

2021

# Is What Works Working? Thinking Evaluatively About the What Works Clearinghouse

Andrew A. Nelson
*University of Kentucky*, andrew.nelson.226@uky.edu
Digital Object Identifier: https://doi.org/10.13023/etd.2021.052

STUDENT AGREEMENT:

I represent that my thesis or dissertation and abstract are my original work. Proper attribution has been given to all outside sources. I understand that I am solely responsible for obtaining any needed copyright permissions. I have obtained needed written permission statement(s) from the owner(s) of each third-party copyrighted matter to be included in my work, allowing electronic distribution (if such use is not permitted by the fair use doctrine) which will be submitted to UKnowledge as Additional File.

I hereby grant to The University of Kentucky and its agents the irrevocable, non-exclusive, and royalty-free license to archive and make accessible my work in whole or in part in all forms of media, now or hereafter known. I agree that the document mentioned above may be made available immediately for worldwide access unless an embargo applies.

I retain all other ownership rights to the copyright of my work. I also retain the right to use in future works (such as articles or books) all or part of my work. I understand that I am free to register the copyright to my work.

REVIEW, APPROVAL AND ACCEPTANCE

The document mentioned above has been reviewed and accepted by the student's advisor, on behalf of the advisory committee, and by the Director of Graduate Studies (DGS), on behalf of the program; we verify that this is the final, approved version of the student's thesis including all changes required by the advisory committee. The undersigned agree to abide by the statements above.

Andrew A. Nelson, Student

Dr. Kelly D. Bradley, Major Professor

Dr. Jane M. Jensen, Director of Graduate Studies

# IS WHAT WORKS WORKING?
# THINKING EVALUATIVELY
# ABOUT THE WHAT WORKS CLEARINGHOUSE

---

DISSERTATION

---

A dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy in the
College of Education
at the University of Kentucky


By

Andrew A. Nelson

Lexington, Kentucky

Co- Directors: Dr. Kelly D. Bradley, Professor of Educational Policy Studies and Evaluation

and Dr. Shannon O. Sampson, Assistant Professor of Educational Policy Studies and Evaluation

Lexington, Kentucky


2021

ABSTRACT OF DISSERTATION

IS WHAT WORKS WORKING?
THINKING EVALUATIVELY
ABOUT THE WHAT WORKS CLEARINGHOUSE

Since the mid-twentieth century, the U.S. Department of Education has drafted and enacted policies to bridge the research-practice gap—that is, the gap between "what works" according to educational research and what is actually practiced by teachers and their administrators (e.g., Dirkx, 2006; Joyce & Cartwright, 2019; Tseng, 2012). One of the latest manifestations of this "what works" political legacy is the What Works Clearinghouse (WWC), which took shape as part of the Institute of Education Sciences (IES) in 2002. The WWC's mission is to be a "central and trusted source of scientific evidence for what works in education" (WWC, 2020d, p. 1) while, at the same, helping the IES "…increase [the] use of data and research in education decision-making" (IES, n.d.-a). The purpose of this dissertation is to evaluate the extent to which the WWC has realized its own mission as well as contributed to the IES's larger goal.

Guided by principles of evaluative thinking (Vo & Archibald, 2018) and premises of the Two-Communities theoretical tradition (Caplan, 1979; Farley-Ripple et al., 2018), this project used a theory-based evaluation approach called contribution analysis (Mayne, 2008, 2012b, 2019) to investigate three guiding questions. Those questions inquired into (a) the extent of the WWC's impact among educators, (b) the reasons why its impact may be wanting, and (c) the changes it could make to maximize its impact. To investigate these questions, a six-step procedure was used to both articulate and scrutinize the WWC's theory of change according to available evidence. An array of evidence was considered, including existing publications (e.g., previously published evaluations, literature reviews, and large-scale surveys), analyses of publicly available data (e.g., public data exports, data requested through the Freedom of Information Act, transcripts from congressional hearings), and findings from a preservice teacher survey conducted for this project.

The results of this contribution analysis offered compelling answers to each of the three guiding questions. First, given the WWC's original benchmark for success (e.g., Baldwin et al., 2008), evidence suggested that it is likely failing to fully reach educators and guide their decision-making. This was especially true for teachers. Second, the evidence suggested that the WWC's impact may be wanting because its theory of change depends on several unsupported assumptions. Not only were many of the WWC's causal assumptions refuted by the evidence, but some of its foundational assumptions—such as the belief that systematic research review would be an effective way of bringing educational research to practice—were refuted as well. Finally, because several of its foundational assumptions were refutable, the WWC may only be able to maximize its impact if it fundamentally retools its approach to systematic research review or to educational research more generally. Suggestions for doing so are discussed.

KEYWORDS: What Works Clearinghouse, Evaluative Thinking, Contribution Analysis

Andrew A. Nelson

04/26/2021

# IS WHAT WORKS WORKING?
# THINKING EVALUATIVELY
# ABOUT THE WHAT WORKS CLEARINGHOUSE

By
Andrew A. Nelson

Dr. Kelly D. Bradley
Co-Director of Dissertation

Dr. Shannon O. Sampson
Co-Director of Dissertation

Dr. Jane M. Jensen
Director of Graduate Studies

04/26/21
Date

DEDICATION

To those working on the front lines while I worked at my desk.

# ACKNOWLEDGMENTS

Acknowledging everyone who played a role in this dissertation would double its length. Doing the same for everyone who played a role in my graduate education would quadruple it. I am tremendously grateful—and will continue to be—for the family and friends, the classmates and colleagues, the faculty and staff, and the advisors and counselors that helped me out along the way. Yet, as I put the final touches on this manuscript, I am struck by how indebted I feel to a bunch of people I have never personally met—namely, the essential workers who allowed me the time and safety to work on something far less essential. The COVID-19 pandemic ramped up just as my dissertation writing was. As I wrote, hospital staff were saving lives. Grocery store workers kept shelfs stocked. Delivery drivers brought me food, pantry items, and even toilet paper. Those are the people that I wish to acknowledge as well as those whom I have dedicated this project to. I dedicate it to them not because it will make much difference to their lives, but because it has changed my life so substantially. And because completing it would not have been possible without them.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

# CHAPTER 1: INTRODUCTION

In the summer of 1985, staff at the U.S. Department of Education's Office of Educational Research and Improvement (OERI) received a memo informing them of a new assignment. That assignment, as requested by then-Secretary of Education William Bennett, was to:

> "…take the lead in compiling a set of research findings that could be helpful to the public, to policy makers and to parents as well as to education professionals in their efforts to improve the quality of learning for all children…[to] assemble an array of strong research findings that can benefit educational practice and that are stated so clearly and concisely as to be readily accessible to the general public." (Bennett as cited in Finn & Tomlinson, 1987, p. 24)

By March of 1986, the Office had done exactly that. They had identified "the best available research" on effective educational practices (Reagan, 1986, p. 2), assessed that research for "accuracy, veracity, validity, and importance" (U.S. Department of Education, 1986, p. 5), and translated those research findings into "language that they would be pleased to have the public read" (Finn & Tomlinson, 1987, p. 25). Together these efforts culminated in a brief, 65-page booklet entitled *What Works: Research About Teaching and Learning.*

Despite its brevity—or perhaps because of it—What Works seemed a success. A day after President Reagan remarked on the booklet in a White House address, 14,000 Americans requested a copy (Glass, 1987). Within a year, more than half a million copies had been distributed. Even critics of the booklet conceded that it had become "the most widely read document in the history of educational research" (Glass, 1987, p. 5). Being widely read did not ensure its success, however. As Bennett's original memo makes clear, What Works sought to "improve the quality of learning for all children" (Finn & Tomlinson, 1987, p. 24), presumably by encouraging its readers to enact the research-

based recommendations therein. Given Bennett's goal, did What Works work? Did it succeed at bringing research-based educational practices into the classroom? In his 2008 memoir, Dr. Chester Finn, the OERI Assistant Secretary tasked with overseeing the What Works project, offered an unabashed answer. He writes that "despite all the trees and postage sacrificed to its dissemination, it had little impact on its primary audiences" (Finn, 2008, p. 135). In short, What Works didn't work.

By now, the What Works pamphlet has come and gone. The OERI has been dismantled, and the Institute of Education Sciences (IES) has taken its place. Yet the fundamental intention underlying the "what works" agenda—namely, to improve educational outcomes by bridging the "research-practice gap" (e.g., Dirkx, 2006; Joyce & Cartwright, 2019; Neal, Neal, Mills, Lawlor, & McAlindon, 2018, 2019; Tseng, 2012)—has remained unchanged; if anything, it has intensified. One such example is the What Works Clearinghouse (WWC), which some consider the IES's "primary mechanism" for research dissemination (Baldwin et al., 2008, p. xiii). The WWC enjoys certain dissemination technologies that the What Works booklet didn't (e.g., interactive webpages and social media accounts), but its process is much the same. In an effort to bring research to bear in practice, the WWC (a) identifies all research on a given educational program or practice, (b) assesses that research based on its standards of quality, and (c) disseminates findings from approved research through summaries and reports (WWC, n.d.-g). Like the What Works booklet before it, the WWC hopes to be both read and used. Not only does it wish to be a "central and trusted source of scientific evidence for what works in education" (WWC, 2020d, p. 1), but it also hopes to contribute to the IES's larger goal of "…increas[ing] use of data and research in

education decision-making" (IES, n.d.-a). Given these similarities, both in process and in mission, the same question that was asked of the What Works booklet can be asked of the WWC. So, again, I ask: is What Works working?

This dissertation pursues an answer to that question. Admittedly, I am not the first to do so. Scholars have offered their own takes, with some answering affirmatively (e.g., Hearing Before the Committee on Ways and Means, 2015) and others the opposite (e.g., Biesta, 2007, 2010; Lykins, 2012). In the spirit of "evaluative thinking" (e.g., Vo & Archibald, 2018), however, I pursue my own answer. Guided by Farley-Ripple, May, Karpyn, Tilley, & McDonough's (2018) conceptual framework of the educational research and practice communities, I use a theory-based evaluation approach called contribution analysis (Mayne, 2008, 2012b, 2019) to both articulate and scrutinize the What Works Clearinghouse's theory of change. I consider an array of evidence while doing so, including existing publications (e.g., previously published evaluations, literature reviews, and large-scale surveys), analyses of freely-available data (e.g., public data exports, data requested through the Freedom of Information Act, transcripts from congressional hearings), and findings from a teacher survey conducted specifically for this project. When considered systematically, this evidence brings me closer to an answer while, at the same time, helping me identify areas where the WWC might improve.

## Framing the Problem

Before describing this project any further, I must first introduce the political and theoretical contexts informing it. By considering policy and theory together, I intend to highlight a divergence between how U.S. educational policy has attempted to bridge the research-practice divide and how knowledge utilization theories have suggested doing so.

As will be shown below, the federal government has targeted the research-practice gap with increasing authority and deference to knowledge producers (i.e., educational researchers). In contrast, knowledge utilization theories—particularly in the tradition of Two-Communities theory (e.g., Caplan, 1979)—imply that bridging the research-practice gap requires the consideration and inclusion of knowledge users (i.e., educators). Understanding this divergence is critical because it implicates the What Works Clearinghouse's approach as potentially ineffective.

**The Political Context**

The WWC's approach to identifying, assessing, and disseminating educational research is a remnant of the policies preceding it. It is only the latest iteration of a much larger policy legacy concerning the federal government's relationship with educational research—one that existed long before the *What Works* pamphlet. This relationship has been both long and turbulent but not necessarily unguided (for reviews, see Vinovskis, 1998 and St. Pierre, 2006). Since the passing of the *U.S. Office of Education Act of 1867,* which called for the creation of a federal education agency to both research "the condition and progress" of the nation's schools and "diffuse" that research to the public (p. 434), U.S. education policy has always been committed to bridging the gap between educational research and educational practice. Not only is the WWC a manifestation of such a commitment, but it also exemplifies *how* the government has pursued this commitment with increasing authority and epistemic control. This is evidenced by three landmark policies that helped pave its way, which I review in the following pages. At the same time, I also highlight how each policy expanded the government's role—from

sponsor to agenda-setter to expert to gatekeeper—in its relationship with educational research.

