responded that they were willing to participate. For example, at one institution where five participants eventually volunteered, I contacted every available faculty member from six departments which included more than 190 professors before getting the necessary quantity of participants. Recruited participants were organized into three groups: (1) no-transfer (i.e., three educators); (2) near-transfer (i.e., economist, political scientist, and historian); and (3) far-transfer (i.e., life-scientist and two chemists). Participants' years of teaching after obtaining their PhD ranged from less than one to 43 (M = 12.22, SD = 13.1)

I asked professors to contribute to a research study on knowledge and expertise across academic disciplines by spending approximately 90 minutes of their time reviewing research articles while being audio recorded. The post study questionnaire included questions to gather information about participants' prior knowledge of flipped classroom pedagogy. Two of the nine participants indicated the departments in which they worked were promoting flipped classrooms although neither of them had specifically engaged in the practice. The remaining seven participants indicated cursory familiarity with the concept. Three of those seven indicated they were not familiar with the term, but the concept was familiar. Participants were not compensated for participating.

Protocols

I used think-aloud protocol (TAP; Ericsson & Simon, 1993; Ericsson 2006) to investigate professors' transfer of EC. TAPs require that participants speak continuously throughout data collection. Participants read aloud and described their actions as they reviewed the material. Ericsson and Simon established the use of verbal report data to study cognition arguing that it provided better insight than other methods (e.g., surveys) and was not disruptive to learning. Overall, Ericsson and Simon defined two types of verbal report data that differ primarily in terms of when the data is gathered. The first type, retrospective reporting, typically involves interviews or questioning after a learning task. TAPs, in contrast, gather data as the cognition occurs. Ericsson argued that TAPs can help to alleviate methodological concerns with gathering retrospective data such as learners' failure to accurately remember details of the learning event. Despite concerns that thinking aloud during learning could interfere with cognition, subsequent research, including a meta-analysis (Fox, Ericsson, & Best, 2011), has generally found TAPs not to be intrusive. Because thinking aloud requires that participants verbalize their thinking as it occurs it allows for the capture of cognition concurrently as opposed to other methods which may require individuals to recall their actions and thoughts.

Materials

I have selected several different types of publications to serve as the core material for this study. My goal was to select articles from a variety of source types (e.g., informal or formal) that simultaneously demonstrated varying viewpoints on the flipped classroom pedagogy, represented a diverse array of topics of interest (i.e., statistics, biology, higher education in general) and represented a variety of depths of analysis (e.g., layperson to peer-reviewed). First, I selected an accessible and influential piece (i.e., Berrett, 2012) that, according to an introduction published

on the Vanderbilt University Center for Teaching website (Brame, 2013), was supportive of flipped classroom pedagogy and helped to drive interest in the topic. This relatively informal piece was published in *The Chronicle of Higher Education.* Next, I included a guide to flipped classrooms by Brame (2013) published on the Internet by the Vanderbilt Center for Teaching that served as an informal introduction to the topic including the theoretical basis, a brief literature review indicating the efficacy of flipped classrooms, and a list of the key elements of a flipped classroom course. Next, I selected a peer-reviewed study (Wilson, 2013) that extolled the values of a flipped classroom for teaching undergraduate statistics. Finally, in order to offer a counter balance to the pieces extolling the virtues of flipped classrooms, I selected a study (Jensen et al., 2015) that attributed the gains associated with flipped classrooms to active learning, which Jensen and colleagues argued could be achieved more efficiently using a traditional classroom arrangement. Specifically, they noted the increased resource demands on both teachers and students as factors to consider when opting to employ a flipped classroom.

Procedure

I followed the procedures established by Greene and colleagues (under review) and described step-by-step in Appendix B. The project was approved by the IRB at both institutions where data was gathered. Participants conducted the learning task in their office or in another quiet place of their choosing. I began each session by introducing myself and the project as I setup an audio recording device. After this, I asked them to read and sign a consent form. Next, I explained think-aloud protocol, while closely following the script in Appendix B, and the tasks participants were asked to complete throughout the learning process. Then I provided an opportunity for participants to ask questions. I told participants they should be speaking

throughout the task and that if there is a pause of five seconds or more, I would gently remind them to, “Please keep talking.”

After I gave the participants an opportunity to ask any questions, I presented them pre-selected flipped classroom materials. In order to facilitate the learning task, I brought several printed items to each session. First, I brought a copy of each of the selected reading materials. Additionally, I brought a separate sheet of paper that clearly showed the entire learning task. The printed learning task was available for participants to review throughout the process. I read them the below scenario that remained visible to them on a piece of paper throughout the learning task:

“One of your colleagues is beginning her third-year teaching undergraduates in a department and subject similar to yours at a university considered one of your university’s peers. She has autonomy over her course. She has read about flipped classrooms and wants to know your opinion. She sends you these four publications and asks you to review them before answering questions. You may take up to an hour to complete the task or stop at an earlier time of your choosing. Please be certain to answer each of these questions during the task."

1) What do you think of flipped classroom pedagogy?
2) What are your thoughts on the quality of the arguments presented in these articles?
3) Do you think she should implement flipped classrooms in her new course? Why or why not? Please justify your answer using reasons and evidence.

Professors were afforded up to one hour to complete the task. I made this time limit clear to them in the verbal instructions I provided and clarified they could terminate the session and recording at their discretion (see Appendix B). Throughout data gathering I conducted memoing

in order to record pertinent data about the session for future analysis. In order to avoid influencing data collection, I waited until after recording was completed to ask professors to complete the questionnaire (see Appendix A), which included a question asking participants to describe their prior knowledge about flipped classroom research. It also included questions designed the gather the following information: 1) participants' academic discipline, 2) which courses they teach, 3) how often they teach each semester, 4) their typical pedagogies, and 5) what factors they consider when reviewing research articles for quality. Finally, the questionnaire contained demographic information including years of teaching experience after completion of a Ph.D.

Data Storage, Transfer, and Security

Identifiable information and digital files of audio recordings were stored securely. I utilized UNC-provided services including UNC e-mail and OneDrive storage capabilities. I had all the recorded audio transcribed for further analysis using a reputable outside transcription service and worked closely with the internal review board at both institutions to ensure that the proper non-disclosure agreements were in place with the transcription service.

Data Analysis

Assisted by a colleague, I utilized a similar process to the one employed by Greene and colleagues (under review) to conduct in-vivo thematic analysis of each transcript. We conducted inductive coding using the constant comparative method and memo writing throughout based in grounded theory (Glaser & Strauss, 1967). The second coder and I used our knowledge of Barzilai and Chinn's (2018) Apt-AIR Framework as well as knowledge of existing literature on transfer to guide the selection of valuable codable segments, however, we did not begin with an *a*

priori list of codes. Other researchers have used an *a priori* list of codes in the past, however, most of their research was not focused on either experts or the transfer of EC. These differences in the level of expertise of my participants as well as the context led me to utilize in vivo coding instead of a codebook. We were attentive to the potential for other literatures (e.g., multiple source use; Braasch, Bråten, & McCrudden, 2018) to inform our findings and codes (e.g., integration, corroboration, and source evaluation).

The coding process took place in multiple stages. First, I determined that coding would take place by group (i.e., no-transfer, near-transfer, and far-transfer) to allow for comparisons to be drawn both within and between groups. Each coder reviewed each of the three individual transcripts in each group and selected segments that highlighted participants' apt epistemic performance. Each coder generated a list of quotes and associated aspects of apt epistemic performance. Simultaneously, each coder generated a list of themes that represented the apt epistemic performance of the group, and noted differences between the groups. For example, both coders began by assessing each of the transcripts from the no-transfer group for evidence of apt epistemic performance. Before reconvening, each coder produced two items: 1) a list of important segments with notes describing how the material represented an aspect of apt epistemic performance, and 2) a draft list of potential themes that described the apt epistemic performance demonstrated by members of the group. Then the two coders met to compare their individual assessments of themes for both individuals and for the no-transfer group, synthesize any themes, and reconcile coding before reviewing each transcript in the no-transfer group again. During the second round of coding the no-transfer group, both coders specifically reviewed participants' answers to the prompted scenario questions and focused on reviewing individual

differences within groups. After completing the second round of coding of the no-transfer group, both coders moved on to the near-transfer group. The same pattern was followed for the each of the transfer groups, but during each subsequent coder meeting differences between groups were also discussed. For participants from the two transfer groups (i.e., natural scientists and other social scientists), coders looked for examples of the transfer of knowledge, skills, and actions from professors' academic disciplines as they assessed education research. Finally, we met to finalize themes across transcripts to highlight results. After the overall themes were determined based on input from both coders, I returned to each transcript individually to ensure that all participants' codable segments were included and that the final results included all quotes that were relevant to the research questions and the selected themes. We worked to ensure that participants' cognitions were reflected in the data in order to avoid reflecting our own biases in the findings.

Positionality

I am a Caucasian male graduate student who has studied and researched EC as a member of a research lab for over four years. I have conducted EC research including participating in peer-reviewed conference presentations, publishing in peer-reviewed journals, and engaging as a co-author of book chapters. Additional coding was completed by a fellow graduate student who is also a Caucasian male. He has conducted EC research for more than two years as a member of the same research lab as me. Together, we remained open to the data and tried to avoid biasing our findings based on any expectations. In order to aid in this process, I did not share example findings or potential contributions section with the second coder prior to our analysis.

CHAPTER IV: Results

With assistance from a second coder, I analyzed whether and how experts from a variety of academic disciplines transferred their apt epistemic performance to assess a complex problem in education. Participants were organized into three groups: the no-transfer group (i.e., education experts), the near-transfer group (i.e., other social scientists), and the far-transfer group (i.e., natural scientists). Our analyses revealed several themes. There were differences in the way that each individual dealt with the problem, but, with a few exceptions, findings will be presented by group due to the homogeneity of their responses. Of note, one member of the near-transfer group, the Political Scientist, focused almost exclusively on thoroughly reading the articles in the allotted time and produced few codable examples of epistemic performance. Despite this, evidence from this participant will be included in findings for the near-transfer group where appropriate. All segments of each transcript that provided support for one or more of the selected themes will be presented below. Overall, these segments made up a relatively small portion of the verbalizations made by participants throughout the data gathering process as the majority of the verbalizations were reading. There was variation among the participants both in terms of the extent to which they used the maximum allotted time of one hour and the percentage of time they spent reading the material as opposed to commenting on it. Two other sources of data were captured. First, I took notes during the data gathering sessions to help guide future analysis of the transcripts, however, these notes did not provide additional insights that were not accurately captured in the transcripts. Second, two of the participants, both members of the no-transfer (i.e.,

educators) group, requested permission to write directly on the articles as they reviewed them. These notes were preserved as data but did not provide additional information.

We analyzed the data for themes that best described the transfer of apt epistemic performance using three lenses (i.e., traditional transfer theory, actor-oriented theory, and the preparation for future learning perspective). During analysis, it became clear that due to the methodology employed in this study, the preparation for future learning perspective was not as valuable as initially hypothesized. Specifically, this particular lens was not as useful as predicted because it requires that learners demonstrate transfer after encountering material for a second time. However, in this study, participants were only observed at one point in time, and there was no clear evidence that reading one article prepared participants to read subsequent articles. The preparation for future learning perspective would have provided additional value if there were additional opportunities to engage with the same participants as they encountered education research again.

The themes grouped into three overarching categories: 1) Evidence of positive transfer with distinct differences between the no-transfer group and the other groups (i.e., near-transfer and far-transfer) in the depth of analysis; 2) Common and group-specific evidence of positive transfer; and 3) Evidence of negative transfer. Sub-themes were found under both theme 1 and theme 2. Evidence in support of each theme will be presented below, organized by theme and sub-theme with direct quotes from each participant supporting each theme and connections highlighted between participants' statements and the aspects of the Apt-AIR framework they represent. Selected portions of the same evidence are also presented in tabular form in Appendix C for easier viewing.

The themes linked directly to my research questions. First, with respect to research question 1 (i.e., How do education experts differ from other experts in the EC they employ to evaluate education research?), differences between education experts and experts from other fields appear in theme one and theme three. Specifically, in theme one, evidence of differences between the depth of analysis and complexity of source evaluations employed by education experts (i.e., members of the no-transfer group) and members of the near- and far-transfer groups will be presented. With respect to research question 2 (i.e., To what degree and in what ways do experts from outside of education transfer their apt epistemic performance to the evaluation of education research?), themes two and three contain evidence of the transfer of apt epistemic performance. Themes two and three explicitly address research question 3 [i.e., If apt epistemic performance transfers, is this transfer generally positive (i.e., beneficial to achieve a complex understanding of the topic) or negative (i.e., a hindrance to achieving a complex understanding)?] as I divided the evidence of transfer into positive and negative aspects. Finally, evidence shown across all three themes will specifically address research question 4 [i.e., 4. What are the differences between apt epistemic performance when it is transferred between similar academic disciplines (i.e., near transfer; e.g., social sciences to education) when compared to transfer between less similar academic disciplines (i.e., far transfer; e.g., hard sciences to education)?]. I will support each of these findings with direct quotations and analysis linking the themes to Barzilai and Chinn's (2018) five aspects of apt epistemic performance. Throughout, I will use italics to indicate when participants were reading directly from a provided article.

Positive Transfer: Differences in Depth of Analysis of Conclusions and Source Evaluation

Two themes exemplified positive transfer but with a distinct difference in the depth of analysis demonstrated by the no-transfer (i.e., education) experts and members of the other groups (i.e., near-transfer and far-transfer). First, based on their analysis of the provided readings and their prior knowledge and experience, all participants reached a similar answer to the final and most significant prompted question about whether or not the hypothetical colleague should utilize a flipped classroom, (i.e., "Do you think she should implement flipped classrooms in her new course? Why or why not?"), but the depth of analysis displayed by the no-transfer group members was qualitatively different than that displayed by other participants. Second, all participants conducted some level of source evaluation, but the depth of source evaluation demonstrated by the no-transfer (i.e., education) experts demonstrated specific knowledge of the field of education that distinguished their source evaluations from those conducted by their peers in the near- and far-transfer groups.

Conclusions to primary question.

All participants reached similar conclusions. When answering the critical prompted question of whether or not the hypothetical colleague should flip her classroom or not, all participants concurred that flipped classroom pedagogy is one of many methods that can be used to achieve the aims of student engagement and active learning they deemed crucial factors in building a successful classroom. In general, they stated that flipping the classroom should be carefully considered by the hypothetical colleague and may provide some benefits for particular students, but it is not the only way to achieve an engaged and active classroom. They also stated that if she chooses to flip her classroom, she should know that it is not guaranteed to succeed. Despite these broadly similar conclusions, there were differences in the depth of explanation

between members of the no-transfer group (i.e., educators) and the other participants that will be elaborated upon below. The commonalities of the responses demonstrated that near- and far-transfer groups did engage in transfer of apt epistemic performance, but that the transfer groups were not able to reach the same level of complexity as the no-transfer group.