***Government as Sponsor and The Cooperative Research Act of 1954***

*The Cooperative Research Act* of 1954 represents government's first formal sponsorship of educational research (Stiles, 1962). Although, prior to 1954, the federal government had sponsored data collection efforts via the U.S Office of Education (which later became the U.S. Department of Education), those efforts were largely for record-keeping purposes. In the early 1950s, an internal review "confirmed the agency's deficiencies in handling statistics and research" (Vinovskis, 1998, p. 76), and in response, *The Cooperative Research Act* was designed to refresh the Office's research capacity. To do so, the Act granted the Office of Education authority to:

> "…enter into contracts or jointly financed cooperative arrangements with universities and colleges and State educational agencies for the conduct of research, surveys, and demonstrations in the field of education." (p. 533)

Such a decision was monumental for educational research, though not necessarily novel in its own right. The agricultural model of Cooperative Extension, which invested heavily in agricultural research and its extension from universities to surrounding communities, had existed since 1862 (Rogers, 1988). Although it took nearly one hundred years for this model to be extended to educational research, *The Cooperative Research Act* still signaled government's growing belief that "[r]esearch, both basic and applied, [was] the key to school improvement" (Stiles, 1962, p. 231).

Along with signaling the government's willingness to invest in educational research, *The Cooperative Research Act* afforded greater government control over both the scope and sharing of educational research. For example, any project seeking federal

funds required approval from a Research Advisory Committee, which reviewed project ideas and vetted them according to a set of criteria. Only those that met the Committee's standards regarding project significance, research design, and economic efficiency could be recommended for funding (Clark & Carriker, 1961). The Act also sponsored several innovative dissemination channels, which ensured, among other things, that reports from all funded research were distributed to sixty regional libraries across the nation. By sponsoring the undertaking of research, as well as its wider dissemination, *The Cooperative Research Act* is emblematic of the government's commitment to bridging the research-practice divide. At the same time, by requiring that sponsored research be preapproved, it is also emblematic of how the research-practice divide was addressed with increasing top-down authority.

***Government as Agenda-Setter and The Elementary and Secondary Education Act of 1965***

A second policy development was the passing of *The Elementary and Secondary Education Act* (ESEA) in 1965. Although ESEA is often recognized for spotlighting educational inequities faced by students from disadvantaged backgrounds (i.e., Title I), the law also expressed faith in research's ability to address these inequalities (i.e., Title IV). Examples of this faith abound. For example, in a report that was later credited for motivating the ESEA legislation (for a detailed history, see Kearney, 1967), the director of President Johnson's Task Force in Education urged that:

> "[w]e need a system for continuous renewal, a system in which appraisal and innovation is built in. That is why references to research and development, to innovation and experiment, appear in every chapter of this report." (Gardner, 1964, as cited in Vinovskis, 1998)

Another example is the dramatic increase in funding for educational research, which swelled from $3 million in 1960 to $100 million in 1967 (Vinovskis, 1998). Most of this money went towards establishing the "system" suggested by the Task Force—a system that, once again, tried to narrow the gap between research and practice by setting its own research agenda.

Broadly speaking, the system envisioned by President Johnson and his Task Force consisted of two novel components (e.g., Guthrie, 1989). The first was the creation of national Research and Development Centers (R&D Centers), which were responsible for conducting basic research on educational best practices while still residing under federal jurisdiction. The second was the creation of Regional Education Laboratories (RELs), which were designed to translate the knowledge produced by the R&D Centers into recommendations for educators. Even today, both components remain a part of the government's educational research repertoire. They are also characteristic of how the Office of Education confronted the research-practice gap with increasing top-down authority. The R&D Centers, for example, were a subliminal attempt to reign in the funding protocols of *The Cooperative Research Act,* which afforded "little opportunity for federal officials to shape the nation's education research agenda" (Guthrie, 1989, p. 6). Indeed, although *The Cooperative Research Act* required federally funded educational research to be reviewed by a Research Advisory Committee, the work was still conducted by independent researchers. Because the R&D Centers were both federally managed and federally funded, they could be instructed to pursue research agendas most compelling to the U.S. Office of Education (i.e., research on curricula development, teaching technologies, and teacher education materials; Kearney, 1967).

### *Government as Expert and The No Child Left Behind Act of 2001*

An overview of this policy legacy would be incomplete without mentioning the *No Child Left Behind Act of 2001* (NCLB). NCLB headlined a package of policy developments responding to the federally-commissioned *A Nation at Risk* report (1983), which raised alarms about the nation's education system and its inferiority on the international stage. These policy developments doubled-down on the government's belief that the rigorous execution and dissemination of educational research was the best way to initiate reform. Indeed, the passing of the *Education Research, Development, Dissemination, and Improvement Act of 1994*, as well as its subsequent establishment of the *National Education Research Policy and Priorities Board,* are case and point.

NCLB rode on the coattails of these developments; in fact, it built upon them by facilitating even greater government control over educational research (Zoellner, 2010). Perhaps most infamously, NCLB established a federally approved definition of "scientifically-based research." The definition formalized a set of standards constituting "scientifically-based research," among them the requirement that research be:

> "…evaluated using experimental or quasi-experimental designs in which individuals, entities, programs, or activities are assigned to different conditions and with appropriate controls to evaluate the effects of the condition of interest, with a preference for random-assignment experiments, or other designs to the extent that those designs contain within-condition or across-condition controls." (p. 1965)

By writing these standards into law, the federal government appointed itself an expert in what constitutes a rigorous research approach. Some members of the research community applauded such a step (Slavin, 2005), while others—particularly those who resisted positivism as a scientific epistemology—condemned it (e.g., Lather, 2004). In sum, in its effort to bring "what works" into educational practice, NCLB granted the government

epistemic authority over "what counts" as educational research in the first place (Zoellner, 2010).

***Government as Gatekeeper and The What Works Clearinghouse***

Shortly thereafter, the Bush administration followed NCLB with the *Education Sciences Reform Act of 2002*, which brought the *What Works Clearinghouse* (WWC) into fruition. The WWC remains one of three core IES research dissemination initiatives, along with the RELs and the Education Resources Information Center (ERIC) (NCEE, n.d.). Like its predecessors, the What Works Clearinghouse is concerned with bridging the research-practice divide in order to spur educational reform. Its approach to doing so is unique, however, in that it combines the dissemination focus of *The Cooperative Research Act* with the "agenda-setting" and "expert" roles set forth by ESEA and NCLB, respectively. The WWC's approach is self-described as follows:

> "The What Works Clearinghouse (WWC) addresses the need for credible, succinct information by identifying existing research on education interventions, assessing the quality of this research, and summarizing and disseminating the evidence from studies that meet WWC standards." (U.S. Department of Education, 2019b, p. 1).

Given this description, the WWC can be thought of as a government-controlled systematic review initiative. Systematic review has become a favorite strategy for initiatives trying to bridge the research-practice gap because of its ability to synthesize nuanced findings from several individual studies into a simple, generalized conclusion—an end result that is thought to more effectively communicate research information to users (e.g., Oakley, 2003; Schlosser, 2006; Paulsell, Thomas, Monahan, & Seftor, 2016). Instead of reviewing all research on a given topic before generating its conclusion, however, the WWC's approach involves a gatekeeping step that ensures that only

research of "sufficient" quality is reviewed. Although there is precedent for doing so as part of the systematic review process (see Slavin's [1986] "best-evidence" approach to systematic review), there is far less precedent for that gatekeeping step to fall within the government's jurisdiction. As such, by controlling the identification of research, the review of that research, and its eventual dissemination, the WWC allows the Institute of Education Sciences to act as a gatekeeper between educational research and its potential users. This also means that the government is making judgements *on behalf of* educators about what research is credible and relevant to practice.

**The Theoretical Context**

Nevertheless, the irony of the WWC's political legacy is that its largely technocratic approach to bridging the research-practice gap fundamentally conflicts with how research and theory suggest doing so. This was true as the government's efforts intensified, and it remains true today. Indeed, as recently acknowledged by Gorard and colleagues (2020), many "recommendations for the implementation of evidence [in educational practice] are not themselves based on good evidence" (p. 575). This point will be elaborated upon in Chapter 2. For the remainder of this section, however, I will instead focus on how the WWC's approach diverges from theory—and more specifically, from theories of knowledge utilization.

The 1970s were a "golden age" (Bogenschneider, Corbett, & Parrott, 2019, p. 130) in the social scientific study of the research-policy and research-practice relationships (for a discussion, see Newman, Cherney, & Head, 2016). Some of social sciences' most prolific thinkers, including Carol Weiss (1979, 1980) and David Cohen (e.g., Lindblom & Cohen,1979), can be credited with spearheading this topic of study—a

topic that continues to receive scholarly attention in applied and professional fields (Smith & Wilkins, 2018; Tkachenko, Hahn, & Peterson, 2017). As this body of work has grown, so too have attempts at theory-building, especially in education (for reviews, see Nutley, Jung, & Walter, 2008; Tseng, 2012). Here I review one prominent theoretical tradition, which began as Two-Communities theory (Caplan, 1979; Dunn, 1980) before branching into newer variants (Bogenschneider et al., 2019; Farley-Ripple et al., 2018). This tradition undergirds the current study.

There are several reasons why I have elected to use Two-Communities theory, as well as its subsequent iterations, as my guiding theoretical position. First, despite its inception half a century ago, Two-Communities theory remains current. Bogenschneider et al. (2010, 2019) and Farley-Ripple et al. (2018), both of whom have proposed some of the most prominent knowledge utilization models in recent years, explicitly recognize Two-Communities theory as an influence. Second and relatedly, Farley-Ripple and colleagues' (2018) model is currently serving as the guiding framework for the IES's Center for Research Use in Education, which suggests that the federal government is receptive to the premises of Two-Communities theory (Center for Research Use in Education, n.d.). Third, both Two-Communities theory (e.g., Dunn, 1980), as well as its more contemporary descendants (Bogenschneider et al., 2019; Farley-Ripple et al., 2018), have repeatedly identified five domains where fractures between research producers and research users exist (see Table 1.1 for details). Three of these domains—which, in Farley-Ripple et al.'s (2018) latest iteration, include (a) the problems that research addresses, (b) the nature and quality of research, and (c) the usefulness of research products—align with issues that the WWC is specifically designed to target.

This alignment is yet another reason why the Two-Communities tradition has utility for a project such as mine.

### *Caplan's Two-Communities Theory*

The work of Nathan Caplan and colleagues (1975, 1979) was some of the first to theorize about why social science was so often neglected in policymaking. It attributes the non-utilization of social science to cultural differences between researchers and policymakers. Specifically, proponents of this theory maintain that:

> "…social scientists and policy makers live in separate worlds with different and often conflicting values, different reward systems, and different languages. The social scientist is concerned with 'pure science' and esoteric issues. By contrast, government policy makers are action-oriented, practical persons concerned with obvious and immediate issues." (Caplan, 1979, p. 459)

The research-practice gap is, in effect, the result of each culture refusing to assimilate to the other. In order for the utilization of social science to improve, Caplan (1979) calls for "collaborative arrangements" (Caplan, 1979, p. 468) between knowledge producers and knowledge users. Although Caplan failed to elaborate on what those arrangements might look like, he did call for researchers to work alongside decisionmakers when deciding what research questions to pursue.

Contemporaries such as William Dunn found Caplan's thinking to be a compelling metaphor; in its original form, however, Dunn did not believe that it constituted a testable scientific theory. Accordingly, Dunn (1980) attempted to elevate Caplan's thinking from "metaphor" to "theory" by generating and testing a series of propositional models. The nature of those models is worth elaborating upon here because of how they reappear in later theory (e.g., Bogenschneider & Corbett, 2010; Farley Ripple et al., 2018). In particular, Dunn detected a host of theoretical assumptions

embedded within Caplan's claims, and, in doing so, categorized those assumptions into five conjoint models of knowledge use. Those models included the *product-contingent model*, which suggests that knowledge will be used when it is shared in ways that are desirable and accessible to intended users; the *inquiry-contingent model*, which suggests that knowledge will be used when it is generated using procedures (e.g., research designs, methods, analytical techniques) perceived as trustworthy by intended users; the *problem-contingent model*, which suggests that knowledge will be used when it targets problems of interest to intended users; the *structure-contingent model,* which suggests that knowledge will be used when organizational structures facilitate, rather than impede, use; and, the *process-contingent model,* which suggests that knowledge will be used when knowledge producers and intended users interact in ways that stimulate use.

Importantly, Dunn articulated these models so that they could be empirically tested. Using data from their multi-year case survey of knowledge use in organizational change processes, Dunn and Swierczek (1977) analyzed the extent to which published accounts of knowledge use supported and/or falsified the models specified earlier. Some, though not all, of his findings are in concert with Caplan's thinking. As an example, when studying the *inquiry-contingent model,* Dunn found that knowledge users' perceptions of validity/reliability were associated with research use, yet there was no association between knowledge utilization and "the use of classical procedures of variance control" (p. 530). What this suggests is that practices used to ensure information quality in the culture of knowledge producers (e.g., researchers) are not the same indicators of quality in the culture of knowledge users (e.g., policymakers). Or, as stated by Dunn himself:

> "This suggests a wider conception of reliability and validity than that implied by the simple use or nonuse of experimental or quasi-experimental designs, random sampling, or quantitative methods." (p. 531)

A second pattern of findings relates to the *process-contingent model*. Specifically, Dunn found that "involv[ing] all relevant stakeholders in the problem definition, goal-setting, and evaluation phases" (p. 532) helped facilitate knowledge utilization. This finding aligns with Caplan's suggestion that "collaborative arrangements" between knowledge-producing and knowledge-using communities were critical to bringing research into policymaking.

***Bogenschneider et al.'s Community Dissonance Theory***

In the years following Caplan and Dunn's work, social scientists have continued to both extend (e.g., Wingens, 1990) and critique (e.g., Newman, Cherney, & Head, 2016) Two-Communities Theory. One recent extension is Bogenschneider and colleagues' Community Dissonance Theory (Bogenschneider & Corbett, 2010; Bogenschneider et al., 2019), which begins with the same premise that Caplan did. Specifically, Bogenschneider and colleagues conceptualized:

> "...those who conduct studies and those who consume them as separate communities that live on different islands and see the world through different lenses. Put simply, it is as if the inhabitants of each island speak different languages, march to different drummers, and see the world through different lenses." (Bogenschneider, Corbett, & Parrott, 2019, p. 129)

Similar to Dunn (1980), however, Bogenschneider felt compelled to develop her own iteration of Two-Communities Theory because it was "overly simplistic" in its original form. Community Dissonance Theory sought to remedy this by extending Caplan's thinking in two main ways. First, in order to better capture the complexity of real-world policymaking, Bogenschneider and Corbett (2010) acknowledged the countless other

actors involved in bringing social science to the forefront of policymaking. Whereas Caplan's work relied on a "two islands" metaphor—with academic researchers and policymakers existing in their own distinct communities—Bogenschneider and Corbett imagined an "archipelago" of island chains inhabited by various types of research producers (e.g., think tanks, government agencies, evaluation consultants) and research users (e.g., lobbyists, advocacy groups, philanthropic organizations). Each of these groups still maintains its own institutional and professional culture, but their mere existence is recognized by Community Dissonance Theory.

Community Dissonance Theory also extends Two-Communities Theory by more deeply investigating the cultural dissonance that occurs when inhabitants of different islands interact. Doing so is important because Community Dissonance Theory holds a second, more optimistic premise. Specifically, it contends that interisland travel is possible, and moreover, that these communities are "not destined to remain isolated on their own islands" (Bogenschneider et al., 2019, p. 142). To them, doing so requires that researchers better understand the practice community. As such, Bogenschneider and colleagues have undertaken a decade-long research program investigating how research is accessed, perceived, and used in policy settings. In a recent review of that work, Bogenschneider, Corbett, and Parrott (2019) outline five points of friction between policy and research cultures. Several of those frictions align with the conclusions of Dunn (1980). As an example, just as Dunn found that the indicators of research quality held by researchers (e.g., experimental designs) had little, if any, bearing on the eventual research use of policymakers, Bogenschneider and colleagues (2019) identified differences in how research-producing and research-using actors assess the credibility of evidence. They

refer to these differences as the "epistemology frameworks" (p. 143) privileged by either culture. Whereas researchers are likely to rely on the "canons of science" (p. 144) when assessing credibility, legislators rely on their intuition and understanding of their colleagues' underlying motives. What's more, policymakers prefer to assess the credibility of information through face-to-face negotiations with their colleagues—not through written reports. Thus, although Community Dissonance Theory maintains that this dissonance is reconcilable, it still demonstrates the extent to which the policy and research islands remain isolated from one another.

### *Farley Ripple et al.'s Model*

While the ideas of Two-Communities Theory and Community Dissonance Theory have been applied to policy issues writ large, they have also been applied in explanatory models of specific policy domains. An example in the domain of education can be found in the work of Farley-Ripple and colleagues (2012, 2018), which has since culminated in a conceptual framework (e.g., Farley-Ripple et al., 2018; see also CRUE, n.d.) that, along with guiding the IES's Center for Research Use in Education, also guides this project. Like its predecessors, this framework contends that:

> "…the cultures, contexts, and systems in which researchers and practitioners operate, including institutional goals and professional norms and expectations, differ significantly." (p. 237)

It also recognizes, as Bogenschneider and Corbett (2010) did, that drawing a simple dichotomy between "researchers" and "practitioners" is reductionistic given the diversity of stakeholders involved. Perhaps its biggest contribution to previous theories is that it clearly articulates the bidirectionality of the research-practice gap. Indeed, Caplan (1979) reasoned that "collaborative arrangements" (p. 468) were important to integrating the

research and practice communities, and Bogenschneider and Corbett (2010) called for "open dialogue" (p. 145) between both communities. Yet both failed to convey a point that is central to Farley-Ripple et al.'s model—namely, that "research can inform practice *and* practice can inform research" (p. 242, emphasis mine). By taking this perspective, Farley-Ripple and colleagues also distinguish themselves from scholars (e.g., Kaestle, 1993) who principally blame the state of educational research for the research-practice gap.

Farley-Ripple and colleagues model the research-practice gap as a gulf between two communities—the "research community" and the "practice community" (see Figure 1.1). That gulf exists because each community maintains certain perspectives and/or assumptions that are incompatible with the perspectives and/or assumptions of the other. In contexts where the gap between researchers' and practitioners' perspectives/assumptions is large, research use will be shallow and infrequent; conversely, in contexts where the perspectives/assumptions of the research and practice communities converge, research will be consulted with greater regularity. Five core categories of assumptions/perspectives are identified by Farley-Ripple et al., (2018), who directly aligned their five categories with Dunn's (1980) five models mentioned earlier (see Table 1.1). One of these categories is referred to as "the nature and quality of research." When describing this category, the authors directly reference an assumption about research quality held by the WWC, and moreover, how that assumption might differ among stakeholders:

> "For example, the WWC employs standards that place great weight on internal validity for drawing causal inference (i.e., randomized experiments). In contrast, school-based decision-makers often prefer evidence from organizations or

> contexts (e.g., demographics, location, performance) similar to their own, regardless of study design." (Farley-Ripple et al., 2018, p. 240)

In such a case, where there is a lack of alignment between the assumptions of the WWC and those of its stakeholders, the model would predict shallow and infrequent research use. The extent to which existing literature supports this prediction, as well as the alignment between researchers and practitioners across Farley-Ripple et al.'s (2018) other four categories, will be thoroughly examined in Chapter 2.

## Framing This Project

Farley-Ripple and colleagues' (2018) observation about the WWC is an important one. Not only does it embody the central implication of the preceding section—namely, that there is divergence between how U.S. education policy has attempted to bridge the research-practice gap and how theorists have suggested doing so—but it also helps introduce three points of inquiry that, together, frame this project. First, by recognizing a mismatch between the WWC's standards of research evidence and those favored by school-based decisionmakers, Farley-Ripple et al.'s (2018) model implies—but does not prove—that the WWC is failing to maximize its impact on the research-practice gap. As such, it begs the first question pursued by this project: to what extent is the WWC "working" to bridge the research-practice gap? Second and relatedly, Farley-Ripple et al.'s (2018) model suggests that the WWC's impact may be lacking precisely because its assumptions are misaligned with its stakeholders' own; even so, it does not elaborate on this possibility. As such, it begs a second question pursued by this project: in what ways are the WWC's programmatic assumptions facilitating and/or hindering its attempts to bridge the research-practice gap? Third, Farley-Ripple et al.'s (2018) model infers that the WWC could improve its impact by aligning its approach with the research-related

assumptions and perspectives held by educators. It does not, however, make specific recommendations for how the WWC could do so. As such, it begs a third question pursued by this project: what specific changes could the WWC make to more effectively bridge the research-practice divide?

Answering these questions in their entirety would require a full program of research, not a single project. Therefore, the purpose of this dissertation is to begin the inquiry process by thinking evaluatively about these three questions. Thinking evaluatively means subscribing to central tenants of "evaluative thinking" (ET; for a recent discussion, see Vo & Archibald, 2018). ET is not a new concept, nor is it a particularly well-defined one. Evaluators have long wondered if a certain "logic" (Scriven, 1980) or "habit of mind" (Weiss, 1988, as cited in Vo, Schreiber, & Martin, 2018) might characterize the evaluation discipline and help differentiate it from other, similar ones. In addition, more contemporary scholars have tried to identify the various competencies (for a review, see Jacob & Boisvert, 2010), knowledge bases (Arbour, 2020), and/or ethics (e.g., Morris, 2015) that together embody the discipline's "true north" (Vo et al., 2018, p. 30) Theorists have made much progress in the past decade, particularly in distinguishing between what ET is and what it is not (Vo & Archibald, 2018). Some call it "critical thinking applied in the context of evaluation" (Buckley, Archibald, Hargraves, & Trochim, 2015); others add that it is a "problem-solving approach" (Vo, 2013); still others have modelled it as the cognitive process of moving between values and valuing (Vo et al., 2018). That said, an agreed-upon definition still does not (and may never) exist. Even so, the concept itself is not worthless. As I hope to convey below, the values/motivations, methodology, and intended contributions of this

work are all guided by ET—not necessarily by its definition, but by the principles captured therein. By doing so, this project is the first to deliberately chronicle an application of evaluative thinking in practice.

**Values/Motivations**

ET recognizes that evaluation science is never value-free. As long-held by Scriven (1972) and others (e.g., House & Howe, 1999), the act of evaluating something is, by definition, the act of judging that something's value. Recent conceptualizations of ET agree. For example, Vo and colleagues (2018) proposed that "the overlap between the values that drive *evaluation* and the value *derived from* evaluation in a democratic society differentiates the field from many other professions" (Vo et al., 2018, p. 40, italics in original). A few years prior to that, Vo (2013) differentiated the classic definition of evaluation—that is, the process of investigating the worth or merit of an object (The Joint Committee, 1994)—from her own working definition of ET, which sees it as the process at "arriv[ing] at contextualized value judgments" (p. 107). These judgments are "contextualized" by, among other things, the values held by the evaluator and the values privileged in whatever context the evaluation was conducted in. In a similar way, Vo and colleagues' (2018) conceptual model of ET depicts it as a process that begins with values (i.e., the "societal and discipline-specific standards that serve as catalysts for evaluation") and arrives at the valuing of an evaluand (i.e., an "ascription of merit, worth, significance, importance") (p. 37). As such, ET acknowledges the role values play in both motivating and guiding the evaluation process.