As they worked to reach these conclusions, participants from across the three groups questioned the evidence presented in the readings and its generalizability to the specific prompt in an attempt to better grasp the value of the presented findings. As they did this, all participants demonstrated Aspect 1 (i.e., cognitive engagement in epistemic performance) of the Apt-AIR framework as they evaluated the epistemic value of the presented evidence. Simultaneously, members of the near- and far-transfer groups demonstrated Aspect 2 (i.e., adapting epistemic performance) as they transferred knowledge, skills, and experience developed in their own fields to assess the value of evidence in an educational study.

Evidence of transfer. In reaching their conclusions, members of the near- and far-transfer groups assessed the value of the data presented in the provided articles. Members of the near- and far-transfer groups clustered their decision criteria around several key elements including the relative costs attributed to flipping a class versus using a more traditional teaching method and the influence of class size on the effects of a flipped classroom. As an example of the transfer of knowledge about effective pedagogical techniques, the Historian, a participant in the near-transfer group, addressed questions of the relative cost of flipping a class versus continuing to use traditional methods by noting:

> Okay, this new colleague is teaching a course, uh, similar to mine. Uh, and I would say to her, a flipped classroom is worth considering, um, it's not a cure all. You certainly should

> not undertake it if for instance you've got five preparations per semester. You will be, um you will never get it all done. … Um and think also about ways that you can encourage active learning without necessarily uh doing a full-fledged flip. I would tell her that encouraging active learning is the most important thing, and flipping is not the only way to get there.

Here, the Historian showed the ability to transfer prior experience, and the knowledge gained from it, to provide a nuanced response (i.e., the underlying goal of a flipped classroom is active learning and that can be achieved using other pedagogies) to whether or not the colleague should flip her class. The Economist also provided a detailed response early in the session after reviewing only one article and cited the value of having an active classroom that would allow students to practice the types of skills that may later be tested, stating:

> Students presumably should also have practice doing the thing that we end up testing on, which is applying these things to new contexts. Uh, I ... so it's not, it's not immediately clear to me that you can't do that without, uh, in, you know, without doing a flipped classroom, um, with or without a flipped classroom. It seems like you can sort of still get at those ideas. But I am, uh, I'm coming to it from the perspective that I sort of agree with the goal.

This conclusion that an active classroom is important and may be achieved with or without flipped classroom pedagogy was also evident in the conclusions of a member of the far-transfer group, the Life Scientist, who stated:

> As far as the flipped classroom pedagogy, I think it's a technique. Um, I think the key thing is really, um, whether it's an active learning environment, um based on these

> articles, um, they seem to indicate that the amount of interaction with the students and using active learning techniques like projects within classes, or having to explain their, their homework or questions, um, is more important relatively speaking. Although the requirement to, um, have some, uh, of the information, um, um, be learned outside of class is probably helpful for that.

Then, later, the Life Scientist expanded these conclusions:

> I think in regard to whether she should implement the flipped classroom in her, in her course, I think it really, um, would have to do with how she's currently doing it. If she's standing up and um, you know, kind of lecturing without classroom interaction, then yeah that would probably be a good way for her to change the way she's doing it and get more classroom interaction. If she's already using an active learning, um, technique, where there's at least some portion of the, uh, um, knowledge attainment is outside of class, and then classroom activities are enforcing that, then it probably doesn't matter that much…

Chemist A, a member of the far-transfer group, asked about the role class size played in the presented evidence. Chemist A noted that Wilson (2013) built conclusions purely on evidence gathered in small classes, which led Chemist A to subsequently discount the value of this evidence stating:

> *My classes are limited to 25 students* (Wilson, 2013), and now I'm just thinking, oh my God, what is-what is this? This is crazy. This is ... If you're teaching 25 students, it should be interactive and engaging, and if you're lecturing 25 students, you're just wasting your time. Um ... So, then all of a sudden, I'm like, oh my gosh, this almost doesn't apply.

Clearly, Chemist A had knowledge of pedagogy and higher education, and used this to identify a potential lack of fit between Wilson's findings and the scenario faced by the hypothetical colleague.

Deeper analysis from the no-transfer group. Despite similarities in the overall conclusions, members of the no-transfer group (i.e., educators) conducted a deeper analysis of the question and produced more complex answers than participants from the near- and far-transfer groups. Members of each group asked clarifying questions and stated their answers (i.e., flipped classrooms could help but are not necessary to achieve an active and engaged classroom) were contingent on several factors (e.g., class size). However, the educators consistently posed more complex questions about the evidence provided and presented a more nuanced solution. These included questions about the role of academic discipline, attitude, instructor motivation, instructor training, and facilities on the effectiveness of flipped classrooms. For example, Educator A focused on whether findings were dependent on academic discipline as evidenced in the following quotes:

> if they're gonna compare by class, are they gonna compare by major discipline, because that might have an impact on it?

and

> As I'm reading this, it goes back to that last article wondering if this is disciplinary based because the last one was more biology.

Later in the session, Educator A again asked about the role of academic disciplines specifically with respect to the role of professors in a flipped classroom:

> *Professors remain the best source for guiding students in how to understand* (Wilson, 2013)...So again, as I'm reading this, and they're talking about the faculty, it makes me think of the difference in the disciplinary based. How much of a difference does that make?

Educator A also asked about the role of attitudes in learning:

> *Attitudes toward the course did differ* (Jensen et al., 2015). The interesting thing is I don't think they went and asked about the attitudes, which is towards the course and not towards the instructor even though they said that, that was one of the factors is the, the link and the relationship with the instructor, but most questions, they didn't ask anything about that.

These questions demonstrated Educator A's interest in the role that both academic discipline and student attitudes may play in the efficacy of flipped classroom pedagogy and demonstrated the more nuanced conclusions than no-transfer (i.e., educator/in-domain) experts built about the presented evidence. This demonstrated Aspect 1 (i.e., cognitive engagement in epistemic performance).

Similarly, Educator B wondered about the role of instructor motivation in a potential shift to flipped classrooms:

> I mean, I think I'd have to talk to her, um, to answer the third question. I think the k- key question is like, is somebody telling her t- to flip the course? Um, or is it something that she wants to do? Like, what is the nature of her curiosity?

and

> if the context for this question is like, gosh, I've been teaching, um, for two years and it's been hard and my, you know, I don't think it's working well, my evaluations are poor and I'm worried about my teaching, uh, and I'm gonna flip the classroom. Um, I don't see either in my own experience or in this literature, like an argument, like, okay, that's gonna- think it's gonna solve this for you.

Here, Educator B is demonstrating deep knowledge of why an instructor might consider flipping a class, and in particular the motivations that might drive a decision to exert the effort and endure the consequences (e.g., poor course evaluations). Educator B also noted the role of training and facilities in switching to flipped classrooms, which was not mentioned by members of the near- or far-transfer group, in each of these two examples:

> okay, so now we're talking to this, gets us to this example in, uh, intro calc. Um... So, there's *60 s- small sections of intro calc, …32 students per class*…. They meet *three days a week*. Um, *faculty* (Berrett, 2012) are well trained, that seems important to sort of point out.

and

> Which is to say that, like, if you're doing instruction that's different from the instruction that you were exposed to, um, and, uh, I would tell her that, you know, you really need to invest in that, and so, yeah, I might ask her to look, you know, what doe- what does her university- you know, what does the center for faculty excellence or, you know, what does the local teaching center have to offer?

Here, Educator B clearly demonstrated expertise and aspect 1 (i.e., cognitive engagement in epistemic performance) by carefully evaluating the epistemic value of the presented evidence.

Similar to members of the near- and far-transfer groups, Educator C also addressed the role of class size in making pedagogical choices. However, Educator C also demonstrated more complex thinking by asking more complex questions. For example, Educator C echoed the comments of Educator B by asking about the role of training in determining the value and effects of a pedagogical change to a flipped classroom. For example, Educator C noted that effectively flipping a classroom requires additional training: "That speaks to the level of professional development and, uh, faculty support that would be needed to do this in a meaningful way, um, if it were to be done." Educator C also extended this questioning to inquire about the necessity for novel assessments in flipped classrooms and the additional training required for teachers to implement those assessments:

> So, on top of this, we would need to be thinking about a different method of assessment, which is what I went back to earlier in terms of thinking about that immediate feedback. In particular, thinking about what training are teachers getting for this, particularly if their own education was didactic, multiple choice, that kind of thing. And, and what does it mean to now go in, uh, and how do you test a concept inventory. Does that go back to this idea of multiple choice? Um, even if it does get at concepts, but how are you getting students to sort of explain their strategies?

After reading Brame (2013), Educator C returned to this line of questioning stating: "I still am wondering how that assessment either training or model looks like when you're doing that immediate feedback and how do you assess concepts differently than assess correct answers." Later in the session, Educator C returned to questions about the effects of teacher training on the efficacy of flipping the classroom a third time stating:

> I think one of the things I'm reading again here is that this sort of lack of training, uh, for ... lack of pedagogical training for university faculty is a big deal. Um, and it goes back to what I was saying earlier too, that if you're not trained to do this new way, then there's going to be some serious, uh, lack of effect. And also, I'm thinking about, she's saying, it doesn't have much evidence on student evaluations. Uh, that goes back to what I said earlier about it's changing student outcomes but maybe not their experience, um, or maybe changing their experience but not for the better.

Educator C also asked questions about the role of infrastructure and assessment in determining the value of flipped classrooms:

> What's the infrastructure for it? Because, um, one of the big things I take away from reading these, is that all of those things are necessary to make flipped classrooms the benefit that they claim to be. You have to have classroom infrastructure to do the unconventional. You have to have infrastructure and support for the technology. Um, you have to have some additional training on assessment techniques that you may not be familiar with.

Educator C demonstrated deep knowledge of educational research and made connections between the flipped classroom research presented in this study and adjacent fields of research including questions about the role of technology in flipped classroom pedagogy and the potential value of related findings about technology and asynchronous learning. This demonstrated not only aspect 1 but also aspect 5 (i.e., engaging in collaborative and collective achievement of epistemic aims) as Educator C connected current readings with existing education literature to build a more complex set of conclusions:

> I'm thinking about also some of the things I hear from my colleagues that are engaged with technology about, um, synchronous and asynchronous online learning because this seems to echo some of that. If we're thinking that we can do a lot of that teaching while students are not in the classroom, then that opens new opportunities for things like distance learning, synchronous and asynchronous learning, and competency based digital learning… Thinking about somebody can work through the process at their own speed. Um, but in that case, what does it mean to do that engagement and that deeper application and synthesis of learning on an online platform, not in real time because I think that would have an effect.

Educator C connected to additional education research in an effort to better understand the effects of flipped classrooms and demonstrated expertise and complex thinking in education that was not shown by members of the near- and far-transfer groups.

The complexity and nuance of thought demonstrated by the no-transfer group (i.e., educators) as they delved into the question of whether or not their hypothetical colleague should employ a flipped classroom pedagogy was qualitatively different from that displayed by participants from the near- and far- transfer groups. Similarly, members of all groups conduced source evaluation, but the evaluations conducted by the no-transfer group were qualitatively different and more complex.

Source evaluation. Source evaluation is a critical epistemic process (Barzilai & Zohar, 2014) and members of each group demonstrated it with varying levels of depth. To varying degrees, near- and far-transfer group members highlighted questions about the evidence presented (e.g., about the effect of class size on the arguments presented in the materials), but

members of the no-transfer group (i.e., educators) asked more complex questions about the value of presented evidence and generated deeper and more refined conclusions. Participants from all three groups evaluated sources along a variety of factors including, for example, publication age and type. Sources were also evaluated based on participants' views about the quality of methodology employed, however, there were distinct differences between the near- and far-transfer groups in terms of methodological criticism that will be discussed separately.

Members of both the near- and far-transfer groups demonstrated source evaluation. Their effective source evaluation demonstrated aspect 2 (i.e., adapting epistemic performance) of the Apt-AIR framework as they transferred expertise developed in their own disciplines to assess the quality of the provided education research materials. Simultaneously, their effective use of source evaluation knowledge and skills demonstrated aspect 5 (i.e., participation in epistemic performance together with others) as they applied the norms of scientific research and publishing developed in their own academic disciplines to assess the quality of the presented information and the applicability of that information for future decision making. Members of the two transfer groups primarily assessed sources based on their age, limited knowledge about the journals or other media where the articles were published, and the appearance of the article. For example, the Historian, a member of the near-transfer group, explicitly emphasized source evaluation and demonstrated transfer while assessing the age of sources and then making a determination about the relative value of sources during a transition between reading one article and starting the next:

> Wait a minute, I'd better do something historical here and check the dates on these things. All right. So, this one's 2014. This one's 2012. This one, I can't find a date right away, but all of the dates in the bibliography are before 2012. So, all right. So, this long journal

> article (i.e., Jensen et al., 2015) is the most recent piece of material from 2015. All right, that's also I guess, that's what I would expect. Uh, as time goes on the research gets more sophisticated. That's ... that's good I guess. And then finally I'm reading one from 2013.

This statement demonstrated the transfer and purposeful use of an epistemic process (i.e., source evaluation) the Historian deemed as specifically crucial for historians. Notably, the Historian also stated that newer research should be more sophisticated and that the provided example fit that description, bestowing greater value on one article due its recency and related sophistication. Later, the Historian extended the evaluation of sources to include the type of publication stating, **"**Now, um this is a scholarly article unlike the material from the Chronicle, and it's from uh, *Life Sciences Education*."

Another member of the near-transfer group, the Political Scientist, made a similar assessment of the articles based on their source publications:

> As far as the quality of the arguments, I took the academic articles to be providing the quality of the argument. And that the Chronicle and the other article are simply summations of the literature and sort of a pop science lit review.

Notably, the classification of articles as "academic" was made by the Political Scientist as these articles were not explicitly labeled in the study materials. Earlier in the session, while transitioning from Wilson (2013) to Jensen et al. (2015), the Political Scientist stated a criterion used to classify an article as academic or not: "Actually, the last one appeared in *Teaching of Psychology*. I just took it as an academic work because of the standard SAGE type setting. This is in *Life Sciences Education*, uh, article."