Vo et al.'s (2018) efforts to conceptualize ET found democracy to be a commonly cited value in the ET literature; it features heavily in this work as well. Indeed, this

project emerged from the observation that federal approaches to bridging the education research-practice gap are more often technocratic than they are democratic. In fact, they often diverge from how one of democracy's greatest champions, John Dewey, suggested doing so. In an essay entitled *The Sources of a Science of Education*. Dewey concluded that the ultimate source of a science of education is democratic participation rather than technocratic authority:

> "The sources of educational science are any portions of ascertained knowledge that enter into the heart, head and hands of educators, and which, by entering in, render the performance of the educational function more enlightened, more humane, more truly educational than it was before." (1929, p. 54)

In other words, the only way to ensure that educational research finds its way to educational practice is to engage the hearts, heads, and hands of practitioners themselves. Yet, despite Dewey's plea to place practitioners at the helm of educational science, the federal government has—and continues to—approach educational science with increasing top-down authority. It was this initial realization, especially as it applied to the WWC, that catalyzed this project.

A second, related value motivating this work is that of accountability. Vo and colleagues (2018) found the value of accountability, particularly in the context of Western democracy, to be abundantly referenced in previous conceptualizations of ET. For ET, no object is too sacred, too protected, or too reputable to be its subject. In much the same way, this project is motivated by the belief that even other evaluators should be held accountable—a belief that also underpins the notion of meta-evaluation. Put simply, meta-evaluation is the process of ensuring accountability by evaluating the evaluator (Stufflebeam, 2001). It also encompasses the process of evaluating an evaluative body, such as the WWC (e.g., Scriven, 1994). Although the meta-evaluation tradition was

formalized over half a century ago (Scriven, 1969; Stufflebeam, 1974), it has enjoyed a recent resurgence of recognition (e.g., Scriven, 2009). As an example, when the third edition of the *Joint Committee on Standards for Education Evaluation* was published in 2010, it called for meta-evaluation in two of its three Accountability Standards. The reasons for conducting a meta-evaluation are numerous, among them being a desire to build credibility, enhance quality, or identify areas for improvement (for a review, see Stufflebeam, 2001). Given that the WWC has received remarkably little evaluative attention, these reasons for conducting a meta-evaluation, as well as the broader values of democracy and accountability, were motivators of the current study.

**Positionality**

ET also demands self-reflection. Indeed, the willingness to acknowledge and investigate one's own values—and, in turn, one's own assumptions and biases—is a hallmark of evaluative thinking (Vo et al., 2018, p. 41). Vo and colleagues (2018) explicitly refer to this as "scrutinizing one's own positionality" (p. 43). Evaluators have long recognized the importance of positionality when working within certain paradigms (e.g., constructivist evaluation; Lincoln & Guba, 1989) or utilizing certain methods (such as in qualitative evaluations; Greene, 1998). Prominent evaluation theorists and professional organizations have continued to advocate self-reflection and transparency. For example, some of Mertens' (2015, 2016) recent work on evaluation paradigms calls for "making explicit the assumptions that evaluators make about themselves and their roles as evaluators" (2016; p. 103). Likewise, in the latest iteration of their *Guiding Principles for Evaluators* (AEA, 2018), the American Evaluation Association instructs

evaluators to explicate their values and assumptions in order to satisfy its principles of Systematic Inquiry and Integrity.

To accommodate professional ethical principles as well as the conceptual principles of ET, I elaborate on two aspects of my positionality here. The first is necessary given the Two-Communities (e.g., Caplan, 1979) theoretical tradition informing this project. Specifically, it is important to recognize myself as a knowledge producer—that is, as part of the "research community" responsible for generating knowledge to be used in practice. Aside from serving as a postsecondary teaching assistant, I have never taught in a K-12 setting, nor have I served in an administrative role. Rather, over the past decade, I have been a student of and contributor to the social sciences, conducting research in the fields of social psychology, developmental psychology, human factors research, and educational program evaluation. During that time, however, I have maintained an interest in how findings from social science are used by policymakers and practitioners. Acknowledging this is important given that this project seeks to improve the WWC's usefulness to educators—not social scientists such as myself. Although my ultimate suggestions for doing so (see Chapter 5) are informed by what the scholarly community has learned about educators' research engagement, they are tempered by my lack of insider knowledge. (For a compelling commentary of the WWC written by a member of the practice community, see Sheldon, 2016).

A second aspect of my positionality that requires mentioning is the WWC-related biases I have developed over the past three years. I first learned about the WWC as a graduate student in social psychology, and more specifically, as I became interested in educational interventions developed according to social psychological principles (for a

review, see Yeager & Walton, 2011). At that time, I was often dismayed by the dearth of intervention research employing randomized sampling procedures and experimental designs. My excitement about the WWC's standards of rigor were documented in a cover letter I drafted at that time. In it, I wrote:

> "First, I am excited by the possibility of working and learning in the [name removed]. The [name removed] approach to evaluation is exactly what I desire in my training: rigorous (i.e., based on What Works Clearinghouse standards), locally focused, and committed to utilization-focused evaluation practices."

As my graduate training progressed, however, I was challenged to consider both the value of other research designs as well as the limitations inherent to experimental designs. It was in an Ethics in Education Decision-Making course where I was invited to confront my biases most intentionally. As part of that course's final project, I explored the extent to which the WWC and its review processes satisfied the ethical standards of human subjects research as formalized in the Belmont Report. My paper concluded the following:

> "The purpose of this paper was to question the ethicality of three standards used in the What Works Clearinghouse's (WWC) research review process. The standards in question include: (1) the privileging of desired outcomes at the expense of undesirable risks, (2) the privileging of randomized control designs at the expense of validity and stakeholder needs, and (3) the privileging of quantitative evidence at the expense of contextualizing qualitative evidence and possible acts of corruption. I have charged each of these three standards as violating the guiding principle of beneficence in human subjects research, which has a rich legacy in both theoretical and applied ethics." (Nelson, 2019, p. 25)

Accordingly, though I still harbor biases about the WWC, the nature of those biases has changed. I now look at the WWC less with reverence and more with skepticism. As an application of evaluative thinking, however, this project provides an opportunity to pit these biases against the careful and systematic review of evidence.

## Methodology

ET is methodologically agnostic. To be clear, this is not the same as saying that ET is methodologically ambivalent, or that methodological decisions are unimportant when thinking evaluatively. What it means is that ET rejects the notion of a methodological "gold standard." As asserted by Vo and Archibald (2018), placing randomized sampling and experimental designs at the top of some methodological hierarchy represents "an extreme lack of evaluative thinking" (p. 144). What matters most in ET is not blind faith in a methodological hierarchy, but honest efforts to ensure methodological appropriateness. Methods are appropriate when they are aligned to the nature of one's questions and the desired claims one hopes to make (Archibald & Buckley, 2018).

To reiterate, in this project, I pursue questions about (a) the WWC's impact, (b) its alignment with educators' perspectives and assumptions about educational research, and (c) the changes it can make to maximize its impact. Doing so requires a methodology nimble enough to do several things at once. First, to formulate a claim about the WWC's impact, the methodology must be able to approximate cause-effect logic while also accommodating the messy, real-world policy environment in which it operates. Such a setting makes traditional methods of causal inference, such as experimental designs, untenable (e.g., Lipsey & Cordray, 2000). Second, to determine how the WWC's assumptions align with educators' own, the methodology must recognize the significance of those assumptions (e.g., Nkwake 2013; Nkwake & Morrow, 2016) as well as facilitate the identification of those assumptions in the first place. Third, to generate recommendations for program improvement, the methodology must be able to pinpoint

specific breakdowns in the WWC's underlying logic, as those breakdowns are where program redesign will have the greatest impact.

Given these requirements, this project uses a methodology called contribution analysis (Mayne, 2008, 2012, 2019; Better Evaluation, 2020). The intricacies of contribution analysis and its place in the theory-based evaluation tradition (Chen & Rossi, 1989) are discussed in Chapter 3. For now, it's enough to realize that contribution analysis satisfies all the methodological requirements outlined above. It does so by guiding evaluators through both the articulation and scrutinization of a program's theory of change—that is, the often-implicit logic underlying how a project's activities are expected to result in certain desirable outcomes (for an overview, see Rogers, 2008). This is done by evaluating the veracity of the theory of change according to available evidence. By evaluating the WWC through its theory of change, contribution analysis allows me to make defensible claims about its impact, its assumptions, and its capacity for improvement while also accommodating the resource and time constraints inherent to this project. For example, contribution analysis allows for claims about program impact to be made based on the extent to which empirical evidence validates a program's theory of change; it does not require experimentation and random assignment. Contribution analysis also provides a framework to help evaluators identify programmatic assumptions as well as the role(s) they play as part of program theory; it does not require that program officials share these assumptions with evaluators, nor does it require them to be aware of these assumptions in the first place. Finally, by forcing evaluators to consider program theory, contribution analysis encourages evaluators to look beyond questions of "if" a program works (or not) to questions of "how" or "why" a program works (or doesn't);

knowing this is critical when identifying specific points for program improvement. Because contribution analysis is well-suited for this project's questions and desired claims, its use is justified by ET's principle of methodological appropriateness.

**Intended Contributions**

Finally, ET privileges learning. It does not set out to loudly congratulate effective programs or slap the wrists of struggling ones. Rather, it approaches evaluation as a learning opportunity, such that "the users/clients function as students or learners, the evaluator as teacher; the evaluation process and its findings become the curriculum" (King, 2007, as cited in Vo et al., 2018, p. 42). It shares this sentiment with theorists who have considered evaluation to be a form of "pedagogy" (e.g., Patton, 2017). One way ET's learning orientation is visible is in its emphasis on evaluation use. As noted by Vo et al. (2018), previous conceptualizations of ET describe how it "support[s] the long-term attainment of evaluation use" (p. 42). Not only does this refer to using evaluation findings, but also to using the evaluation process (process use; Patton, 1998, 2007) as a pedagogical exercise—namely, to teach a program or organization how to evaluate itself.

In a similar way, I intend this project's contribution to be multifaceted. Not only is it well-positioned to make contributions to both scholarship and practice, but the intended contributions are meant to be both findings- and process-oriented. Take, for example, the possible scholarly contributions of this project. Its findings help fill a gap in the literature while also contributing to an ongoing conversation among scholars. Specifically, my study responds to a dearth of research on the extent to which practitioners have heard of and/or used the WWC. It also builds upon an existing conversation in the literature, in which scholars have widely critiqued the WWC

(Schoenfeld, 2006; Stockard, 2010; Stockard & Wood, 2017; Slavin, 2017). These critiques are limited, however, because they amplify the perspectives of scholars rather than the perspectives of educators. Using an original survey of new teachers as well as previously collected focus group data, my study adds to this conversation by building an argument using educators' own perspectives about the WWC. Another scholarly contribution emerging from this project is that it offers an empirical test of Farley-Ripple et al.'s (2018) theoretical framework. In effect, it extends Farley-Ripple's (2018, p. 241) own ponderings about if and how the model will behave when applied to research brokers (such as the WWC) rather than researchers themselves. Finally, it offers a process-specific contribution that is also relevant to the scholarly community. Although previous work has documented how evaluative thinking can be fostered among non-evaluators (Bhatti, Dahlgaard, Hansen, & Hansen, 2015; Archibald, Sharrock, Buckley, & Cook, 2016), there are still no clear examples of what evaluative thinking looks like in practice. This project changes that. By providing a literal step-by-step example (see Chapter 4) of how to approach the evaluative thinking process—particularly through the use of contribution analysis—I hope that future students find it useful when structuring their own ET pursuits.

The practical implications of this work are varied as well. The end goal of this project is not to criticize, but rather to develop a list of recommendations for improving the WWC. I envision policymakers being more receptive to these recommendations than to the largely theoretical critiques leveled by Biesta (2007, 2010) and Lykins (2012). Even if my recommendations are stonewalled, though, I at least hope that this project encourages the IES and/or WWC to undertake their own self-evaluation process. The

U.S. Government Accountability Office (GAO) last evaluated the WWC in 2010, and considering the evaluative nature of their recommendations (e.g., to "develop performance measures related to product usefulness"; p. 42), a subsequent evaluation seems prudent. This project also serves the practical purpose of informing teachers—namely, those participating in a short survey as part of the contribution analysis—of the WWC and its mission. Doing so sounds trivial, but I believe that it is important. Not only could it increase the WWC's utilization, but it might also encourage educators to offer their own recommendations for improvement.