The final member of the near-transfer group, the Economist, also evaluated the quality of the Chronicle article (Berrett, 2012) stating:

> This was sort of, uh, nice news type of article about, uh, that I ... you know, gave some background about what a flipping means and, um, uh, you know, it references, some studies, but without more detail on the studies it's sort of hard to evaluate them.

Later, the Economist noted a lack of specific knowledge about the academic journals that published the two scholarly articles used in the study (i.e., Jensen et al., 2015; Wilson, 2013), stating: "let's look at these two, um, uh, papers in journals that I do not know, but, um, that's not my area of expertise."

Similar to the near-transfer group, members of the far-transfer group also assessed the quality of the presented material. For example, the Life Scientist reviewed the types of sources while making a plan to review materials:

> I'm just going to look at these articles, the ones that I haven't read. And just take a look and see where they're from and what the titles are. Okay, so Flipping the Classroom by Cynthia J. Brame, CFT instructor. So, this looks like it's just something taken off the internet I'm guessing, it doesn't look like an actual journal article as far as I can tell. Another one is from Life Science Education. It looks like a journal. Article title is Improvements from a Flipped Classroom May Simply be the Fruits of Active Learning. Um, yeah that makes sense, just by the title. Um, and then another journal article from the Society for Teaching of Psychology. Um, The Flipped Class: A Method to Address the Challenges of an Undergraduates Statistics Course.

After this evaluation of the articles, the Life Scientist opted to read the article published in the *CBE-Life Sciences Education* journal due to the alignment between the journal title and the Life Scientist's own interests. Later, while discussing conclusions to the primary prompt, the Life Scientist again assessed the quality of the provided sources, stating that if an instructor was already employing active teaching techniques then switching to a flipped classroom:

> Probably doesn't matter that much based on, you know, the, these, the two principal articles that actually kind of describe specific experiments. The other ones, um, you know, the one in the Chronicle for Higher Education is mostly just sort of anecdotes.

Likewise, Chemist A evaluated the quality of sources early in the session to determine how to approach the materials, stating:

> All right, well, the other thing I notice right away is two of these articles are like, real, serious scientific articles, and two of them are like, high-level, um, summaries.

When transitioning from reviewing Brame (2013) to reviewing Berrett (2012), published in the *Chronicle of Higher Education,* Chemist A stated:

> so now I'm reading the Chronicle one, which is the other, um ... which is the other sort of ... not research article, but article that's sort of a high level, and then I guess I will say the first thing that- the first paper that was given was- it was- it looked like something like, printed out off of a blog page or something. And you know, I look at The Chronicle one, and I- and I-I'm not going to accept it at face value, but I do know it's a very thoughtful, well-researched thing. So, I guess my ... I'm sort of intrigued to see what this says and thinking that it might be a slightly more reputable source than just someone's blog.

This statement exemplified Chemist A's transfer of knowledge about the quality of articles published in *The Chronicle of Higher Education* and implied an increased trust in the value of content published in *The Chronicle* over material published in a blog post (i.e., Brame, 2013). Chemist A skimmed the Berrett article for a short time before determining that the article did not contain additional valuable information. After reviewing the titles of the two "research" articles, Chemist A decided to read Jensen et al. (2015):

> I'm reading through this, and so I'm kind of thinking like, well, I'm not- this is not a good use of my time, I'm going to stop, which I ... if I were doing this in real life, that's what I would do. All right, so now I'm looking at the two research articles and um ... and actually, the first title, um ... uh, already intrigues me because it's sort of ... I am really interested in why it works and for whom it works, and I'm always interested in making classes more engaging, and I think there are a lot of ways of making a class engaging.

This statement from Chemist A included an evaluation of the remaining two articles (i.e., Jensen et al.; Wilson, 2013) as "research" articles and the selection of Jensen and colleagues' article as the next one to read based on interest in the title. The statement also included an initial mention of what later became Chemist A's final conclusion that flipped classrooms are a "fad" and engaging students matters more than flipping. To this point, the examples of Chemist A's source evaluation have demonstrated positive transfer including continued assessments of the predicted quality of resources based on where they were published (i.e., *The Chronicle of Higher Education* versus a blog). These examples closely resemble the types of evaluations conducted by members of the no-transfer group. However, as Chemist A continued to evaluate Wilson's article the first examples of negative transfer of knowledge about research methods appeared:

"As I read the last article, the more qualitative one, I just- I-I sort of got to the end of it and thought, oh my gosh, this is not really a serious article." This dismissal of findings from qualitative sources emerged as a separate theme for members of the far-transfer group as they tended to devalue qualitative findings. More examples of the negative transfer of research methods knowledge will be discussed in the negative transfer section.

Deeper source evaluation from the no-transfer group. Members of the no-transfer group (i.e., educators) also conducted source evaluation, however, because they were grounded in their knowledge of education research and publications, their evaluations were deeper and more contextualized. Members of the no-transfer group evaluated the materials, including assessment of the age and type of publication similar to the appraisals performed by members of the near- and far-transfer groups. However, they extended these assessments by demonstrating knowledge of the literature and changes in the literature that occurred in certain time periods as well as knowledge of authors cited within the materials. Their assessments, built on expert, in-domain knowledge, demonstrated aspect 1 of the Apt-AIR framework. For example, Educator B began the task by assessing the provided readings:

> Um, so I see- I see two- I see two, um, peer-reviewed articles, I see a Chronicle article, um, by Dan Barrett, and I'm gonna just- curious about who the author is, just if I can get a sense. But I don't see that there, um, so Chronicle article by someone who I don't know. Um, if I was sitting at my desk I'd probably, I might start by googling him just to know whether it's a- whether I'm hearing about, um, you know, hearing an instructor's perspective or research perspective kind of before I decide how much energy to spend on that article. Um, and then I h- have this, um, l- looks to me like a- a nice and concise, um,

> lit review (i.e., Brame, 2013). Um, I'm just curious to see whether... It looks like it's maybe a little bit dated, um, and the, um, and the- the- the, uh, peer-reviewed article seemed to be more recent.

Here, Educator B described two articles by type (i.e., peer reviewed) and then noted that *The Chronicle* article (Berrett, 2012) was written by an unknown author. Then Educator B stated that outside the constraints of the study the typical next step would be to search for more information about the author given he (i.e., Berrett) was unknown. Then, Educator B correctly classified the fourth provided item (Brame, 2013) as a literature review. Finally, Educator B assessed the age of each source before immediately creating a plan for reviewing the articles in the allotted time. This more nuanced process of assessing the literature is notably more complex than the processes conducted by members of the near- and far-transfer groups, specifically because Educator B assessed the age, type, and authors of each article before building a plan to review them. Based on these evaluations, Educator B opted to begin by reading the literature review (Brame) for "a little grounding" before reviewing the more recently published "peer-reviewed articles as kind of updates." While reviewing Brame, Educator B demonstrated sociological knowledge of the field by describing familiarity with cited authors, "Eric Mazur, I know his name, Catherine Crouch I know their names. I know they're kinda big in this are in, um, I think Mazur's a physicist, kind of big in this area and STEM instruction." This knowledge extended to an assessment of not just the article (i.e., Brame) but also the quality of the cited authors such as: "This Mazur stuff always struck as being, um... Uh, wishful thinking, I guess" that was grounded in knowledge of the field as demonstrated by this quote: "I kinda know this literature, that part of the literature a little bit." After reviewing the Brame article, Educator B then moved to *The Chronicle* article (Berrett)

providing this assessment before reading it, “I expect it to be better written.” Here Educator B specifically compared Brame’s publication with an expectation for what would be published in *The Chronicle of Higher Education* and implied that expectation was grounded in knowledge of the publication itself. Then, after reviewing it, Educator B confirmed the perceived value of the article, stating: “So I like that Berrett article.” Following the review of Berrett, Educator B transitioned to Wilson (2013) describing it as, “the stuff that looks a little bit more academic.”

Similarly, Educator C reviewed the ages of the provided materials and demonstrated knowledge of the field while conducting source evaluation after reviewing two of the articles (i.e., Brame, 2013; Berrett, 2012) and determining what to review next:

> Um, this one is Jensen and it is spring 2015. And the other is Wilson and it's 2013. I see that there is still no date on this other one that I don't know what source it's from. All right. I am gonna read the most recent one.

Additionally, Educator C reviewed the types of materials and commented on the effect a lack of information about the sources had on their trustworthiness:

> Generally, I would want to check the source to make sure that this is from either a publication or a blog that I trust. Um, I see that it says CFT. I don't know too much about that. So, I would want to know more about the source, where it came from.

Then, when preparing to review Jensen et al. (2015), Educator C described a lack of familiarity with the specific journal but described an accepted way to determine the value of the publication: “This is published in *Life Sciences Education*. I have not heard of that, but I would probably want to know if there's an impact factor.” Later, while preparing to review the Berrett (2012) article, Educator C indicated familiarity with the publication demonstrating the role of

experience in source evaluation, “This is from *The Chronicle of Higher Education*, which is something that I am used to perusing on a day to day basis.” However, immediately after expressing familiarity with *The Chronicle of Higher Education* and implying that it was a frequently used source (i.e., perused daily), Educator C expressed concern about the age of the publication and demonstrated nuanced sociological knowledge of the field of educational research by highlighting vast changes in both the research base and technological capabilities necessary for the implementation of flipped classrooms since the article’s publication in 2012. “I do note that it's from 2012 and the flipped classroom check, uh, scholarship and even the technology to implement it has evolved significantly since 2012. So that will certainly shade my thinking about this.”

Source evaluation was an explicit component of the study as participants were asked to answer a question about the quality of the arguments presented in the materials, However, members of all three groups demonstrated source evaluation beyond that required specifically by the prompted questions. When members of the near- and far-transfer groups conducted source evaluation they demonstrated positive transfer of the apt epistemic performance built in their own academic disciplines, and, with the notable exception of members of the far-transfer group consistently dismissing qualitative findings, they succeeded in assessing the quality of the materials in ways similar to those utilized by members of the no-transfer group. The positive transfer examples from both the near- and far-transfer groups demonstrated both aspect 2 (i.e., adapting epistemic performance) as well as aspect 5 (i.e., participating in epistemic performance together with others) when experts from the near- and far-transfer groups applied epistemic criteria to evaluate the credibility of provided information. However, unlike the positive transfer

examples described in the next section, there were distinct differences in the nuance and complexity of evaluations demonstrated by the transfer groups in comparison with the no-transfer (i.e., in-domain education) experts. These differences were derived from sociological knowledge of the field, for example, as an education expert applied knowledge of education research to assess an article's value toward achieving an epistemic aim (i.e., gain knowledge about flipped classroom efficacy) not only based on the age of the article but specifically on the relative change in education research since publication. There were other examples of positive transfer where members of the transfer groups effectively utilized expertise they brought to the task from their own training and education and the difference between the in-domain experts and the transfer group members was negligible.

Positive Transfer: Common and group-specific evidence of positive transfer

Broadly, this section contains examples of near- and far-transfer participants exhibiting aspect two of the Apt-AIR framework (i.e., adapting epistemic performance) as they transferred epistemic performance built in their own academic disciplines and utilized that expertise to effectively assess education research. Evidence from members of the no-transfer group (i.e., educators) will also be presented to provide context and offer examples of how an in-domain expert dealt with this complex problem.

Both groups: Statistical knowledge. Participants from all groups exhibited levels of statistical knowledge that allowed them to interpret the findings presented in the provided articles. For members of the near- and far-transfer groups, the application of statistical knowledge built in their own academic disciplines demonstrated positive transfer (i.e., aspect 2 of Apt-Air). Two of the provided articles (i.e., Berrett, 2012; Brame, 2013;) were intended for a

lay audience and were not statistically complex. Wilson (2013) was more complex including basic descriptive statistics (i.e., means and standard deviations) as well as the results of t-tests. Jensen et al. (2014) utilized more complex statistics including independent sample t-tests, analysis of variance (ANOVA), analysis of covariance (ANCOVA), and the Mann-Whitney U-test. Participants from across all three groups (i.e., no transfer, near-transfer, and far-transfer) demonstrated the statistical knowledge necessary to interpret these results.

The Political Scientist predominantly read without additional comment throughout the data gathering processing. Despite this lack of comment, statistical knowledge was demonstrated. For instance, Brame (2013), cited statistical evidence using shorthand notations in text that the Political Scientist interpreted correctly:

> *He found that students taught with interactive engagement methods exhibited learning gains almost two standard deviations higher than those observed in the traditional courses.* And so, he gives the margin of error around the point estimates for each of the courses.

Here, the Political Scientist appropriately interpreted the shorthand notation used by Brame (i.e., "0.48+/-0.14 vs. 0.23+/-0.04", p. 2) as margins of error. Later, while reviewing Wilson (2013), the Political Scientist demonstrated knowledge of the levels of measurement (i.e., nominal, ordinal, interval, or ratio) and critiqued Wilson's labeling of a series of variables used as an in-class demonstration for statistics students:

> So, I'm going to stop reading to look at the table right now. And this is looking at the examples, uh, from the nation and the world. They got the number of homes destroyed by hurricanes as ratio. They're calling it discrete. The ... Uh, not sure about that. Uh, metro

> and state, they have forecasted high temperature, they're calling it interval and continuous. That one I agree with. I have sports, the number of goals scored on a soccer game as a ratio. They're calling that discrete. That's to the letter correct. Um, I guess they prob- ... So, it's probably less of an applied view. Um, okay. So, home and garden, hottest zip codes, nominal. Good. To your health, they reported musculoskeletal symptoms and call it nominal.

This transfer of basic statistical knowledge was essential to the Political Scientist's understanding of this material and allowed for analysis of the provided materials. Despite the fact that levels of measurement are a basic concept in statistics, this analysis demonstrated transfer of expertise that laypeople do not possess about statistics.

The Life Scientist, a member of the far-transfer group, also demonstrated basic statistical knowledge (i.e., interpreting *p*-values and histograms) while reviewing a figure presented by Jensen and colleagues (2015). Jensen and colleagues used the figure to detail the differences in unit exam scores between students taught using a flipped classroom technique and those taught using traditional classroom techniques:

> Figure two, unit exam scores. None of the differences is significant. Error bars represent 95% confidence interval. So, uh, exam one, exam two, and exam three. The grades trend downward, but are nearly equivalent for flipped and nonflipped classes. Um, for the final exam scores, um, another bar graph, uh, indicating percentage correct. Total scores, as well as low level score and high-level scores are, um, very close to equivalent for both the nonflipped and flipped.

Chemist A also demonstrated the transfer of basic statistical knowledge to the interpretation of research findings in education when interpreting results reported in Brame (2013) stating:

> *They gather data from 14 introductory physics courses, taught by traditional methods.* Um ... and I'm skimming here trying to find out what the actual data are ... yeah, and then this is the thing that I always get, that these things are always reported, um, like it's ... *exhibit higher learning gains* and then they report numbers like *.48 plus or minus 1.4 and .23 plus or minus .04,* so it's like just barely significant.

Notably, these examples from members of both the near- and far-transfer groups demonstrated the transfer of apt epistemic performance to successfully interpret the statistical findings reported by education researchers exhibiting the use of a reliable process learned in one field to understand a complex question in another field.

Both groups: Career management. Members of all three groups commented on the effect flipping classrooms may have on the hypothetical colleague's career. The majority of participants commented on the critical nature of the third year of any academic's career and the importance of accomplishing professional development milestones. For example, Educator B, a member of the no-transfer group, noted the role of student evaluations on potential tenure decisions:

> And she may actually be thinking about whether she's- she's gettin' ready for tenure. Um, I could imagine the conver- the conversation going into this conversation thinking, like, oh, I didn't think enough about the attitudinal data because I don't think it, um, reflects what I care about. Um, but if I'm on a tenure track, I actually do care about it because the

evaluations matter. Um, so I may actually, you know, also have her think about, um, if she's really f- if that's- if the context here is that she's feeling bad about her instruction or her use, it might actually also send her, um, to research about bias in student evaluations, and let her think about whether or not the students are evaluating her fairly.

Educator C also commented on professional development timelines noting:

> But nonetheless, in her third year of teaching, if she is a tenured track professor, um, she's going to be going into third year review. Um, so unfortunately you need to be thinking about what's going on for you in terms of preps, in terms of what you're getting out.

These comments about professional development were also exhibited by members of both transfer groups. For example, the Political Scientist comingled comments about the effects of flipping a class on student outcomes with the potential effects of a flipped classroom on the hypothetical colleague's career:

> Um, and so *do you think do you implement the flipped course*? Uh, if it fits with her personality and what she would like to do as an instructor, the, I think the evidence suggests that it will improve given that you want to improve it. Um, and so I think that evidence where people were flipping courses, they were able to obtain, um, better outcomes. And if the, you know, those are the ones that you are going for, then, um, there might be some risk. This is, uh, more professional and pedagogical here, but there's risk to doing something that might hurt your teaching evaluations if that's going to affect your promotion through, uh, to get tenure or something like that. So, I think the disposition of the individual that's taking the risk to take on a new type of teaching format, um, has to be cognizant of where they are in their career. And if that's sort of, uh, negative feedback

> on surveys or from students or need things that they're willing to potentially take on. Um, but as far as learning outcomes, it seems like this reorganization on a whole, um, benefits students.

Here, the Political Scientist described the potential tensions between flipping the classroom to achieve possible improvements in student outcomes with the need for a tenure track faculty member, specifically one at this relatively early (i.e., third-year) stage of a career, to receive positive student feedback. This demonstrated the transfer of knowledge about professional milestones and the pressures of obtaining tenure that exist across academic disciplines. Likewise, the Historian posited the potential challenges of implementing flipped classroom pedagogy in an important year for young academics noting:

> I would say to her, a flipped classroom is worth considering, um, it's not a cure all. You certainly should not undertake it if for instance you've got five preparations per semester. You will be, um you will never get it all done. Um and also turn your dissertation into a book, which a third-year uh college professor has to do or lose her job. Um, and especially at a university considered one of your university's peers. Um, think carefully about whether, not only whether you have the time to ... to create a flipped class.

The Historian was the only participant to specifically note that third-year professors should turn their dissertations into books, but others also highlighted the theme of balancing other career requirements with the challenges of flipping a class. For example, Chemist A, a member of the far-transfer group, noted the time commitments required of a third-year professor and the challenges the hypothetical colleague would face balancing flipping the class with other requirements:

> She's starting her third-year teaching, um so if that means third year, um, as a faculty member, um, you know, she's got a finite amount of time and energy in her hands, and so you have to balance all these things.

This demonstrated the transfer of expertise about academic careers in their respective disciplines to assess the hypothetical colleagues' potential challenges with flipping a classroom.

Near-transfer group: Methodological assessment. Members of the near-transfer group were able to successfully transfer their knowledge of research methods to effectively assess the quality of both the quantitative and qualitative methodologies used in the presented articles. This was in sharp contrast to members of the far-transfer group who exhibited positive transfer to effectively assess some research methodologies but expressly dismissed findings built using qualitative methodologies. This negative transfer will be highlighted in the next section.

Educator B's assessment exemplified the no-transfer group's review of research methodologies. For example, Educator B assessed the measures used by Jensen and colleagues (2015), demonstrating nuanced knowledge of educational research:

> I can see here when I get to the tables, that there is a- there's some kind of control, um, but you know the data are student reports of looks to me like, yea, student evaluations of course and instructor… I don't, I just don't care about those data. Like, I don't believe them.

This demonstrated Educator B's knowledge of educational research methods, specifically knowledge about the controversies regarding the validity and utility of student evaluations as an assessment tool. Later, when describing the measures used to capture student affect, Educator B said, "This is nice, I mean, this like, well-measured, the stuff that you kind of care about. Um, it's

as good as I've seen in this literature." As noted earlier, this quote supported Educator B's overall evaluation of Jensen as a source, but, here, serves as evidence of a methodological assessment. Immediately after this statement, Educator B clarified, "Well... I wish the affect stuff was a little better. It's really course evaluations." Additionally, Educator B later critiqued Wilson's (2013) data and analysis stating:

> I kind of think the empirical stuff is not particularly convincing here… Still not really very clear on what the reference category is, but I don't really care that much 'cause I don't believe this (i.e., evidence built from student evaluations) empirically, in- in any sort of really deep way. Um, so they seem to have done better, but compared to what?

Educator B continued assessing methodologies and demonstrated knowledge of research design, specifically the need to use an experimental design to establish the causal effect of flipping the classroom, at the start of a review of the Jensen and colleagues (2015) article:

> So now I'm turning to this biology article. Um... Uh, so this is a *quasi-experimental design* (Jensen et al., 2015). Um, we're comparing, and what that means here is we're comparing a... Oh, this is nice. Um, now it gets really- gets at what's... Flipping is the thing that seems to really vary. Um, like, there's one where they're using the sort of active learning, conceptual learning in the classroom but isn't flipped. Um, and one where it's- where it's flipped. Um, and you know, yeah, so this is so not surprising to me. Um... The, when you do this kind of, is it flipped or is it not flipped, um, there's really no difference in outcomes

Here, Educator B reviewed the design as explained by Jensen et al. (i.e., quasi-experimental) then expressed approval of the overall design and the effectiveness with which Jensen and

colleagues isolated the flipped classroom as the variable of interest and employed a quasi-experimental design to allow for the assessment of the effects of flipping the class. This contrasted with Educator B's previous negative assessment of Wilson's measures and demonstrated an understanding of research methodologies and the implications of those methodologies on the veracity of reported findings. Later, while continuing to review Jensen et al., Educator B provided an additional assessment of Jensen and colleagues' execution of the study stating:

> So, *two sections were put into a test control situation*. Um, same instructor, they're taught back to back, same classroom. Um, I don't have a better design but I hate this design. It's the design they use all the time.

Then, Educator B continued to assess the methodology and reviewed possible confounds such as the effects of changing two variables at one time:

> Okay, this is kind of difference in difference design. Um, 'cause you've got, um, you also have two courses where *prior semester*, um, which is all not- not flipped... Um... Although, actually they're tinkering with two things here, so the priority here, their consistent design um, but, um, there's, uh, less structure.

Finally, Educator B approved of the authors' measures and demonstrated knowledge of the research methodologies used in education research. Specifically, Educator B by assessed Jensen and colleagues' measures, stating: "*Measures of student affect*. This is nice, I mean, this like, well-measured, the stuff that you kind of care about. Um, it's as good as I've seen in this literature."

Clearly, members of the no-transfer (i.e., educators) groups were able to engage in thorough epistemic evaluation of the research design. Two members of the near-transfer group (i.e., the Historian and the Economist) effectively demonstrated positive transfer as they assessed methodologies and the value of findings generated using these methodologies throughout the articles. Both the Historian and the Economist demonstrated research methods knowledge and the ability to transfer it from their own academic disciplines to assess education research. As an example, the Historian noted:

> Uh, the um Jensen, Kummer and Godoy article is um you know, balanced uh on the one hand. On the other hand, I thought it was a well-controlled study, uh if I'm remembering it right. They had a controlled class, yeah. It was another active learning type classroom that wasn't flipped, and I thought that was a useful way to um ... to isolate flipping as a, as an effective ... uh as a meaningful variable.

In this example, the Historian noted one of the same key points selected by Educator B (i.e., Jensen et al. effectively isolated flipping as the variable of interest) and then made a conclusion about the value of the findings of that article that matched the conclusion reached by the in-domain expert (i.e., Educator B).

The Economist, another member of the near-transfer group, was very focused on methodologies and demonstrated a complex understanding of research methods. This assessment began with the first article that the Economist selected: Brame (2013). After completing a review of Brame, the Economist stated:

> This was sort of, uh, nice news type of article about, uh, that I ... you know, gave some background about what a flipping means and, um, uh, you know, it references, some

> studies, but without more detail on the studies it's sort of hard to evaluate them. Um, uh, the one that it does cite and talks about, uh, is not a randomized study. So, it was sort of hard to put much weight on this. It's mostly anecdotal evidence.

Here, the Economist clearly demonstrated knowledge of research methodologies including a preference for randomized studies as evidence for a causal effect. Throughout the session, the Economist consistently returned to methodological questions and critiqued the value of evidence based on the methodologies used to obtain it. For instance, during a review of Brame (2013), the Economist stated:

> So I mean here, there is, uh, uh, a lot of results, but, uh, um, not a lot of discussion of the methods that does not seem like, um, these were sort of, again, it doesn't seem like these were randomly assigned a, um, evaluations and, you know, um, my primary concern with evaluating these would be that you, uh, um, you know, if you are, um, if you're just letting the, the instructors choose which method to do, maybe the teachers who are sort of better and more motivated, will try something new. And so, you're actually getting better gains from people who tried something, uh, um ... who, who were sort of more motivated teachers. And it's not about the method itself. And so, you know, really randomly assigning teachers to trying these methods or not, and following through would be kind of a better way to do this. It's really unclear from reading this, whether that's what happened or not. Um, it does not seem like it. Um, so it's sort of hard for me to, uh, uh, trust the, the evidence here.

The Economist's questions about research methodologies continued with an assessment of a study cited by Brame (i.e., DesLauriers, Schelew, & Wieman, 2011) that involved a comparison

between two large-enrollment physics classes, one of which used flipped and the other not of two large-enrollment physics classes. The comparison showed large gains in both student engagement and academic performance among students in the flipped classroom when compared with those in a traditional class.

> So, this is fine, but it's still just one observation. Effectively, it's one section or not. And uh, um, uh, it's uh, uh, you know, even if you flipped a coin to pick which, uh, which section got this, it would be nice to have a lot more data, a lot more classrooms and not just one.

These critiques continued after the Economist segued to the Wilson (2013) article about teaching undergraduate statistics and specifically focused on a topic similar to one highlighted by Educator B:

> So, at this point I would like to know ... it sounds like this teacher sort of really changed a lot of things at once. It would be nice to know that at least, you know, they kept something like the final exam, the same to have some measure to see whether at least comparing students before and after that they got better on some sort of fixed, uh, fixed task. It's great that you want to reevaluate your learning objectives but maybe reevaluating them, but then keeping the other things the same. So, you have some baseline before you change. Sort of just the teaching methodology to a flipped classroom would be kind of a cleaner test, of, of the flipped classroom at the moment. It just looks like a test of whether, you know, putting a lot of effort into your course and changing a bunch of things seems to improve, uh, outcomes.

After completely reviewing Wilson's article, the Economist reiterated previous methodological critiques highlighting concerns that too many variables were changed in the classroom from one iteration to the next and that these changes prevented the creation of reliable conclusions about the efficacy of flipped classrooms, stating:

> Here we have an article that has a, you know, one experiment that seems to, uh, by comparing sort of students, uh, in prior years to students in new years where there's been a number of changes to the course, including this sort of flipped classroom that these changes seem to have improved the outcomes on exams moderately. Um, it's, uh, again, it's sort of hard to isolate that that's really from like flipping the classroom and not from some of these other changes.

The Economist concluded the session by reviewing the methodologies used in the provided articles and ultimately stated that the evidence was not convincing due to methodological concerns:

> I don't think I've seen, uh, one study here that, uh, is sort of convincing, strong, uh, large scale, randomized evidence that, that, that flip is, uh, is effective for learning. But there's definitely a bunch of sort of suggestive case studies, um, uh, uh, that, you know, some of these pieces might help. Um, the, this last article sort of, uh, suggests that that flip itself is not, um, uh, is not particularly helpful. Um, uh, and, uh, it seems like a, a fairly a, uh, well done study from, uh, looking at it, uh, briefly. Um, uh, but again, sort of more evidence would be useful.

Overall, the Economist's review of the provided materials demonstrated effective transfer of knowledge about research methods developed in economics to effectively assess education

literature. This was mirrored by the Historian who also applied knowledge of research methods developed in a different academic discipline to build complex understandings of the provided education literature. This was in contrast to the assessment of research methodologies demonstrated by the far-transfer group that included positive transfer (i.e., statistical knowledge and career management) but also demonstrated negative transfer when members of the far-transfer group largely ignored findings reported in the literature due exclusively to the author's use of qualitative methodologies.

Negative Transfer: Rejection of Qualitative Methods

Thus far, the provided examples of transfer have been positive (i.e., experts from academic disciplines outside of education effectively utilized their knowledge to understand education research). However, the analysis of participants in the far-transfer group demonstrated the effects of negative transfer as they consistently dismissed findings and sources generated using qualitative methods. This contrasted with the behaviors of members of the no-transfer and near-transfer groups who recognized Wilson (2013) as a predominantly qualitative source and still found value in the presented findings. Specifically, members of the no-transfer and near-transfer group noted that Wilson's article was a single author's report of findings based on her own experience flipping the classroom and applied this knowledge effectively to understand where Wilson's finding may be generalizable and where conclusions based on Wilson's results should be carefully considered.