## Organization of this Dissertation

The above chapter has introduced my project's main aims, as well as the political and theoretical contexts informing them. Chapter 2 builds from this introduction by reviewing literature on educators' research engagement. Specifically, it uses the five assumption categories specified by Farley-Ripple et al. (2018) to structure a review of what is known about educators' assumptions, perspectives, and use of educational research. Takeaways from this review will, in turn, assist with evaluating the WWC's theory of change in Chapter 4. Following the literature review, Chapter 3 elaborates on my project's methodological approach. I describe the technique of contribution analysis (Mayne, 2008, 2012b, 2019) in greater detail before overviewing the evidence sources used as part of this project's analysis. While doing so, I pay special attention to two evidence sources—a survey of early-career educators and a collection of focus group transcripts—given that those sources were developed/analyzed specifically for this project. Chapter 4 presents findings from the contribution analysis. Finally, Chapter 5 concludes by translating my findings into recommendations for WWC improvement.

## Chapter 1 Tables and Figures

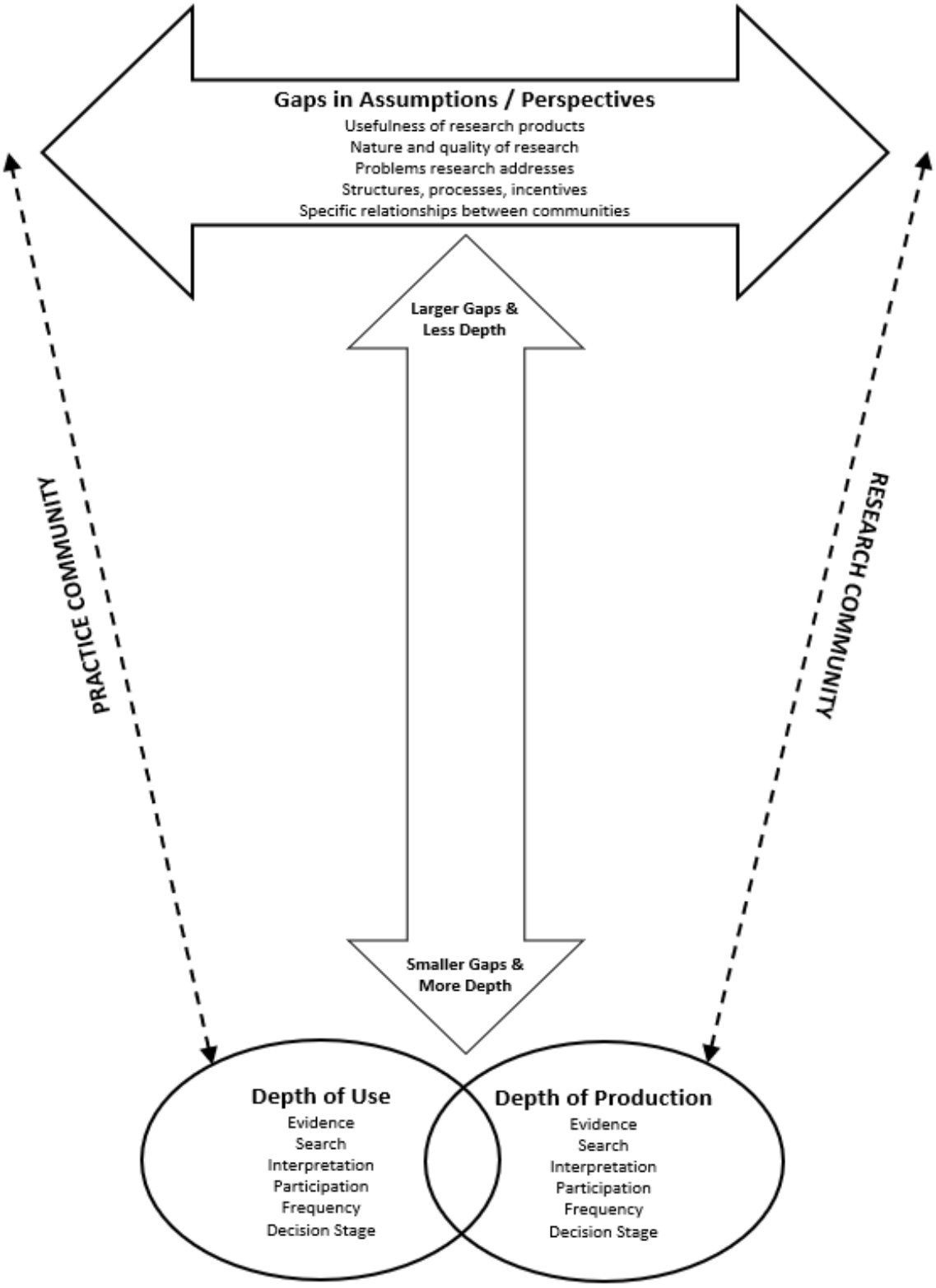
Table 1.1 - Dimensional Alignment Between Three Models in the Two-Communities Theoretical Tradition

| **DIMENSION** | **THEORIES OF USE** | | |
|---|---|---|---|
| | **Two-Communities Theory** contends that variations in policymakers' research use are a function of… | **Community Dissonance Theory** contends that, in order to facilitate research use, the research community must better understand the policy community. Specifically, researchers must reflect on the following questions: | **Farley-Ripple et al.'s (2018) Model** contends that educators' research use depends on: |
| Topic(s) Addressed | *Problem-Contingent Model*: "…the nature and the complexity of policy problems whose resolution may require particular kinds of knowledge" (Dunn, 1980, p. 522). | *Focal Interests:* "Which substantive topics or challenging problems attract interest and attention?" (Bogenschneider et al., 2019, p. 135) | *Problems Research Addresses:* "…the extent to which the evidence produced by the research community is timely and relevant to the problems confronting real schools" (Farley-Ripple et al., 2018, p. 240). |
| Indicator(s) of Quality | *Inquiry-Contingent Model:* "…the quality of procedures used to monitor and evaluate the diffusion, adoption, and implementation of innovations" (Dunn, 1980, p. 522). | *Credibility of Evidence:* "Which processes and methods do individuals use to know what to believe?" (Bogenschneider et al., 2019, p. 135) | *Nature and Quality of Research:* "…how the two communities value different qualities of research, including issues related to internal and external validity as well as conclusiveness of findings" (Farley-Ripple et al., 2018, p. 240). |

Table 1.1 (continued) - Dimensional Alignment Between Three Models in the Two-Communities Theoretical Tradition

| | | | |
|---|---|---|---|
| Products Used for Dissemination | *Product-Contingent Model:* "…the form in which information is embodied or stored" (Dunn, 1980, p. 521). | *Interactional Preferences:* "Which communication channels are preferred?" (Bogenschneider et al., 2019, p. 135) | *Usefulness of Research Products:* "…the degree to which products produced and valued by researchers aligns with those preferred by practitioners" (Farley-Ripple et al., 2018, p. 240). |
| Relationships Fostered | *Process-Contingent Model:* "…the nature and types of interaction among social scientists and policy-makers in `various phases of the policy-making process. (Dunn, 1980, p. 527) | *Interactional Preferences: "*How important are interpersonal relationships, and how do relationships compare to other influences on getting the job done?" (Bogenschneider et al., 2019, p. 135) | *Relationships Between Communities:* "…the relationship between communities in the production of research and in education decision-making" (Farley-Ripple et al., 2018, p. 241) |
| Structures Present | *Structure-Contingent Model:* "...the differences in the formal structure, procedures, and incentive system of organizations." (Dunn, 1980, p. 526) | *Work Environment:* "Which kinds of pressures do inhabitants face in their world? What is the dominant pace of activity and to which kinds of time pressure are they exposed?" (Bogenschneider et al., 2019, p. 135) | *Structures/Processes/Incentives:* "…the context in which researchers and practitioners operate and what influences researchers to produce certain kinds of research and what influences practitioners to use research or other evidence" (Farley-Ripple et al., 2018, p. 240). |

Figure 1.1 - Depiction of Farley-Ripple and Colleagues' Conceptual Framework. Adapted from Farley-Ripple el al. (2018).



# CHAPTER 2 – LITERATURE REVIEW

As explained in Chapter 1, federally sponsored efforts to bridge the educational research-practice gap have generally preferred technocratic, top-down enforcement to democratic, bidirectional collaboration. Doing so ignores the guidance of Farley-Ripple et al's (2018) conceptual framework, which suggests that knowledge producers—as well as "intermediary organizations" responsible for disseminating that knowledge (p. 241)—must accommodate the attitudes and assumptions of knowledge users. Though ignoring these attitudes and assumptions is regrettable, it is also understandable given how little we know about how educators access research, how they vet it for quality, how they perceive its relevance to practice, and how they eventually use it. Indeed, over time, reviews of educators' research use have commented on the literature's limited quantity, quality, and scope (Hemsley-Brown & Sharp, 2003; Dagenais, Lysenko, Abrami, Bernard, Ramde, & Janoz, 2012; Cain, 2016; Schaik, Volman, Admiraal, & Shenke, 2018; Gorard, See, & Siddiqui, 2020). Even so, the extant literature still provides insights into how the WWC's approach both aligns with and diverges from the attitudes and assumptions of its stakeholders.

The purpose of this chapter, then, is to review what is known about educators' research engagement while also reviewing what the WWC does to elicit that engagement. To organize this review, I evoke each of the five dimensions proposed by Farley-Ripple and colleagues (2018) as possible areas of dissonance between educators and the WWC. These dimensions include (a) problems addressed by research, (b) nature/quality of research, (c) usefulness of research products, (d) relationships between communities, and (e) structures/processes/incentives (see Figure 1.1). Within each dimension, literature

addressing educators' research engagement is discussed, followed by an overview of how the WWC enacts each dimension; literature published after 2000 is prioritized. By overviewing both the literature and the WWC procedures together, I do justice to Farley-Ripple et al.'s (2018) framework by allowing the reader to compare how the WWC has enacted these dimensions as part of the "research community" with how practitioners have approached these dimensions as members of the "practice community." As a reminder, differences in the assumptions and perspectives of these communities are expected to hinder both the frequency and depth of educators' research use.

Furthermore, in an effort to adhere to Farley-Ripple et al.'s (2018) framework, I bracket the scope of this literature review by using the word "educator" in a limited way. Just as Farley-Ripple and colleagues' (2018) definition of the "practice community" includes "school district administrators, principals, interventionists, and teachers" (p. 237), I also focus on literature pertinent to these groups.

## Problems Addressed by Research

Among other things, Farley-Ripple and colleagues (2018) postulate that research use depends on how well the questions, topic areas, and outcomes addressed by researchers align with the questions, topic areas, and outcomes of interest to practitioners. They write that "[t]he extent to which the evidence produced by the research community is timely and relevant to the problems confronting real schools is an indicator of this dimension of the gap" (Farley-Ripple et al., 2018, p. 240). Accordingly, two bodies of academic literature are reviewed here. The first concerns practitioners' attitudes about educational research and its general relevance to the problems they experience in practice. The second explores educators' specific research interests—that is, the specific

research questions, topic areas, and outcomes that they find most relevant and compelling. Together these literatures provide some idea of what research questions and topics the WWC should feature if it hopes to align itself with the practice community.