For example, Educator B characterized Wilson as a case study and expressed excitement about reviewing it after reading other work that was based largely on student evaluations, stating: "There's just so little information in a student evaluation, um, that I think that the case study's

probably gonna do, um, be the most interesting part of it." Educator B later continued this initial assessment and specifically found value in Wilson's case study approach, stating: "And I like to, I like that it's written in the first person. And so, it's like a, it's a kind of direct case study." Then, Educator B connected Wilson's recommendations for employing the flipped classroom to specific strategies recommended by Brame (2013) and found explicit advice to pass on to the hypothetical colleague about how to effectively implement a flipped classroom:

> So, h- here we've got some nice strategies for, um, you know the first piece of advice, um, in the Brame article. Or, second piece of advice, which is to provide incentives. Um, she's got, um, quizzes, um, before each class period.

In contrast to Educator B's view of Wilson as a valuable resource, members of the far-transfer group dismissed it. In this instance, members of the far-transfer group failed to demonstrate aspect 2 of the Apt-AIR framework (i.e., adapting epistemic performance) as they continued to rely on definitions for evidence, data, and findings built in their own fields of expertise while working on a difficult question in a different field (i.e., education). For example, Chemist A quickly reviewed the article looking for numerical data and implied that only numbers could be data:

> And now I'm quickly looking to see if there's actually any data in the f- in the paper, and the answer is it isn't ... it's more- it's very descriptive. I don't see anything that looks like actual data, um ...

This negative assessment continued as Chemist A reviewed the article including questions about why Wilson employed qualitative methodologies and a general dismissal of this article:

> Yeah, it's this very, very qualitative, um ... and-and this also makes me just wonder, like why the heck is it so qualitative? Are the results- Do they have to really parse things? So anyway, I'm not reading this article. I will say that if this were in my field and this was something I really wanted to dive into, I would never make broad conclusions based on just scanning a paper, but-but in fact, I'm a busy person. I mean, you know, life is busy, and I- I don't have time to go into details. Um ... and so, my final shot is that I'm sort of a little skeptical of this article, but I recognize it's for no good reason. It's just sort of general skepticism.

Chemist B's review of the value of Wilson's (2013) case study about flipping an undergraduate statistics course was more succinct. After rapidly reviewing the article, Chemist B described it by stating: "This doesn't look like a traditional article. That's what, it's almost like a translation or script from the presentation."

Likewise, the Life Scientist largely dismissed Wilson (2013). At the conclusion of the session, in response to the prompted question that requested participants assess the quality of the arguments in the provided materials, the Life Scientist stated:

> So, I think, um, the arguments are reasonable. I think it's just, you know, how um, well controlled, so the article, uh, in Life Science Education (i.e., Jensen et al., 2015) is very kind of tightly controlled experiment whereas the article, um, in the Society for Teaching of Psychology (i.e., Wilson, 2013), that's really more of just kind of a description of, um, how, um a professor changed his teaching style.

Overall, each of the three members of the far-transfer group described Wilson as a narrative description of what one instructor had done and found little value in the methodology employed.

Grounded in this distrust of findings from qualitative research, members of the far-transfer group largely dismissed sources based on qualitative methodologies. This hindered their ability to find value in these sources. For example, Chemist A began to discount the value of Wilson's (2013) article while reviewing the abstract, stating: "Okay, so right away, it's clear the level of this ... from the abstract, it seems like the level of sophistication of this article is not very high." Likewise, Chemist A determined that Wilson's findings were not valuable because of the size of the class used as a sample:

> Oh, and then I'm finally learning something really important, um ... *My classes are limited to 25 students*, and now I'm just thinking, oh my God, what is-what is this? This is crazy. This is ... If you're teaching 25 students, it should be interactive and engaging, and if you're lecturing 25 students, you're just wasting your time. Um ... So, then all of a sudden, I'm like, oh my gosh, this almost doesn't apply.

Chemist A then explicitly stated that Wilson's article would not be valuable for the hypothetical colleague to review:

> I would tell this person that I think these publications ... you know, publications are fine, but in this particular case, um, they're not the most relevant thing actually for-for decision.

Chemist B concurred with this assessment of Wilson after reviewing it in response to the second prompted question:

> Second question is that, "*what are your thoughts on the quality of the arguments presented in these articles?"* Hm, the first one, if I remember, was more like a script coming from the person's presentation in a conference, so there were some evidence,

> some data presented, but was not written in a way that, uh, more like a scientific publication.

Here, Chemist B described Wilson's article, which was a scholarly publication, as just a script from a conference presentation specifically because it lacked data of the kind Chemist B sought. Overall, members of the far-transfer group utilized schema for what research and data look like in their own natural science fields to assess the quality of qualitative research methodologies and subsequently the sources built on evidence gathered using those methodologies. This negative transfer hindered their ability to gather information from these sources and to draw the types of complex conclusion exhibited by the no-transfer experts.

Summary of Results

My analysis of the data gathered in this study has answered the proposed research questions. With respect to research question 1 (i.e., How do education experts differ from other experts in the EC they employ to evaluate education research?), several examples emerged from the data. First, education experts demonstrated more complex analysis as they reached conclusions about whether the hypothetical colleague should shift to using flipped classroom pedagogy. Education experts used their knowledge of education research to ask more complex questions than other experts (i.e., members of the near- and far-transfer groups) including the role of academic discipline, the effects of training and facilities, and the importance of instructor attitude and training on the efficacy of flipping a classroom. Similarly, educators utilized their knowledge of the field to create more complex evaluations of the provided sources than members of the near- and far-transfer groups. Specifically, educators applied knowledge about individual

researchers and changes in the field of education research that have occurred in a given time period to draw more complex conclusions about the value of the presented sources.

Several examples of positive transfer were captured that provided answers to research question 2 (i.e., To what degree and in what ways do experts from outside of education transfer their apt epistemic performance to the evaluation of education research?). Multiple examples emerged from the data that demonstrated the transfer of knowledge built in other disciplines by members of the transfer groups, which allowed them to build a complex understanding of an education question. Notably, experts from the near- and far-transfer groups effectively transferred statistical knowledge to make sense of the provided education literature and construct advice for the hypothetical colleague. Similarly, participants from across the two transfer groups transferred knowledge about academic careers and effectively applied this knowledge to provide advice to their hypothetical colleague. Additionally, near-transfer group members also transferred apt epistemic performance about research methods to effectively assess the research methodologies employed in the provided materials.

With respect to research question 3 [i.e., If apt epistemic performance transfers, is this transfer generally positive (i.e., beneficial to achieve a complex understanding of the topic) or negative (i.e., a hindrance to achieving a complex understanding)?], evidence of both positive and negative transfer emerged from the data. Evidence for positive transfer took two forms. First, positive transfer was demonstrated by members of both groups (i.e., near- and far-transfer) that did not reach the same levels of depth of analysis as that demonstrated by member of the no-transfer group. These findings included the depth of conclusions reached by members of each group and the complexity of the source evaluation demonstrated. In the second form of positive

transfer, which included the transfer of knowledge about statistics, academic careers, and methodologies, there were not qualitatively differences in the depth of analysis demonstrated by members of the no-transfer group and members of the near- and far-transfer groups. Instead, participants across all three groups reached similar conclusions and demonstrated similar levels of apt epistemic performance. Finally, negative transfer emerged from the data as members of the far-transfer group dismissed findings built using qualitative methodologies which reduced the complexity of their overall conclusions.

Finally, in answer to research question 4, [i.e., What are the differences between apt epistemic performance when it is transferred between similar academic disciplines (i.e., near-transfer; e.g., social sciences to education) when compared to transfer between less similar academic disciplines (i.e., far-transfer; e.g., hard sciences to education)?] the evidence from this study demonstrated that members of the near-transfer group did not engage in negative transfer. However, members of the far-transfer group did exhibit negative transfer as they dismissed a form of research that is utilized less often in their own fields. In contrast, members of the near-transfer group were able to make connection between qualitative methodologies employed in their own disciplines and avoid the effects of negative transfer.

CHAPTER V: Discussion

In the modern world, there is great interest, among both scholars and laypeople, in truth and falsehood. The ability to effectively assess information, particularly in the modern world's complex information environment, requires apt epistemic performance (i.e., "performance that achieves valuable epistemic aims through competence"; Barzilai & Chinn, 2018, p. 353). Questions about how to best promote apt epistemic performance drove Barzilai and Chinn (2018) to describe the Apt-AIR framework. Barzilai and Chinn stated that competence in learning and performance necessarily includes an ability to adapt in order to achieve epistemic aims in a variety of contexts (i.e., aspect 2 of Apt-AIR). This adaptation requires the transfer of knowledge, skills, and abilities acquired in one context to novel contexts. Transfer is essential to the success of formal education as most of what is learned in school will eventually be applied outside of the classroom (Barnett & Ceci, 2002). Researchers have investigated transfer for over 100 years, and the complexities of the modern world have only increased interest in transfer research (Lobato, 2012). My goal for this study was to investigate whether and, if so, how professors from the social sciences (i.e., the near-transfer group) and natural sciences (i.e., the far-transfer group) transferred the apt epistemic performance developed as they gained expertise in their own fields to understand and answer a complex question in education. In order to investigate this, I utilized a methodology similar to one employed by Greene and colleagues (under review). Greene et. al. investigated the transfer of apt epistemic performance as non-psychologist social scientists (e.g., an anthropologist) and natural scientists (e.g., a physicist) worked to reach conclusions about the replication crisis in psychology. They asked a group of

psychologists (i.e., their no-transfer group) the same questions about the replication crisis and used their answers as a baseline for comparison. Similarly, I selected flipped classroom pedagogy as the topic for experts from education (i.e., no-transfer), other social sciences (i.e., near-transfer), and natural sciences (i.e., far-transfer) to investigate. The growing interest in flipped classrooms, from both experts and laypeople, contributed to my selection of it as the best topic for participants to investigate. However, beyond simply the growing interest in flipped classrooms, the use of a question specifically about college teaching, as opposed to different education topics (e.g., measurement or reading comprehension), allowed college professors with expertise in fields other than education to apply the competencies and experiences they have built in college teaching to assess education research. Despite their familiarity with college teaching, this study required the epistemic evaluation of education research beyond that which is required as college professors build the competence necessary to teach. It was not clear whether the apt epistemic performance professors have developed as they assess research in their own fields would transfer. Thus, this study differs from the one conducted by Greene et al. (under review) because it investigates a topic that professors outside of education know about, but may not understand at the same depth as education professors. My intention was to contribute to the understanding of both apt epistemic performance and transfer while simultaneously supporting an argument for the use of the Apt-AIR framework in empirical investigations.

Barzilai and Chinn's (2018) Apt-AIR framework described five aspects of apt epistemic performance as crucial to achieving epistemic aims through competence. Each of these aspects requires that learners set epistemic aims and values, establish the epistemic ideals necessary to determine whether those aims have been met, and engage in reliable processes to achieve the

aims. Barzilai and Chinn developed the Apt-AIR framework in order to establish goals for epistemic education. The 2018 article that introduced the Apt-AIR framework (i.e., Barzilai & Chinn) has been cited as part of studies focused on a variety of topics including the role of argumentation in learning (Iordanou, Kuhn, Mato, Shi, & Hemberger, 2019) and the effects learner's determinations of source comprehensibility (i.e., easy to understand or difficult to understand) on decisions about the reliability of those sources (Scharrer, Stadtler, & Bromme, 2019). However, to date, little empirical investigation of the Apt-AIR framework has been conducted. Several of the selected themes in the current study connected directly with the findings of the only other known project that specifically addressed the transfer of apt epistemic performance among experts using the Apt-Air framework (i.e., Greene et al., under review). After analyzing their data, Greene and colleagues organized their findings into three themes: 1) Understanding of social practices in psychology; 2) Knowledge of psychology research methods, statistics, and theory; and 3) Depth of justifications for conclusions. There are several connections between the themes highlighted by Greene and colleagues and those I found. For example, Greene and colleagues highlighted the role an understanding of social practices in psychology played in aiding participants to reach a complex understanding of the replication crisis. As expected, the psychologists in their study demonstrated the most complex understanding of social practices in psychology. The other social scientists (i.e., near-transfer group members) demonstrated a more complex understanding of social practices in psychology than the natural scientists (i.e., far-transfer group members) and, thus, reached more complex conclusions about the replication crisis in psychology. Similarly, as captured in the first theme of my study, an understanding of education research allowed the members of the no-transfer group

(i.e., education experts) to reach more complex conclusions (e.g., the role faculty training and school resources may play in the success of flipped classrooms) about whether the hypothetical colleague should switch to a flipped classroom pedagogy than those reached by members of the transfer groups. The findings from this study begin to establish the utility of conducting empirical investigations into the transfer of apt epistemic performance using the five aspects of the Apt-AIR framework as a starting point. Specifically, similar to the findings from Greene et al., my findings indicate there is value in parsing knowledge transfer along the five aspects of apt epistemic performance described by Barzilai and Chinn and noting differences and similarities between the ways that in-domain (i.e., no-transfer) experts engage with a complex problem when compared with experts from outside the domain.

At the beginning of the project, I proposed three lenses of transfer (i.e., traditional cognitive theory, actor-oriented theory, and preparation for future learning) as valuable ways to envision the transfer of experts' apt epistemic performance. After analyzing the data, I determined that the preparation for future learning (Bransford & Schwartz, 2012) perspective was not as valuable as originally conceptualized. Predominantly, this was attributable to methodological decisions I made about this study. Using the preparation for future learning lens requires that learners revisit material after being exposed to an opportunity to learn something novel. However, in this study near- and far-transfer group members were exposed to the educational content just once. Thus, it was not possible to capture a change in their performance over time that could have demonstrated transfer through a preparation for future learning lens.

In contrast, the traditional cognitive theory of transfer (Barnett & Ceci, 2002) did provide value for this study. Notably, the distinctions between positive and negative transfer were

defined by researchers working with the traditional cognitive theory and those distinctions were crucial to defining the major themes of this project. The traditional cognitive theory of transfer has previously been utilized primarily to study transfer in a very specific way. That is, students are taught new knowledge in one learning environment then asked to utilize that knowledge in a novel situation (Lobato, 2012). Researchers employing this methodology have typically had preconceived notions of the correct answers to the questions posed in the new learning environment and measured transfer simply by assessing whether learners achieved these answers in the new environment. In this study, instead of using that research paradigm, I asked participants from the near- and far-transfer groups to apply knowledge obtained throughout their scholarly careers to answer novel questions to an education topic. Another distinction between the way that I used the traditional cognitive theory of transfer and the ways it has typically been used in the past is that there was neither direct instruction of new learning nor a preselected correct solution. Instead, members of the no-transfer (i.e., in-domain, education experts) were asked to investigate the same problem as the members of the two transfer groups, and then I compared the epistemic processes exhibited by members of each group, as captured using think-aloud protocol.

Overall, I determined the actor-oriented perspective (Lobato, 2012) was the most valuable of the three transfer lenses. Similar to the ways that other researchers have used the actor-oriented perspective, I employed a qualitative methodology and analyzed the data to look for transfer without relying on an *a priori* list of coded correct answers or acceptable types of transfer. In the actor-oriented perspective, transfer is not simply the application of a discrete piece of knowledge or a skill built in one location to a novel problem as it was defined in the

traditional cognitive view of transfer. From the actor-oriented perspective, transfer more broadly encompasses the generalization of knowledge and skills and the use of that knowledge to perform a novel task. This definition of transfer was the most valuable for measuring experts' use of their acquired knowledge and experiences to answer complex questions built from research in a different academic domain. Despite the value of the actor-oriented perspective, there were still challenges with adequately capturing the transfer of apt epistemic performance among experts. These challenges imply multiple future directions and additional research about transfer theories and their applicability to the transfer of apt epistemic performance across varying levels of expertise.

I found that experts from both the near- and far-transfer groups did transfer their apt epistemic performance and, in the process, demonstrated aspect 2 of the Apt-Air framework (i.e., adapting epistemic performance). This transfer also included examples of cognitive engagement in epistemic performance (i.e., aspect 1), and, in certain limited instances, aspect 5 (i.e., participating in epistemic performance together with others). I organized these findings into three themes and connected these themes to four research questions.

Before gathering data, I proposed three potential outcomes. First, if apt epistemic performance transfers, then I predicted that members of the near- and far-transfer groups would engage deeply with the content and implement reliable epistemic processes to achieve their established aims. Second, if apt epistemic performance only near transfers then I predicted distinctly different results among the two transfer groups. Finally, I predicted that if apt epistemic performance did not transfer there would be clear differences in the performance of education experts when compared with members of both transfer groups. The data indicated that

apt epistemic performance did transfer in certain instances, supporting my first predicted outcome, and exposed some interesting distinctions between the performance of the three groups. First, there was evidence of the positive transfer of apt epistemic performance as members of both the near- and far-transfer groups transferred statistical knowledge and knowledge about academic careers to advise the hypothetical colleague. There was also evidence to support the argument that some aspects of apt epistemic performance only near transfer as members of the near-transfer group applied their knowledge of qualitative research methods to assess education literature in ways that aligned with the assessments performed by the in-domain education experts. This contrasted with members of the far-transfer group (i.e., natural scientists) who demonstrated negative transfer by steadfastly applying epistemic ideals about the research methodologies utilized in the natural sciences to assess the value and complexity of qualitative research in education. Finally, there were distinct differences between the no-transfer group and members of both transfer groups in the complexity of conclusions and source evaluations, supporting the value of in-domain expertise and demonstrating that there are limits to the transfer of apt epistemic performance.

Limitations

This study was limited by several important factors. First, due to the intensive and exploratory nature of the research methods, a relatively small sample of only nine participants was selected. The small nature of the sample makes it difficult to draw robust conclusions about the findings with respect to the population. Similarly, the expert-expert interview process necessarily precludes making conclusions about the ability of non-experts to transfer knowledge. Another limitation emerged during data analysis as the transfer of statistical knowledge arose as

an important theme. The provided materials required only a rudimentary understanding of statistics and that limited the level of expertise required to excel at the task. This limited the ability of participants to demonstrate apt epistemic performance and possibly prevented more nuanced findings about the transfer of apt epistemic performance related to statistics that may have been revealed if the provided readings contained more complex statistics.

The selected methodology also limited my ability to draw conclusions about all of the five aspects of the Apt-AIR framework. Specifically, Barzilai and Chinn's (2018) aspect 5 (i.e., participating in epistemic performance together with others) describes engaging in epistemic performance with others including developing, justifying, critiquing and applying reliable processes. This aspect also includes an emphasis on achieving epistemic aims together with others and recognizing epistemic injustice. There were limited opportunities for participants in this study to engage in aspect 5 because there were no opportunities for participants to engage with others. The limited demonstrations of aspect 5 that did occur revolved around the application of accepted norms in a given field of research that were established among scholars.

Implications and Future Directions

Several different avenues could be pursued in order to advance the findings of this study. For example, similar future studies may include a larger sample size in order to potentially enhance the generalizability of the findings. Additionally, future researchers may elect to investigate the transfer of apt epistemic performance among non-experts in order to compare these results with the existing evidence about the transfer of apt epistemic performance among experts. Future researchers may also use a different set of academic disciplines both to establish the groups and as the basis of the complex problem. That is, Greene et al. (under review) selected

a psychological problem (i.e., understanding the replication crisis in psychology) and established psychologists as the no-transfer group before asking experts to work on a complex problem. In this study, driven by the connections between professors who have research expertise in one field and the task inherent in their role as professors to teach classes, an education topic was selected. Future researchers may select from a variety of topics to serve as the no-transfer discipline that exist along a continuum of familiarity for the transfer participants providing opportunities for researchers to test multiple hypotheses about whether and how apt epistemic performance transfers. Since both Greene et al. (under review) and I generated findings about the transfer of apt epistemic performance related to statistics, future researchers may set out to specifically investigate the transfer of statistical knowledge and provide source material with complex statistics in order to ensure that participants have the opportunity to transfer statistical expertise.

Similarly, future researchers may focus on the role of experience in the transfer of expertise. In this study, participants were divided by the academic discipline of their expertise, but not by their years of experience as instructors after receiving their PhDs. There may be distinct differences in the ways that experienced experts transfer their apt epistemic performance when compared with novice PhDs. Future researchers may choose to specifically investigate other aspects of apt epistemic performance (e.g., aspect 5) that appeared infrequently in this study but are important aspects of learning and knowledge. These investigations may require different methodologies including, for example, a learning task that involves collaboration with others to allow for the investigation of cooperative aspects of apt epistemic performance. Finally, researchers may build on this research to investigate the effects of training on transfer. This research could take many forms. For example, participants could be explicitly trained in transfer

and asked to transfer their existing knowledge in order to assess the effectiveness of the training. Alternately, researchers may opt to specifically train participants on epistemic cognition and apt epistemic performance and measure if and how effectively participants are able to employ this training to a complex problem in a different area.

Conclusion

I believe this study makes a contribution to the literature for both epistemic cognition and transfer and contributes to an argument that the Apt-AIR framework has utility for use in empirical investigations. I identified ways that experts from a variety of disciplines enacted epistemic performance to answer questions about an education topic. In the process, I found evidence that apt epistemic performance does transfer. Typically, this transfer was positive although in one instance the negative transfer of epistemic performance prevented participants from achieving a complex answer to the question. My findings connect with similar findings from Greene et al. (under review) and provide support for the use of Barzilai and Chinn's aspects of the Apt-AIR framework to conduct empirical research.

APPENDIX A: Questionnaire

Post Event Questionnaire:
Participant #: _____
Area of Expertise: _____
Demographics:

1. How many years have you taught after completing your PhD? ____
2. On average, how many courses do you teach in a semester? _____
3. What is your age? _____

Additional questions:
1. Describe your typical pedagogy (e.g., lecture). _____
_____
_____
2. Describe the processes you generally employ to evaluate the quality of research. _____
_____
_____

3. Describe your knowledge of flipped classroom pedagogies prior to beginning this study.
_____
_____
4. Describe the types of research you have conducted.
_____
_____

APPENDIX B: Procedures

The following step-by-step process will be employed with each participant:

1. Occupy a quiet, private place to gather data for approximately 90 minutes.
2. Introduce myself.
3. Review and sign IRB Consent form
   - Ensure confidentiality
   - Ask for participants' help to maintain confidentiality by not mentioning names or identifiable information during the recording
4. Explain think-aloud protocol by reading the following statement:
   - "We will be using think-aloud protocol to capture today's session. Think-aloud protocol requires that you as the participant continually verbalize your actions throughout the learning process. This means that you should read and think aloud. You should be verbalizing what you are thinking and reading not explaining it to me or narrating your actions. I am here to ensure that the equipment is working and to prompt you if you are quiet for too long, but you should otherwise act like I'm not here. If you are quiet for more than five seconds, I will remind you to please keep talking. Do you have any questions?"
5. Clip lapel microphone to shirt, begin recording
   - "Today is (date), (time) and this is participant number (x)."
6. Scenario:
   - I am required to read the scenario to you. After that, we will place the scenario in a place where you can view it throughout our time together. After I review the scenario, I can answer any questions.
     "One of your colleagues is beginning her third year teaching undergraduates in a department and subject similar to yours at a university considered one of your university's peers. She has autonomy over her course. She has read about flipped classrooms and wants to know your opinion. She sends you these four publications and asks you to review them before answering questions. You may take up to an hour to complete the task or stop at an earlier time of your choosing. Please be certain to answer each of these questions during the task."

     i. What do you think of flipped classroom pedagogy?

     ii. What are your thoughts on the quality of the arguments presented in these articles?

     iii. Do you think she should implement flipped classrooms in her new course? Why or why not? Please justify your answer using reasons and evidence.

7. Do you have any questions?

8. “You may take up to an hour to review the materials and answer the questions. You can stop the recording at any time. Once you finish and answer the questions I will stop recording and provide a brief survey for you to complete. I will let you know as time runs out prompting you when there are five and two minutes remaining. Please remember to think-aloud, and please try to answer your colleague’s questions before the end of the task.”
9. “Please remember to talk throughout the process. I will prompt you if you are quiet for 5 seconds or more.”
10. Once recording has stopped:
    a) Remove the microphone
    b) Complete Demographic survey
    c) Answer questions

APPENDIX C: Results

Table 1

Positive Transfer: Difference in depth of analysis of conclusions and source evaluation

Sub-theme	All	No-Transfer (Educators)	Near-transfer (Other Social Scientists)	Far -transfer (Natural Scientists)
Conclusions to primary questions	Flipped classroom pedagogy is one way to achieve active learning and student engagement	Added questions about academic discipline, student attitude, instructor motivation, instructor training, and facilities Educator A: if they're gonna compare by class, are they gonna compare by major discipline, because that might have an impact on it? (Aspect 1, Ideals) Educator A: *Professors remain the best source for guiding students in how to understand* (Wilson, 2013)...So again, as I'm reading this, and they're talking about the faculty, it makes me think of the difference in the disciplinary based. How much of a difference does that make? (Aspect 1, Ideals)	Questions about role of class size and relative costs Historian: Okay, this new colleague is teaching a course, uh, similar to mine. Uh, and I would say to her, a flipped classroom is worth considering, um, it's not a cure all. You certainly should not undertake it if for instance you've got five preparations per semester. You will be, um you will never get it all done. … Um and think also about ways that you can encourage active learning without necessarily uh doing a full-fledged flip. I would tell her that encouraging active learning is the most important thing, and flipping is not the only way to get there.	Questions about role of class size and relative costs Life Scientist: As far as the flipped classroom pedagogy, I think it's a technique. Um, I think the key thing is really, um, whether it's an active learning environment, um based on these articles, um, they seem to indicate that the amount of interaction with the students and using active learning techniques like projects within

Sub-theme	All	No-Transfer (Educators)	Near-transfer (Other Social Scientists)	Far -transfer (Natural Scientists)
		Educator A: *Attitudes toward the course did differ* (Jensen et al., 2015). The interesting thing is I don't think they went and asked about the attitudes, which is towards the course and not towards the instructor even though they said that, that was one of the factors is the, the link and the relationship with the instructor, but most questions, they didn't ask anything about that. (Aspect 1, Reliable process)	(Aspect 1, Ideals) Economist: Students presumably should also have practice doing the thing that we end up testing on, which is applying these things to new contexts. Uh, I ... so it's not, it's not immediately clear to me that you can't do that without, uh, in, you know, without doing a flipped classroom, um, with or without a flipped classroom. It seems like you can sort of still get at those ideas. But I am, uh, I'm coming to it from the perspective that I sort of agree with the goal.	classes, or having to explain their, their homework or questions, um, is more important relatively speaking. Although the requirement to, um, have some, uh, of the information, um, um, be learned outside of class is probably helpful for that. (Aspect 1, Ideals)
		Educator B: I mean, I think I'd have to talk to her, um, to answer the third question. I think the k- key question is like, is somebody telling her t- to flip the course? Um, or is it something that she wants to do? Like, what is the nature of her curiosity? (Aspect 1, Aim)	(Aspect 1, Ideal)	Chemist A: *My classes are limited to 25 students* (Wilson, 2013), and now I'm just thinking, oh my God, what is-what is this? This is crazy. This is ... If you're teaching 25 students, it should be interactive and engaging, and if you're lecturing 25
		Educator B: okay, so now we're talking to this, gets us to this example in, uh, intro calc. Um...		