**Educators' Attitudes About Educational Research**

Perhaps unsurprisingly, researchers studying the research-practice gap have repeatedly found the perceived relevance of educational research to predict its eventual use by educators (for a review, see Schaik et al., 2018). This is true of work examining both hypothetical and actual research use. For example, Williams & Coles' (2007b) study of teachers' information literacy discovered that positive attitudes about educational research were associated with heightened confidence in using that research to inform practice. Likewise, when exploring the bidirectional nature of this association, they discovered that teachers' previous research experience—especially in participatory research efforts (as also reported by Cousins & Walker, 2002)—was associated with more positive attitudes about educational research. Subsequent work has replicated this finding (e.g., Judkins, Stacey, McCrone, & Inniss, 2014; Lysenko et al., 2014). Lysenko, Abrami, Bernard, Dagenais, and Janosz's (2014) study is particularly conclusive given its analytical strategy and the diversity of its sample. In their sample of Canadian teachers, administrators, and education professionals, they found participants' opinions about research-based information—including both its relevance and timeliness—to be the strongest explanatory factor of research use. Their regression model demonstrated this finding across dimensions of research use (i.e., instrumental, conceptual, and symbolic) and after controlling for various participant- (i.e., years of teaching experience) and school-level (e.g., size, language) characteristics. Hence, considering the attitudes

educators harbor about educational research is critical to efforts looking to improve research use.

So what, then, does the literature suggest about the nature of these attitudes? A closer look at Lysenko and colleagues' (2014, 2015) findings reveals a second, more concerning pattern about the educators' research-related attitudes. Not only did participants report engaging with research relatively infrequently (i.e., once or twice over the past year), but they expressed a general ambivalence towards educational research in the first place. In 2014, Lysenko et al.'s sample showed some variability in their attitudes. Teachers—compared to administrators and educational professionals—reported slightly more favorable endorsements of statements such as "Research-based information is relevant to your reality" and "Research offers timely information." In context, however, these attitudes were still somewhat neutral, and a subsequent study (Lysenko et al., 2015) using the same instrument found educators' ratings of research relevance and timeliness to be similarly ambivalent.

It is important to note how Lysenko and colleagues' (2014) pattern of findings, with teachers having slightly more positive attitudes about educational research than administrators, actually reverses the pattern commonly noted in the literature. Administrators typically report more positive attitudes about research than their teachers (for an early review, see Hemsley-Brown & Sharp, 2003). Biddle and Saha's work (2002, 2006) found virtually their entire sample—which featured primary- and secondary-school principals from both Australia and the United States—to hold research-based knowledge in high regard. A more recent survey by Penuel et al. (2017), as well as subsequent interview research (Penuel, Farrell, Allen, Toyama, & Coburn, 2018), also found that

administrators across levels of leadership (i.e., school and district levels) valued educational research. Even so, it is worth noting that other work (i.e., Coburn & Talbert, 2006) have documented a much greater array of attitudes—including skepticism—among education administrators, and that the nature of these attitudes may be contingent upon the level of administration in question (i.e., school-level vs. district- or state-level).

Skepticism about the relevance of educational research is far more apparent in literature examining teachers' attitudes about educational research. Reviews of the literature have spoken of this skepticism repeatedly (Hemsley-Brown & Sharp, 2003; Dagenais et al, 2012; Cain, 2016; Schaik et al., 2018; Gorard et al., 2020). Indeed, in the open-ended portion of Lysenko et al.'s survey (2015), the authors found that teachers' neutrality towards educational research were complicated by open-ended responses that "add rather a negative connotation to their perceptions" (p. 35). Nearly 40% of responses alluded to the irrelevance of research, claiming that researchers had "lost touch with the realities of the classroom" (p. 47). Elsewhere, in qualitative work by Cain (2017) and Joram, Gabriele, and Walton (2020), there are further indications that some teachers are averse to—rather than ambivalent about—educational research. For example, in Cain's interviews with 28 British schoolteachers, he characterized the common attitude held by his participants as one of "dissent" (p. 21), meaning that they exhibited an active contestation towards research findings. Similarly, Joram and colleagues (2020) interviewed teachers who had just finished a graduate-level course in educational research. Many of their interviewees voiced "discomfort" (p. 1) towards research, and more specifically, towards the idea of generalizing findings from research that had been conducted in a setting other than their own.

**Reasons for Teachers' Skepticism Towards Educational Research**

Joram et al.'s (2020) observations, as well as those featured in Cain's (2017) work, highlight the possible reasons why teachers may hold neutral and/or oppositional attitudes towards educational research. One reason, which is noted by Joram et al. (2020) as well as others (e.g., Le Fevre, 2014), is that teachers are notoriously risk adverse. This aversion to risk, especially the risk of changing classroom practices that teachers are already comfortable with, can discourage them from buying in to educational research as well as programs that are research-based. As an example, Le Fevre (2014) found that teachers' perceptions of risk—particularly risks to their own practice—influenced their implementation of a primary school literacy intervention, even when the intervention aimed for seemingly uncontroversial outcomes. The intervention, which encouraged teachers to grant students greater agency in classroom activities, was met with opposition by teachers because of their concerns over losing control over students; to teachers, a loss of control could, in turn, allow for additional classroom disruptions (Le Fevre, 2014). Relatedly, Pareja Roblin and colleagues (2018) found that teachers expressed similar concerns about using tablet PCs in their classrooms. Although previous research had documented that tablets "worked" as a teaching resource, interviews with practitioners revealed widespread practical concerns that tablets "…might significantly increase students' temptation to go off-task" (Roblin et al., 2018, p. 11). This work is notable because it reminds us that educational interventions can have unintended side effects for both teachers and students (Lykins, 2012; Pondiscio, 2019; Zhoa, 2017, 2018). It further suggests that teachers do, in fact, care about these side effects.

A second, related factor fueling teachers' skepticism about research is that, in some cases, the evidence-based recommendations afforded by research contradict the practice-based observations made by teachers. This sort of "non-congruence" was noted by Cain (2017) as the chief reason why teachers in his study contested educational research. He shared an extended example of one teacher whose experiences with gifted and talented students contradicted a piece of literature Cain (2017) asked her to read as part of his study. This was true of other participants in Cain's (2017) study as well—several of whom, though willing to accept certain research findings as credible, still doubted how useful they would be to their *own* students. A similar sentiment was found in focus groups conducted by Behrstock-Sherratt, Drill, and Miller (2011). Just as Lysenko et al's (2016) sample did, teachers in Berhstock-Sherratt et al.'s (2011) study perceived educational research to be inconsequential to the daily realities they encountered as practitioners. As explained by one of their participants:

> "A lot of times, research is done by people who don't spend time in classrooms and who don't know students. It only takes you so far. Then, you feel like, "I'm living in it. I have a better sense of what students need and what works with students than someone who is just looking at a bunch of numbers on a piece of paper." (Behrstock-Sherratt et al., 2011, p. 5)

As such, many teachers may find educational research to be less useful than their own professional judgements.

A third concern underlying teachers' skepticism towards educational research relates to the issue of generalizability/transferability. As will be explained later in the chapter, one of the criteria used by educators to vet the credibility of an educational research study is the extent to which the study shares contextual similarities to their own (i.e., ecological validity; Bracht & Glass, 1968). In a similar way, some teachers balk at

the value of educational research because they view the characteristics of their students, as well as the context in which they practice, as unique. The theme of uniqueness featured heavily in Joram and colleagues (2020) interviews with teachers. In those conversations, Joram et al. (2020) encountered teachers who firmly believed in the uniqueness of each student, and accordingly, in the inability of educational research to accommodate that uniqueness. The teachers in Cain's (2017) work, though slightly more moderate in their views, still questioned the value of educational research conducted in other settings or at other points in time. One respondent, for example, discussed how they believed that "pupils have got more and more confident over the years", and accordingly, that some of the research findings they were asked to review were outdated (Cain, 2017, p. 15). Another one of Cain's (2017) teachers spoke of the uniqueness of their students, asking rhetorically: "Does anything work for every single person? No…"

**Educators' Research Interests**

Despite there being some opposition to educational research, many educators still consult it. Unfortunately, compared to the literature examining educators' attitudes about research, markedly less has inquired into the specific research topics educators find most interesting. The extant literature again suggests that educators' research interests vary depending on their role (i.e., teacher vs. administrator); within both groups, however, there is still a great deal of heterogeneity in topic areas of interest. Beginning first with administrators, Biddle & Saha's (2002, 2006) early work with principals asked them to report on various topics they had sought research on. Various topics were listed, with research on effective schools being a common one. In more recent work with district-level administrators, Penuel and colleagues (2018) used both a survey and follow-up

interview to learn more about the types of research administrators had used over the past year. Although a large number of these administrators had mathematics-specific roles, mathematics-related research was (surprisingly) not the most common research topic reported. Instead, issues of school organization/improvement were cited as the administrators' most popular research topic, followed by research on mathematics, and then by research on equity/diversity/inclusion. What this ordering of interests suggests is that administrators, at least at the district level, may be more interested in studies examining school- and teacher-level factors rather than student-level ones.

A U.S. survey conducted by the EdTech Evidence Exchange (i.e., Barton & Tindle, 2019) presents similar findings regarding administrators' research interests while also distinguishing those interests from the interests of teachers. In that work, over one thousand educators were surveyed about their attitudes towards, usage of, and access to educational research. For reporting purposes, the survey differentiates between two groups of educators in their sample. The first is referred to as "convening attendees," and the second is referred to as "association members." Convening attendees were significantly more likely to identify as PreK-12 teachers than association members. (For additional details about these groups, see Chapter 3). When comparing between the two groups, differences in their research topics of interest emerged. Specifically, the top three research interests stated by association members including (1) pedagogical practices, (2) special education, and (3) socio-emotional support/practices. All three of these, like the topics identified by Penuel et al's (2018) administrators, are not specific to student-level outcomes. Convening attendees, however, sought research on (1) pedagogical practices, (2) student achievement, and (3) technology's effects on students—two of which are

explicitly student-focused. “Student engagement” was listed as their fourth topic of interest. Again, the nature of this finding suggests that educators’ research interests differ according to their roles.

Like the EdTech Evidence Exchange survey, the limited academic literature explicating teachers’ research interests also suggests that they are concerned about student-specific issues or outcomes. This has been known for quite some time. For instance, Everton, Galton, and Pell (2000) asked teachers to look both retrospectively (i.e., at research they had accessed in the past) as well as prospectively (i.e., at educational issues they would like to be researched). Both types of responses focused on student-specific topics, yet the nature of those topics differed. Teachers’ retrospective responses most commonly identified research related to “aspects of children’s learning,” such as aspects of Piagetian theory or Gardner’s framework of multiple intelligences. When answering prospectively, however, educators reported that research on improving student motivation and curbing student disengagement would be the most research areas moving forward. These issues were considered twice as important as any of the other prospective options in Everton et al.’s survey (2000)—including the response option “managing children’s learning performance.” This finding adds nuance to the EdTech Evidence Exchange survey’s findings by reaffirming teachers’ interest in research on student success, while, at the same time, suggesting that outcomes like “motivation” and “engagement” might matter as much as “achievement.”

Beyond issues of student success, the literature also finds teachers to be interested in research related to students’ demographics/identities and issues of diversity. In open-ended questions about research findings they had implemented into practice, practitioners

studied by Ion and Iucu (2014) frequently mentioned "aspects of multicultural and intercultural education" (p. 340). Likewise, in Barton and Tindle's (2019) survey as part of EdTech Evidence Exchange, "Special Education" and "Diversity, Inclusion, & Acceptance Concerns" were identified as core research interests for both samples. Ultimately, however, more research is needed to better understand (a) the topic areas of research that educators have chosen to use, and (b) the topic areas they wish research would address.

**Problems Addressed by the WWC**

To understand the WWC's own scope as a research clearinghouse, it seems important to reiterate that, as its namesake suggests, the WWC is fundamentally concerned with research addressing the question of "what works in education?" (WWC, n.d.-g). This "what works" agenda did not emerge from thin air; it is, itself, evidence-based. According to Grover Whitehurst, the first appointed director of IES, this focus on "what works" emerged from a stakeholder survey he conducted. Shortly after taking office as IES director, Whitehurst claims to have surveyed a sample of school administrators, state-level education officers, and legislative policymakers about how to make educational research more relevant and useful to them. His data—which, to my knowledge, was never formally published—called for a clearinghouse addressing "…questions of effectiveness. In other words, what works best, for whom, under what circumstances?" (Whitehurst, 2003, p. 6).