Sub-theme	All	No-Transfer (Educators)	Near-transfer (Other Social Scientists)	Far -transfer (Natural Scientists)
		So, there's *60 s- small sections of intro calc, …32 students per class*…. They meet *three days a week*. Um, *faculty* (Berrett, 2012) are well trained, that seems important to sort of point out. (Aspect 1, Ideal) Educator C: So, on top of this, we would need to be thinking about a different method of assessment, which is what I went back to earlier in terms of thinking about that immediate feedback. In particular, thinking about what training are teachers getting for this, particularly if their own education was didactic, multiple choice, that kind of thing. And, and what does it mean to now go in, uh, and how do you test a concept inventory. Does that go back to this idea of multiple choice? Um, even if it does get at concepts, but how are you getting students to sort of explain their strategies? (Aspect 1, Ideals)		students, you're just wasting your time. Um ... So, then all of a sudden, I'm like, oh my gosh, this almost doesn't apply. (Aspect 1, Ideal)

Sub-theme	All	No-Transfer (Educators)	Near-transfer (Other Social Scientists)	Far -transfer (Natural Scientists)
		Educator C: What's the infrastructure for it? Because, um, one of the big things I take away from reading these, is that all of those things are necessary to make flipped classrooms the benefit that they claim to be. You have to have classroom infrastructure to do the unconventional. You have to have infrastructure and support for the technology. Um, you have to have some additional training on assessment techniques that you may not be familiar with. (Aspect 1, Reliable process)		

Sub-theme	All	No-Transfer (Educators)	Near-transfer (Other Social Scientists)	Far -transfer (Natural Scientists)
Source evaluation		Deeper Source Evaluation Educator B: Um, so I see- I see two- I see two, um, peer-reviewed articles, I see a Chronicle article, um, by Dan Barrett, and I'm gonna just- curious about who the author is, just if I can get a sense. But I don't see that there, um, so Chronicle article by someone who I don't know. Um, if I was sitting at my desk I'd probably, I might start by googling him just to know whether it's a- whether I'm hearing about, um, you know, hearing an instructor's perspective or research perspective kind of before I decide how much energy to spend on that article. Um, and then I h- have this, um, l- looks to me like a- a nice and concise,	Historian: Wait a minute, I'd better do something historical here and check the dates on these things. All right. So, this one's 2014. This one's 2012. This one, I can't find a date right away, but all of the dates in the bibliography are before 2012. So, all right. So, this long journal article (i.e., Jensen et al., 2015) is the most recent piece of material from 2015. All right, that's also I guess, that's what I would expect. Uh, as time goes on the research gets more sophisticated. That's ... that's good I guess. And then finally I'm reading one from 2013. (Aspect 1 & 2, Reliable process)	Life Scientist: I'm just going to look at these articles, the ones that I haven't read. And just take a look and see where they're from and what the titles are. Okay, so Flipping the Classroom by Cynthia J. Brame, CFT instructor. So, this looks like it's just something taken off the internet I'm guessing, it doesn't look like an actual journal article as far as I can tell. Another one is from Life Science Education. It looks like a journal. Article title is Improvements from a

Sub-theme	All	No-Transfer (Educators)	Near-transfer (Other Social Scientists)	Far -transfer (Natural Scientists)
		um, lit review (i.e., Brame, 2013). Um, I'm just curious to see whether... It looks like it's maybe a little bit dated, um, and the, um, and the- the- the, uh, peer-reviewed article seemed to be more recent. (Aspect 1, Reliable process) Educator C: Um, this one is Jensen and it is spring 2015. And the other is Wilson and it's 2013. I see that there is still no date on this other one that I don't know what source it's from. All right. I am gonna read the most recent one. (Aspect 1, Reliable process) Educator C: Generally, I would want to check the source to make sure that this is from either a publication or a blog that I trust. Um, I see that it says CFT. I don't know too much about that. So, I would want to know more about the source, where it came from. (Aspect 1, Reliable process)	***Political Scientist:*** As far as the quality of the arguments, I took the academic articles to be providing the quality of the argument. And that the Chronicle and the other article are simply summations of the literature and sort of a pop science lit review. (Aspect 1 & 2, Ideal) ***Political Scientist:*** Actually, the last one appeared in *Teaching of Psychology*. I just took it as an academic work because of the standard SAGE type setting. This is in *Life Sciences Education*, uh, article. (Aspect 1 & 2, Ideal) Economist: This was sort of, uh, nice news type of article about, uh, that I ... you know, gave some background about what a flipping means and, um, uh, you know, it references, some studies, but without more detail on the studies it's sort of hard to evaluate them	Flipped Classroom May Simply be the Fruits of Active Learning. Um, yeah that makes sense, just by the title. Um, and then another journal article from the Society for Teaching of Psychology. Um, The Flipped Class: A Method to Address the Challenges of an Undergraduates Statistics Course. (Aspect 1 & 2, Reliable process) Life Scientist: Probably doesn't matter that much based on, you know, the, these, the two principal articles that actually kind of describe specific experiments. The other ones, um, you know, the one in the Chronicle for Higher

Sub-theme	All	No-Transfer (Educators)	Near-transfer (Other Social Scientists)	Far -transfer (Natural Scientists)
			(Aspect 1 & 2, Reliable process)	Education is mostly just sort of anecdotes. (Aspect 1 & 2, Reliable process) Life Scientist: All right, well, the other thing I notice right away is two of these articles are like, real, serious scientific articles, and two of them are like, high-level, um, summaries. (Aspect 1 & 2, Reliable process) Chemist A: so now I'm reading the Chronicle one, which is the other, um ... which is the other sort of ... not research article, but article that's sort of a high level, and then I guess I will say the first thing that- the first paper that was given was- it was- it

Sub-theme	All	No-Transfer (Educators)	Near-transfer (Other Social Scientists)	Far -transfer (Natural Scientists)
				looked like something like, printed out off of a blog page or something. And you know, I look at The Chronicle one, and I- and I-I'm not going to accept it at face value, but I do know it's a very thoughtful, well-researched thing. So, I guess my ... I'm sort of intrigued to see what this says and thinking that it might be a slightly more reputable source than just someone's blog. (Aspect 1 & 2, Reliable process) Chemist A: I'm reading through this, and so I'm kind of thinking like, well, I'm not- this is not a good use of my time, I'm going to stop,

Sub-theme	All	No-Transfer (Educators)	Near-transfer (Other Social Scientists)	Far -transfer (Natural Scientists)
				which I ... if I were doing this in real life, that's what I would do. All right, so now I'm looking at the two research articles and um ... and actually, the first title, um ... uh, already intrigues me because it's sort of ... I am really interested in why it works and for whom it works, and I'm always interested in making classes more engaging, and I think there are a lot of ways of making a class engaging. (Aspect 1 & 2, Reliable process)

Table 2

Positive transfer: Common and group-specific evidence of positive transfer

Sub-theme	No-Transfer (Educators)	Near-transfer (Other Social Scientists)	Far -transfer (Natural Scientists)
Near- and far-transfer groups: Statistical knowledge	Successfully interpreted statistics	Political Scientist: *He found that students taught with interactive engagement methods exhibited learning gains almost two standard deviations higher than those observed in the traditional courses.* And so, he gives the margin of error around the point estimates for each of the courses. (Aspect 1 & 2, Reliable process) Political Scientist: So, I'm going to stop reading to look at the table right now. And this is looking at the examples, uh, from the nation and the world. They got the number of homes destroyed by hurricanes as ratio. They're calling it discrete. The ... Uh, not sure about that. Uh, metro and state, they have forecasted high temperature, they're calling it interval and continuous. That one I agree with. I have sports, the number of goals scored on a soccer game as a ratio. They're calling that discrete. That's to the letter correct. Um, I guess they prob- ... So, it's probably less of an applied view. Um, okay. So, home and garden, hottest zip	Life Scientist: Figure two, unit exam scores. None of the differences is significant. Error bars represent 95% confidence interval. So, uh, exam one, exam two, and exam three. The grades trend downward, but are nearly equivalent for flipped and nonflipped classes. Um, for the final exam scores, um, another bar graph, uh, indicating percentage correct. Total scores, as well as low level score and high-level scores are, um, very close to equivalent for both the nonflipped and flipped. (Aspect 1 & 2, Reliable process) Chemist A: *They gather data from 14 introductory physics courses, taught by traditional methods.* Um ... and I'm skimming here trying to find out what the actual data are ... yeah, and then this is the thing that I always get, that these things are always reported, um, like it's *... exhibit higher learning gains* and then they report numbers like *.48 plus or minus 1.4 and .23 plus or minus .04,* so it's like just barely significant. (Aspect 1 & 2, Reliable process)

Sub-theme	No-Transfer (Educators)	Near-transfer (Other Social Scientists)	Far -transfer (Natural Scientists)
		codes, nominal. Good. To your health, they reported musculoskeletal symptoms and call it nominal. (Aspect 1 & 2, Reliable process)	
Near- and far-transfer groups: Career management	Educator B: And she may actually be thinking about whether she's- she's gettin' ready for tenure. Um, I could imagine the conver- the conversation going into this conversation thinking, like, oh, I didn't think enough about the attitudinal data because I don't think it, um, reflects what I care about. Um, but if I'm on a tenure track, I actually do care about it because the evaluations matter. Um, so I may actually, you know, also have her think about, um, if she's really f- if that's- if the context here is that she's feeling bad about her instruction or her use, it might actually also send her, um, to research about bias in student evaluations, and let her think about	Historian: I would say to her, a flipped classroom is worth considering, um, it's not a cure all. You certainly should not undertake it if for instance you've got five preparations per semester. You will be, um you will never get it all done. Um and also turn your dissertation into a book, which a third-year uh college professor has to do or lose her job. Um, and especially at a university considered one of your university's peers. Um, think carefully about whether, not only whether you have the time to ... to create a flipped class. (Aspect 1 & 2, Aims)	Chemist A: I would say to her, a flipped classroom is worth considering, um, it's not a cure all. You certainly should not undertake it if for instance you've got five preparations per semester. You will be, um you will never get it all done. Um and also turn your dissertation into a book, which a third-year uh college professor has to do or lose her job. Um, and especially at a university considered one of your university's peers. Um, think carefully about whether, not only whether you have the time to ... to create a flipped class. (Aspect 1 & 2, Aims & Reliable process)

Sub-theme	No-Transfer (Educators)	Near-transfer (Other Social Scientists)	Far -transfer (Natural Scientists)
	whether or not the students are evaluating her fairly. (Aspect 1, Aims) Educator C: But nonetheless, in her third year of teaching, if she is a tenured track professor, um, she's going to be going into third year review. Um, so unfortunately you need to be thinking about what's going on for you in terms of preps, in terms of what you're getting out. (Aspect 1, Aims)		
Near transfer group: Methodological assessment	Educator B: I can see here when I get to the tables, that there is a- there's some kind of control, um, but you know the data are student reports of looks to me like, yea, student evaluations of course and instructor… I don't, I just don't care about those data. Like, I don't believe them. (Aspect 1, Ideals)	Historian: Uh, the um Jensen, Kummer and Godoy article is um you know, balanced uh on the one hand. On the other hand, I thought it was a well-controlled study, uh if I'm remembering it right. They had a controlled class, yeah. It was another active learning type classroom that wasn't flipped, and I thought that was a useful way to um ... to isolate flipping as a, as an effective ... uh as a meaningful variable. (Aspect 1, Ideals)	N/A

Sub-theme	No-Transfer (Educators)	Near-transfer (Other Social Scientists)	Far -transfer (Natural Scientists)
	Educator B: I kind of think the empirical stuff is not particularly convincing here… Still not really very clear on what the reference category is, but I don't really care that much 'cause I don't believe this (i.e., evidence built from student evaluations) empirically, in- in any sort of really deep way. Um, so they seem to have done better, but compared to what? (Aspect 1, Reliable processes) So now I'm turning to this biology article. Um... Uh, so this is a *quasi-experimental design* (Jensen et al., 2015). Um, we're comparing, and what that means here is we're comparing a... Oh, this is nice. Um, now it gets really- gets at what's... Flipping is the thing that seems to really vary. Um,	Economist: This was sort of, uh, nice news type of article about, uh, that I ... you know, gave some background about what a flipping means and, um, uh, you know, it references, some studies, but without more detail on the studies it's sort of hard to evaluate them. Um, uh, the one that it does cite and talks about, uh, is not a randomized study. So, it was sort of hard to put much weight on this. It's mostly anecdotal evidence. Economist: So I mean here, there is, uh, uh, a lot of results, but, uh, um, not a lot of discussion of the methods that does not seem like, um, these were sort of, again, it doesn't seem like these were randomly assigned a, um, evaluations and, you know, um, my primary concern with evaluating these would be that you, uh, um, you know, if you are, um, if you're just letting the, the instructors choose which method to do, maybe the teachers who are sort of better and more motivated, will try something new. And so, you're actually getting better gains from people who tried something, uh, um ... who, who were sort of more motivated teachers. And it's not about the method itself. And so, you know, really randomly assigning teachers to trying these methods	

Sub-theme	No-Transfer (Educators)	Near-transfer (Other Social Scientists)	Far -transfer (Natural Scientists)
	like, there's one where they're using the sort of active learning, conceptual learning in the classroom but isn't flipped. Um, and one where it's- where it's flipped. Um, and you know, yeah, so this is so not surprising to me. Um... The, when you do this kind of, is it flipped or is it not flipped, um, there's really no difference in outcomes (Aspect 1, Ideals)	or not, and following through would be kind of a better way to do this. It's really unclear from reading this, whether that's what happened or not. Um, it does not seem like it. Um, so it's sort of hard for me to, uh, uh, trust the, the evidence here. (Aspect 1, Ideals) Economist: So, this is fine, but it's still just one observation. Effectively, it's one section or not. And uh, um, uh, it's uh, uh, you know, even if you flipped a coin to pick which, uh, which section got this, it would be nice to have a lot more data, a lot more classrooms and not just one. (Aspect 1, Ideals) Economist: So, at this point I would like to know ... it sounds like this teacher sort of really changed a lot of things at once. It would be nice to know that at least, you know, they kept something like the final exam, the same to have some measure to see whether at least comparing students before and after that they got better on some sort of fixed, uh, fixed task. It's great that you want to reevaluate your learning objectives but maybe reevaluating them, but then keeping the other things the same. So, you have some baseline before you change. Sort of just the teaching methodology to a flipped	

Sub-theme	No-Transfer (Educators)	Near-transfer (Other Social Scientists)	Far -transfer (Natural Scientists)
		classroom would be kind of a cleaner test, of, of the flipped classroom at the moment. It just looks like a test of whether, you know, putting a lot of effort into your course and changing a bunch of things seems to improve, uh, outcomes. (Aspect 1, Ideals) Economist: Here we have an article that has a, you know, one experiment that seems to, uh, by comparing sort of students, uh, in prior years to students in new years where there's been a number of changes to the course, including this sort of flipped classroom that these changes seem to have improved the outcomes on exams moderately. Um, it's, uh, again, it's sort of hard to isolate that that's really from like flipping the classroom and not from some of these other changes. (Aspect 1, Ideals)	

Table 3

Negative transfer: Rejection of qualitative methods

Far -transfer (Natural Scientists) All demonstrate Aspect 1, Ideals
Chemist A: And now I'm quickly looking to see if there's actually any data in the f- in the paper, and the answer is it isn't ... it's more- it's very descriptive. I don't see anything that looks like actual data, um ...
Chemist A: Yeah, it's this very, very qualitative, um ... and-and this also makes me just wonder, like why the heck is it so qualitative? Are the results- Do they have to really parse things? So anyway, I'm not reading this article. I will say that if this were in my field and this was something I really wanted to dive into, I would never make broad conclusions based on just scanning a paper, but-but in fact, I'm a busy person. I mean, you know, life is busy, and I- I don't have time to go into details. Um ... and so, my final shot is that I'm sort of a little skeptical of this article, but I recognize it's for no good reason. It's just sort of general skepticism.
Chemist B: This doesn't look like a traditional article. That's what, it's almost like a translation or script from the presentation.
Life Scientist: So, I think, um, the arguments are reasonable. I think it's just, you know, how um, well controlled, so the article, uh, in Life Science Education (i.e., Jensen et al., 2015) is very kind of tightly controlled experiment whereas the article, um, in the Society for Teaching of Psychology (i.e., Wilson, 2013), that's really more of just kind of a description of, um, how, um a professor changed his teaching style.