Dissecting the "what works" phrase helps demonstrate the various research topics and outcomes currently addressed by the WWC. The "what" refers to educational interventions. The WWC (n.d.-c) defines "intervention" as "[a]n educational program,

product, practice, or policy aimed at improving student outcomes". The "works" refers to educational outcomes, which the WWC (n.d.-c) defines as "[k]nowledge, skills, attitudes, and other desired benefits that are attained as a result of an activity." These interventions and outcomes are categorized according to various "topic areas," of which the WWC has twelve, including: literacy, mathematics, science, behavior, children and youth with disabilities, English learners, teacher excellence, charter schools, early childhood, kindergarten through twelfth grade, path to graduation, and postsecondary. Together these topic areas represent various academic subjects, student types, and levels of education. Finally, it's worth noting that these topic areas emerge from the WWC's "review protocols". Review protocols are formal documents "…developed to determine the scope of a review [of research]. Protocols include key definitions, express the types of research studies and interventions reviewed, set reviewed parameters, and explain methodology" (WWC, n.d.-c).

Together this information contains several insights about how the WWC's scope aligns with what is known about educators' research interests. The first is regarding its "what works" focus. As shown above, questions of effectiveness and efficacy are of interest to educators, but so are others. For example, given the literature documenting teachers' risk aversion, it seems that the question of "what hurts?"—that is, what interventions may actually harm student outcomes—could be of interest as well. However, because the WWC equates "outcomes" with an intervention's "desired benefits," it explicitly forsakes the reality of these undesirable harms. A second insight can be gleaned from how the WWC's review protocols, and in turn, its topic areas, are decided upon. The WWC identifies topics for review protocols based on the following:

> "…nominations received from the public to the WWC Help Desk; input from meetings and presentations sponsored by the WWC; suggestions presented to IES or the WWC by education associations; input from state and federal policymakers; patterns of searches for education topics on the WWC website or on the Internet more generally; and scans of the literature or of research funded by the U.S. Department of Education" (WWC, 2020d, p. A-2).

This suggests that, while there are avenues through which educators can nominate topic areas of interest (i.e., the WWC Help Desk), there are numerous other avenues through which other stakeholders determine the WWC's foci.

Such an observation is especially notable given how the WWC's current areas align with educators' own. Indeed, some do—such as a topic area related to Special Education, which was noted as a topic of interest in Barton and Tindle's (2019) survey. At the same time, however, there is evidence of misalignment. For example, the WWC's dedicated topic areas of math, English, and science seem to overemphasize specific academic subjects, which are not often named in the previous literature on educators' research interests. In addition, some desired topic areas—like diversity and/or multicultural education—are either underrepresented or missing altogether from the WWC. Thus, based on the available literature, there are some ways in which the WWC's topics and outcomes of interest aligns with educators' own, but there are also ways in which they do not.

### Usefulness of Research Products

A second dimension featured in Farley-Ripple et al.'s (2018) framework refers to the usefulness of research dissemination methods. Specifically, the extent to which "the products produced and valued by researchers aligns with those preferred by practitioners" (p. 240) is hypothesized as a contributor to subsequent research use. In the case of research brokers, like the WWC, findings from research articles and reports are usually

translated into more practitioner-friendly resources. This means that, when a research broker is involved, the research is generally disseminated using a product produced by an intermediary rather than the researcher themselves. Still, it's worth reviewing how empirical literature suggests disseminating research to practitioners. I review three areas of literature below. One area examines educators' preferred research products or "access points"; another identifies characteristics that make those access points effective. Before reviewing those areas, however, I first discuss what is known about how educators use—or, at least, intend to use—the research featured in those products. Doing so is important because of how different types of use demand different types of products (e.g., Levin, 2011). When considered together, these literatures help inform inferences made in Chapter 4 regarding the usefulness of the WWC's own research products.

**The Nature of Educators' Research Use**

Theorists of research use have long acknowledged the various ways in which use can and does occur. The work of Carol Weiss (1979, 1998, 2008) was some of the first to distinguish the myriad ways that social science can be used outside of the academy. These types of research use can broadly be categorized as instrumental, conceptual, symbolic, and imposed. *Instrumental use* refers to situations where research directly informs action taken to solve a target problem. *Conceptual use* refers to situations where research indirectly influences how people think about or engage with the world. *Symbolic use* refers to situations where research is used as a backdrop or guise for an unrelated (often political) aim. *Imposed use* refers to situations where research evidence is used in order to comply with certain policy demands or funding stipulations. Although theorizing about the research-practice gap has, admittedly, focused on instrumental use (as noted by

Farley-Ripple et al., 2018), Weiss (1998) and others (e.g., Cain, 2015) have noted how other types of research use are both important and common.

Indeed, much of the extant literature on educators' research use either implicitly fits with or explicitly aligns itself with Weiss's categories. Some of this work has inquired into the intentions with which educators engage research. For instance, in Coburn and Talbert's (2006) case study of a large, urban school district in the United States, researchers asked administrators about how they believed research on student learning should be used. Interviewees differentiated between their desires to use research instrumentally (e.g., to inform student placement decisions and classroom instruction) and the more common reality of symbolic and/or imposed uses. One administrator explained how his school often consulted research *after* decisions had been made in order to justify how their Title 1 funds were spent. Another theme identified by Coburn and Talbert (2006) was the extent to which research use was facilitated by accountability demands—an example of imposed use. Thus, even in instances where educators aspire to instrumental research use, the use that occurs in reality may not fully meet these aspirations.

This reality is substantiated by additional literature exploring the ways educators use research in practice. Reviews of this literature conclude that non-instrumental types of research use are just as common—if not more so—than instrumental research use (for reviews, see Dagenais et al., 2012; Cain, 2016). Yet, as before, this may again differ based on the type of educator examined. For example, Penuel and colleagues (2017) survey of school-, district-, and state-level administrators found instrumental uses of research to occur with the greatest amount of frequency, though they were closely

followed by instances of imposed use. Subsequent work focusing on three U.S. school districts (Penuel et al., 2018) found instances of imposed use (e.g., "to fulfill mandates to use research") to be reported with the greatest frequency by all three; however, instances of instrumental use (e.g., "to select standards to give greater focus and attention" or "to improve existing programs") were reported with the second highest frequency by two districts. Taken together, these findings indicate that administrators seek out research products to help them make instrumental decisions, but they do so for compliance reasons as well (e.g., Gorard et al., 2020).

On the other hand, studies of teachers' research use find their use to be largely conceptual. Survey work by Lysenko et al. (2015) found that their sample—82% of which identified as teachers—seldom consulted research-based information. Those who did reported doing so for conceptual reasons more than instrumental or symbolic reasons. Interviews conducted by Cain (2015) furthers this observation when noting that teachers' instrumental and symbolic uses occurred with much less frequency than their conceptual uses. As explained by Cain (2015):

> "The distinguishing feature of this [conceptual] category was the teachers' intellectual engagement with the [research] articles: Rather than using them to explain or justify actions, they used the research texts to think about their experience and practice, individually and in discussion." (p. 13)

As a result, Cain theorized that research primarily functioned as a "third voice" in conversations between teachers and their colleagues—meaning that is served a reflective rather than a decision-making function.

Regardless of the type of use examined, administrators tend to use research with greater regularity than teachers. This has been recognized since Hemsley-Brown and Sharp's (2003) early review of the literature, which concluded that classroom teachers

tended to be irregular users of research while administrators actually appeared to be "regular, thoughtful users of research knowledge" (Hemsley-Brown & Sharp, 2003, p. 9). Hemsley-Brown and Sharp's observation is echoed in more recent work. For example, direct comparisons of teachers' and administrators' research use, such as those reported by Lysenko et al. (2014), have found administrators to report more regular usage, whereas teachers' usage was deemed "sporadic" (p. 14). One reason for this discrepancy may be the reality of imposed research use among administrators (Penuel et al., 2017, 2018)—a reality that Gorard and colleagues (2020) still find to be "the most effective way to get evidence-into-use" (p. 28).

**Educators' Research Access Points**

Teachers' propensity for conceptual research use, as well as administrators' experiences with instrumental and imposed use, are likely to influence the types of research products they choose to consult. Because some dissemination approaches are not so much products as they are events (e.g., professional conferences), I believe the term "access points" (i.e., Barton & Tindle, 2019) is most appropriate here. Understandably, and as acknowledged in Dagenais and colleagues' (2012) review, the literature on educators' research access points is unwieldy precisely because new types of access points continue to emerge. For example, early work by Everton and colleagues (2000) found that teachers continued to rely on print sources—such as books, journals, and other professional publications—to access educational research. Nevertheless, as the Internet has grown more ubiquitous, so too have studies finding that teachers most commonly access research through electronic sources (Williams & Cole, 2007; Dagenais et al., 2012); evidence suggests that research brokers are relying more heavily on web-based

dissemination strategies as well (e.g., Cooper, Edelstein, Levin, & Leung, 2010). Accordingly, I prioritize findings from more recent literature here.

One finding that has held consistent over time is the shear diversity of research access points used by educators. Lysenko et al.'s (2015) survey of Canadian educators found that five out of ten research access points inquired about in their survey—including the Internet, scholarly documents, in-service trainings, professional publications, and conferences—were all selected by roughly 80% of respondents as an access point to research over the past year. Likewise, subsequent work by Penuel and colleagues (2017) found that administrators accessed research "through a variety of sources" (p. 8), though they mainly relied on aspects of their professional network such as colleagues, conferences, and their professional associations. Also similar to the findings of Lysenko et al. (2015) are those emerging from the EdTech Evidence Exchange survey (Barton & Tindle, 2019), which found that over 90% of respondents had accessed research from blogs, journal articles, colleagues, and professional conferences over the past year. In sum, this work demonstrates that educators rely on a variety of mediums when accessing educational research.

It would be shortsighted, however, to ignore some of the nuances embedded within this variety. Two are worth addressing here. The first concerns differences between the preferred access points of teachers and administrators. Although both report using a variety of access points, which do they use the most? An earlier survey by Lysenko and colleagues (2014) found "significant variations in groups' self-reports about their use of RBI [research-based information], implying a divide between school administrators, on the one hand, and teachers and professionals, on the other" (p. 11).

Specifically, they found that administrators in their sample consulted traditional dissemination mediums—such as scholarly documents and professional publications—with greater frequency than teachers. In contrast, teachers reported accessing research through the Internet and multimedia more often than administrators did. In an almost identical fashion, comparative analyses from the EdTech Evidence Exchange survey indicated that their subgroups differed as well. Although comparative information about their frequency of access was unavailable, results indicated that "association members" were almost twice as likely to report accessing research through journal subscriptions. Convening attendees, on the other hand, were more likely to have accessed research via social media (though this difference was not statistically significant; Barton & Tindle, 2019). Therefore, there is some evidence to reinforce a pattern of findings already emerging in this literature review—namely, that teachers and administrators access educational research in different ways.

A second nuance concerns the ubiquity of access points that involve interpersonal relationships—a pattern found in the research-seeking behavior of teachers and administrators alike. Despite the diversity of research access points reported by respondents, Lysenko and colleagues (2014) concluded that their findings speak to "the importance given to peer exchanges of practice-relevant information, with the open-ended reports implying that colleagues (both near and far) are a primary source of information" (p. 49). The importance of collegial networks was noted in their earlier work (Lysenko et al., 2014) as well as the work of others (e.g., Drill, Miller, & Behrstock-Sherratt, 2012). Consider, also, the findings of Penuel and colleagues (2017), which speak to the importance of interpersonal relationships as research access points for

administrators. In that work, over half of participants indicated that they accessed research "often" or "all of the time" via their professional connections. Finally, in recent work by Barton and Tindle (2019), results suggest that colleagues are an access point that educators rely on repeatedly. 91% of respondents report using their colleagues as an access point at least once a year. Of those 91%, 19% reported doing so once or twice, whereas 33% reported doing so seven or more times. The appeal of colleagues of research access points is well-articulated by a teacher in Drill et al.'s (2012) focus groups, who confessed that:

> "It's a lot easier to walk down the hall and ask my colleagues who've had similar experiences or more experience than I do. I just sat down with a colleague the other day, to go over history knowledge that I could've looked up in a journal or a book, but I'd rather go to someone who I know has taught it and can explain it to me in a way that would help my students the best. And it's efficient, because I already see them" (p. 6).