REFERENCES

Abd-El-Khalick, F. (2013). Teaching with and about nature of science and science teacher knowledge domains. *Science & Education, 22*(9), 2087-2107.

AERA (2018). 2019 Annual meeting theme. Retrieved from: http://www.aera.net/Events-Meetings/Annual-Meeting/2019-Annual-Meeting-Theme

Alexander, P. (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. 100-110). New York: Routledge.

Baepler, P., Walker, J. D., & Driessen, M. (2014). It's not about seat time: blending, flipping, and efficiency in active learning classrooms, *Computers & Education, 78*, 227-236.

Barnett, S. M., & Ceci, S. J. (2002). When and where to we apply what we learn? A taxonomy for far transfer. *Psychological Bulletin, (128) 4, 612-637.*

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.

Barzilai, S., & Zohar, A. (2014). Reconsidering personal epistemology as metacognition: A multifaceted approach the analysis of epistemic thinking. *Educational Psychologist, 49*(1), 13-35.

Barzilai, S., & Zohar, A. (2016). Epistemic (metacognition): Ways of thinking about knowledge and knowing: in J.A. Greene, W.A. Sandoval, & I. Bråten (Eds.), *Handbook of Epistemic Cognition* (pp. 409-424). New York, NY: Routledge.

Baxter Magolda, M. B. (1987). The affective dimension of learning: Faculty-student relationships that enhance intellectual development. *College Student Journal, 21,* 46-58.

Belenky, M. F., Clinchy, B. M., Goldberger, N. R., & Tarule, J. M. (1986). *Women's ways of knowing: The development of self, voice, and mind.* New York, NY: Basic Books.

Berrett, D. (2012). How 'flipping' the classroom can improve the traditional lecture. *The Chronicle of Higher Education, 12*, (19), 1-3.

Braasch, J. L. G., Bråten, I., & McCrudden, M. T. (2018). *Handbook of Multiple Source Use.* New York, NY: Routledge.

Braasch, J. L. G., Rouet, J. F., Vibert, N., & Britt, M. A. (2012). Readers' use of source information in text comprehension, *Memory & Cognition, 40,* (3), 450-465.

Brame, C. J. (2013). *Flipping the classroom.* Retrieved from https://cft.vanderbilt.edu/guides-sub-pages/flipping-the-classroom/

Bransford, J. D., & Schwartz, D. L. (1999). Chapter 3: Rethinking transfer: A simple proposal with multiple implications. *Review of Research in Education*, *24*(1), 61-100.

Bråten, I., Britt, M. A., Strømsø, H. I. & Rouet, J-F. (2011). The role of epistemic beliefs in the comprehension of multiple expository texts: Toward an integrated model. *Educational Psychologist, 46,* 48-70.

Bråten, I., Ferguson, L.E., Strømsø, H. I., & Anmarkrud, O. (2013). Justification beliefs and multiple-documents comprehension. *European Journal of Psychology of Education, 28* (3), 879-902.

Bråten, I. (2016). Epistemic cognition interventions: Issues, challenges, and directions. In J. A. Greene, W. A. Sandoval, & I. Bråten (Eds.), *Handbook of Epistemic Cognition* (pp. 360-371). New York, NY: Routledge.

Carraher, D., & Schliemann, A. D. (2002). The transfer dilemma. *The Journal of the Learning Sciences, 11*(1), 1-24.

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.

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). New York, NY: Routledge.

Chinn, C. A., Rinehart, R. W., & Buckland, L. A. (2014). Epistemic cognition and evaluation information: Applying the AIR model of epistemic cognition. In D. Rapp & J. Braasch (Eds.), *Processing inaccurate information.* Cambridge, MA: MIT press.

Day, S. B. & Goldstone, R. L. (2012). The import of knowledge export: Connecting findings and theories of transfer of learning. *Educational Psychologist, 47*(3), 153-176.

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,* 281-312.

Deslauriers, L., Schelew, E., & Wieman, C. (2011). Improved learning in a large-enrollment physics class. *Science, 332*(6031), 862-864.

Donald, J. G. (1990). University professors' views of knowledge and validation processes. *Journal of Educational Psychology, 82*(2), 242-249.

Dweck, C. S. (2008). *Mindset: The new psychology of success.* Random House Digital, Inc.

Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review, 95*(2), 256-273.

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 implication for practice* (pp. 409-434). New York, NY: Cambridge University Press.

Elby, A., Macrander, C., & Hammer, D. (2016). Epistemic cognition in science. In J. A. Greene, W. A. Sandoval, & I. Bråten (Eds.), *Handbook of Epistemic Cognition* (pp. 113-127). New York, NY: Routledge.

Ericsson, K. A. (2006). Protocol analysis and expert thought: Concurrent verbalizations of thinking during experts' performance on representative tasks. In K.A. Ericsson, N. Charness, R. R. Hoffman, & P. J. Feltovich (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 223-242). Cambridge, MA: Cambridge University Press.

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data (revised edition).* Cambridge, MA: The MIT Press.

Fitzpatrick, M. (2012, June 24). Classroom lectures go digital. *The New York Times,* p. 24.

Flavell, J. H. (1979). Metacognition and cognitive monitoring. *American Psychologist, 34*, 906-911.

Forsey, M. Low, M. & Glance, D. (2013). Flipping the sociology classroom towards a practice of online pedagogy. *Journal of Sociology, 49*(4), 471-585.

Fox, E., Ericsson, K.A. & Best, R. (2011). Do procedures for verbal reporting of thinking have to be reactive? A meta-analysis and recommendations for best reporting methods. *Psychological Bulletin*, 137(2), 316-344.

Gallup/Knight Foundation. (2018). *American Views: Trust, Media and Democracy*. Retrieved from https://kf-site-production.s3.amazonaws.com/publications/pdfs/000/000/242/original/KnightFoundation_AmericansViews_Client_Report_010917_Final_Updated.pdf

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science, 7,* 155-170.

Glaser, B. G. & Strauss, A. L. (1967). *The Discovery of Grounded Theory: Strategies for Qualitative Research.* Chicago: Aldine.

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.

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.

Greene, J. A., Chinn, C., & Deekens, V. M. (under review).

Greene, J. A., Muis, K. R., & Pieschl, S. (2010) The role of epistemic beliefs in students' self-regulated learning with computer-based learning environments. *Educational Psychologist, 45*(4), 245-257.

Greene, J. A., Sandoval, W.A., & Bråten, I. (2016). Introduction to epistemic cognition. In J.A. Greene, W.A. Sandoval, & I. Bråten (Eds.), *Handbook of Epistemic Cognition* (pp. 1-15). New York: Routledge.

Greene, J. A., Yu, S., & Copeland, D. Z. (2014). Measuring critical components of digital literacy and their relationships with learning. *Computers and Education, 76*, 55-69.

Goldstone, R. L., & Day, S. B. (Eds.) (2012a). New conceptualizations of transfer of Learning [Special issue]. *Educational Psychologist, 47*(3).

Goldstone, R. L., & Day, S. B. (2012b). Introduction to "New conceptualizations of transfer of learning." *Educational Psychologist, 47*(3), 149-152.

Hao, Y. (2016). Exploring undergraduates' perspectives and flipped learning readiness in their flipped classrooms. *Computers in Human Behavior, 59,* 82-92.

Hammer, D. & Elby, A. (2002). On the form of a personal epistemology. In B. K. Hofer & P. R. Pintrich (Eds.), *Personal Epistemology: The Psychology of Beliefs about Knowledge and Knowing* (pp. 169-190). Mahwah, NJ: Erlbaum.

Hammer, D. & Elby, A. (2003). Tapping epistemological resources for learning physics. *The Journal of the Learning Sciences, 12*, 53-90.

Hammer, D. M., Elby, A., Scherr, R. E., and Redish, E. F. (2005). Resources, framing and transfer. In J. Mestre (Ed.), *Transfer of learning from a modern multidisciplinary perspective,* (89-120). Greenwich, CT: Information Age Publishing.

Herrenkohl, L. R. & Cornelius, L. (2013). Investigating elementary students' scientific and historical argumentation. *The Journal of the Learning Sciences, 22,* 413-461.

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

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

Hofer, B. K. & P. R. Pintrich (Eds.) (2002). *Personal epistemology: The psychology of beliefs about knowledge and knowing.* Mahwah, NJ: Erlbaum.

Iordanou, K., Kuhn, D., Matos, F., Shi, Y. & Hemberger, L. (2019). Learning by arguing. *Learning and Instruction, 63,* 101207.

Jensen, J. L., Kummer, T. A., & Godoy, P. D. D. M. (2015). Improvements from a flipped classroom may simply be the fruits of active learning. *CBE-Life Sciences Education, 14*(1), 1-12.

Kahneman, D. (2011). *Thinking, fast and slow.* London: Penguin Books.

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

King, P.M. & Kitchener, K.S. (2004). Reflective judgment: Theory and research on the development of epistemic assumption through adulthood. *Educational Psychologist, 39*, (1), 5-18.

Kitchener, K. S. (1983). Cognition, metacognition, and epistemic cognition. *Human Development, 26*, 222-232.

Kitchener, R. F. (2002). Folk epistemology: An introduction. *New Ideas in Psychology, 20*, 89-105.

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

Kuhn, D. Cheney, R. & Weinstock, M. (2000). The development of epistemological understanding. *Cognitive Development, 15,* 309-328.

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). Mahwah, NJ: Lawrence Erlbaum Associates.

Lave, J. (1988). *Cognition in practice: mind, mathematic, and culture in everyday life.* New York, NY: Cambridge University Press.

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

Lising, L., & Elby, A. (2005). The impact of epistemology on learning: A case study in introductory physics. *American Journal of Physics, 73,* (4), 372-382.

List, A., & Alexander, P. A. (2017). Analyzing and integrating models of multiple text comprehension. *Educational Psychologists, 53,* (3), 143-147.

Lobato, J. (2006). Alternative perspectives on the transfer of learning: History, issues, and challenges for future research. *The Journal of the Learning Sciences, 15*(4), 431-449.

Lobato, J. (2012). The actor-oriented transfer perspective and its contributions to educational research and practice. *Educational Psychologist, 47*(3), 232-247.

Mason, L. (2016). Psychological perspectives on measuring epistemic cognition. In J.A. Greene, W.A. Sandoval, & I. Bråten (Eds.), *Handbook of Epistemic Cognition* (pp. 375-392). New York: Routledge.

Muis, K. R. (2007) The role of epistemic beliefs in self-regulated learning, *Educational Psychologist, 42*(3), 173-190.

Muis, K. R. (2008). Epistemic profiles and self-regulated learning: Examining relations I the context of mathematics problem solving. *Contemporary Educational Psychology, 33,* 177-208.

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,* 3-54.

National Academies of Sciences, Engineering, and Medicine. (2018). *How People Learn II: Learners, Contexts, and Cultures.* Washington, DC: The National Academies Press. doi: https://doi.org/10.17226/24783.

National Research Council. (2000). *How People Learn: Brain, Mind, Experience, and School: Expanded Edition.* Washington, DC: National Academy Press.

Perry, W. G. Jr. (1968/1999). *Forms of ethical and intellectual development in the college years: A scheme.* New York: Holt, Rinehart, and Winston.

Pew Research Center (2018, 10 September). News use across social media platforms 2018. Retrieved from http://www. journalism.org/2018/09/10/news-use-across-social-media-platforms-2018/

Portnoy, L. B., & Rabinowitz, M. (2014). What's in a domain: Understanding how students approach questioning in history and science. *Educational Research and Evaluation, 20*(2), 122-145.

Puntambekar, S., Stylianou, A., & Goldstein, J. (2007). Comparing classroom enactments of an inquiry curriculum: Lessons learned from two teachers. *Journal of the Learning Sciences 16*, (1), 81-130.

Sandoval, W. (2012). Situating epistemological development. *The future of learning: Proceedings of the 10th international conference of the learning sciences* (Vol.1, pp. 347-354.) Sydney, Australia: International Society of the Learning Sciences.

Scharrer, L., Stadtler, M., & Bromme, R. (2019). Judging scientific information: Does source evaluation prevent the seductive effect of text easiness? *Learning and Instruction, 63*, 101215.

Schoenfeld, A. H. (1999). Looking toward the 21st century: Challenges of educational theory and practice. *Educational Researcher, 28*(7), 4-14).

Schommer, M. (1990). Effects of beliefs about the nature of knowledge on comprehension. *Journal of Educational Psychology, 82,* 398-504.

Schommer, M. (1993a). Comparisons of beliefs about the nature of knowledge and learning among postsecondary students. *Research in Higher Education, 34*(3), 355-370.

Schommer, M. (1993b). Epistemological development and academic performance among secondary students. *Journal of Educational Psychology, 85*(3). 406-411.

Schraw, G. J., & Olafson, L. J. (2008). Assessing teachers' epistemological and ontological worldviews. In M. S. Khine (Ed.), *Knowing, knowledge and beliefs: Epistemological studies across diverse cultures* (pp. 25-44). New York: Springer.

Shreiner, T. L. (2014). Using historical knowledge to reason about contemporary political issues: An expert-novice study. *Cognition & Instruction, 32*(4), 313-352.

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. 100-110). New York: Routledge.

Thompson, P. W. (2011). Quantitative reasoning and mathematical modeling. In L. L. Hatfield, S. Chamberlain & S. Belbase (Eds.), *New perspectives and directions for collaborative research in mathematics education.* WISDOMe Monographs (Vol. 1, pp. 33-57). Laramie: University of Wyoming Press.

Thorndike, E. L. (1924). Mental discipline in high school studies. *Journal of Educational Psychology, 15,* 83-98.

Thorndike, E. L. & Woodworth, R. S. (1901). The influence of improvement in one mental function upon the efficiency of other functions. *Psychological Review, 8,* 247-261.

Wagner, J. F. (2010). A transfer-in-pieces consideration of the perception of structure in the transfer of learning. *The Journal of Learning Sciences, 19* 443-479.

Weber, K., & Mejia-Ramos, J. P. (2011). How and why mathematicians read proofs: An exploratory study. *Educational Studies in Mathematics, 76,* 329-344.

Wilson, S. G. (2013). The flipped class: A method to address the challenges of an undergraduate statistics course. *Teaching of Psychology, 40*(3), 193-199.

Yilmaz, R. (2017). Exploring the role of e-learning readiness on student satisfaction and motivation in flipped classroom. *Computers in Human Behavior, 70*, 251-260.