**Characteristics of Favored Access Points**

Other literature has focused its attention on characteristics of research access points that may account for their popularity. Because these characteristics are important determinants of eventual research use (for a review, see Dagenais et al., 2012), researchers have striven to identify what they are and how they work. One characteristic implicated above is the access point's applicability to practice, which is especially important for teachers given their general skepticism about research's relevance to the problems they encounter in the classroom (e.g., Schaik et al., 2018). As implied by Hammersley (2001b), Biesta (2007, 2010), and others (e.g., Cain, 2016), to be applicable to practice, research access points must act as a translator—specifically, to convert the "what is" scope of most education research to the "what to do" concerns of practitioners. Conveniently, much of the work examining educators preferred research access points

also examines the practices that make those access points most alluring; their applicability to practice was a reoccurring theme. For example, in Lysenko et al.'s (2015) survey, two of the three highest rated activities for making teachers aware of educational research included (a) accompanying research results with recommendations, and (b) demonstrations of how to *apply* those recommendations. When an adapted version of these same survey items was administered to educators in the EdTech Evidence Exchange (Barton & Tindle, 2019) survey, the same pattern of results emerged: respondents indicated that "[r]esearch results accompanied with clear, explicit directions" (p. 17) would be the most useful way of informing them about educational research. Although these findings should not be conflated with mandating teachers to act in certain ways (which can actually discourage research use; Joram et al., 2020), classroom teachers are more receptive to research-based recommendations than to raw research findings. Teachers' colleagues may thus be a preferred access point because they are better able to offer practice-oriented suggestions than researchers themselves.

A second, albeit somewhat more rudimentary characteristic of effective research products is that they offer approachable, non-technical summaries of research (see Gorard et al., 2020 for a discussion). Williams and Coles' (2003, 2007a, 2007b) studies of teachers' information literacy documented that teachers are, in general, lacking in confidence when it comes to evaluating and using educational research. Participants voiced a need for "predigested" (Williams & Coles, 2007a, p. 812) rather than raw research. In fact, one of their respondents asked for an initiative that would, in effect, function like the WWC:

> "Part of the solution would be the establishment of an Internet site devoted to giving summary statements concerning the various strands of research currently being undertaken." (p. 201)

Subsequent work has suggested similarly. The work of Vanderlinde and van Braak (2010) is particularly compelling because it included members of both the practice and research communities. Specifically, in their focus groups with Dutch teachers, administrators, educational researchers, and intermediaries (e.g., editors of practitioner journals), Vanderlinke and van Braak (2010) identified the "technical and complex language usage" (p. 307) of most academic writing to be a barrier to educators' research use. In turn, their conversations with researchers revealed that, while they recognized this problem, they still found it easier to produce technical documents than plain-language summaries for practitioners. Intermediaries, then, saw it as their job to develop "teasers" of research information for practitioners. This may be why, in their review of the research, Gorard et al. (2020) concluded that "[p]roviding access to raw research evidence or even slightly simplified evidence is not generally an effective way of getting it used" (p. 1).

A third characteristic that helps ensure the use of research access points is their timeliness—that is, the extent to which they feature up-to-date information about current educational issues. Timeliness is important because, as found by Behrstock-Sherratt, Drill, and Miller (2012), teachers often consult research access points when they encounter pressing practice-related concerns, such as how to best teach an unfamiliar content area or how to effectively teach students from certain demographic groups. The importance of timeliness has been noted in previous reviews (e.g., Dagenais et al., 2012) as well as in recent large-scale surveys (Penuel et al., 2017; 2018). For example, although

educators in Penuel and colleagues (2017) largely agreed that educational research was relevant, more than half also endorsed the statement "[b]y the time research is published it is no longer useful to me" (p. 35). Given how lengthy the research publication process can be, it is of little surprise that many educators chose to "walk down the hall" and consult a colleague instead (Drill et al., 2012, p. 6).

**Products Produced by the WWC**

Since its conception, the WWC has not only sought to be an evaluator of educational research, but also a disseminator. In fact, Whitehurst (2004) went as far as to claim that the WWC's "…sole purpose is to deliver solid research into the hands of educators, policymakers, and the public" (p. 15). To what extent, then, does the WWC disseminate research in ways that are known to facilitate use? The WWC's dissemination strategy can be considered as having two primary components: (a) its website, and (b) its written reports (which are typically access through its website). The core of its website involves the "Find What Works" database—an interactive system that "allows users to identify programs, policies, and practices that have been shown to improve student outcomes" by using filters and keyword searches (IES, 2018c). The website has been revamped in recent years to improve usability (WWC, n.d.-f). Once users find a certain topic or intervention of interest, three types of written reports are typically used to convey related information. Those reports include (a) Reviews of Individual Studies, (b) Intervention Reports, and (c) Practice Guides. *Reviews of Individual Studies* are the most technical of the three documents. They focus on communicating the effect size and statistical significance level of a single study or evaluation. *Intervention Reports* summarize findings across multiple studies to make general claims about whether an

intervention “works”, and in some cases, for whom an intervention “works” for. These reports exemplify the WWC’s function as a systematic review initiative. Intervention Reports are often coupled with short, one-page *Intervention Snapshots* and slightly more comprehensive *Intervention Briefs*, which are usually three to four pages. Finally, findings from these Intervention Reports are sometimes translated into actionable recommendations for practitioners through *Practice Guides*. These products, which are freely available, are the WWC culminating attempt to bring research into the hands of educators.

At first glance, the WWC’s products share some of the same characteristics as those preferred by educators. For example, its Intervention Reports offer practitioners predigested summaries of research that are, at least by intention, easier to understand than the research would be in its raw form. Practice Guides do one better. They offer recommendations for how a teacher might implement a research-based practice in their classroom, and literature suggests that this is more appealing to teachers than raw or summarized research information. Even so, these products still do not account for the interpersonal nature of most research-sharing that occurs among educators. Teachers talk to one another; administrators converse at conferences. The WWC’s written reports rely on a unidirectional flow of information, not an exchange of ideas that seems to encourage research use (see below). Furthermore, a danger with written products is that their conclusions grow stale as new research is produced. There is thus some reason to believe that educators’ desire for timely information is incompatible with the WWC’s predominant dissemination approach.

## Nature/Quality of Research

Another dimension in Farley-Ripple et al.'s (2018) model distinguishes between indicators of research quality held by the research community and those prioritized by the practice community. It suggests that "the extent to which researcher standards [of quality] and practitioner preferences are similar or different" will determine practitioners' use of educational research (p. 240). So, do practitioners privilege the same types of research designs, methodologies, and analytical approaches as researchers? If not, what other criteria do they use when judging the quality of an educational research study? Literature reviewed in this section offers insight into how the WWC's standards of research quality both converge with and diverge from educators' own.

### Credibility of Educational Research

Before introducing educators' standards of research quality, it is important to first review what is known about the perceived quality of educational research among educators. This is a complementary, yet distinct literature to the one examining educators' skepticism of research's relevance (see above). Indeed, educators can be skeptical of a study's relevance while still perceiving its findings to be empirically sound. Early reviews of the literature (i.e., Hemsley-Brown & Sharp, 2003) has differentiated between the trust in research expressed by administrators and the trust felt by teachers—with administrators generally holding research in greater esteem than their teachers. To an extent, this pattern has held consistent in subsequent work (e.g., Coburn & Talbert, 2006). Even so, more recent work tends to suggest that many administrators and teachers are, at best, lukewarm to the credibility of research-based evidence. With regard to administrators, Lysenko et al.'s survey (2014) found endorsement for the statement

"Research is reliability and trustworthy", though ratings indicated only modest agreement (i.e., a mean rating of 3.37 on a five-point Likert scale). In a similar way, administrators in Penuel and colleagues (2017) survey did perceive educational research as credible, though the authors were quick to note that the finding was not without caveats. As an example, unlike ratings of research's value and relevance, administrators' ratings of credibility were not associated with greater instrumental research use. Furthermore, despite this rating of credibility, Penuel et al.'s (2017) sample was hesitant to endorse the statement that "[e]ducational researchers are unbiased," with only 51% agreeing or strongly agreeing. This suggests that, while administrators may trust research findings, they still recognize them as fallible.

Previous research reviews have concluded that many teachers also question the credibility of educational research—perhaps to a greater extent than administrators (i.e., Hemsley-Brown & Sharp, 2003; Gorard et al., 2020). Teachers' hesitance is especially clear in qualitative and mixed-methods work. For example, in focus groups conducted by Boardman, Argüelles, Vaughn, Hughes, and Klingner (2005), special education teachers voiced a great deal of mistrust towards educational research. Like administrators, they hinted at the notion that even researchers have their own agenda, with one respondent going as far as to ask:

> "How much credence do you lend to research and how much credence do you lend to the numbers that can be manipulated any way you want to manipulate them?" (Boardman et al., 2005, p. 176)

Other work has found that teachers implicate educational research's "lack of rigour" (Lysenko et al., 2015, p. 47) as a reason for their skepticism. Indeed, not only did teachers in Lysenko and colleagues (2015) survey hesitate to endorse the "credibility and

trustworthiness" of educational research, but in their open-ended responses, they also criticized its methodological quality. Given the perceptions of both teachers and administrators, it appears that, in general, educators do not trust the quality of educational research as much as researchers would like.

**Epistemological Differences Between Educators and Researchers**

Why this lack of trust? One reason that is implicit, yet central to the Two-Communities theoretical tradition is that the practice and research communities possess differing—and in some cases, conflicting—epistemological viewpoints. Put another way, researchers and practitioners differ in how they view and interpret knowledge, especially knowledge arising from scientific inquiry. This is, in fact, exactly what Farley-Ripple and colleagues (2018) use as justification for the "nature/quality of research" dimension of their model, writing that "the two communities value different qualities of research, including issues related to internal and external validity as well as conclusiveness of findings" (p. 240).

Indeed, investigations by Joram (2007), Borg (2008, 2009, 2012), and Bråten and Ferguson (2015) have studied these epistemological differences in detail. Joram's (2007) work is especially notable, as it used a directed interview protocol to compare how preservice teachers, practice teachers, and teacher education professors believed knowledge should inform teaching. Differing perspectives across groups were found in several areas. When asked to consider how they would determine whether or not a teaching approach "worked," a majority of preservice teachers discussed a kind of "trial and error" process where the approach would be attempted, and based on the reactions of students, the teacher would continue or modify the approach. Not a single professor

alluded to such a process. Instead, nearly all of them said they would use an experimental design where two groups of students who be taught differently and their reactions would, in turn, be compared. Practicing teachers alluded to both approaches. Joram's (2007) work thus speaks to differences in how educators and researchers prefer practice-related knowledge to be generated in the first place; other research, such as Bråten and Ferguson's (2015) finding that teachers preferred experiential knowledge sources to theoretical ones, reinforces this difference as well.

A second epistemological difference uncovered by Joram (2007) relates to how individual student differences were perceived by teachers compared to professors. Specifically, the idea that "each student learns differently" was noted by half of the preservice teachers interviewed by Joram (2007); conversely, only two of the seven professors interviewed endorsed this idea and/or acknowledged how individual student differences might make the implementation of research-based practices challenging. This may also be why nearly half of the professors in Joram's (2007) study believed that research could identify universal educational principles, whereas only 7% of preservice teachers and 11% of practicing teachers believed this to be true. These findings align with literature discussed above (i.e., Cain, 2017; Joram et al., 2020) suggesting that teachers doubt the relevance of educational research because of the perceived idiosyncrasies of their students.

A third observation made by Joram (2007) logically emerged from this second one. Because students are unique, teachers believed that the utility of research-based knowledge was confined to the setting it was conducted in; it cannot and should not be generalized to other contexts. Practicing teachers were actually more extreme in this

belief than preservice teachers, with 78% (compared to 43% of preservice teachers) mentioning this context-specificity. On the other hand, only 14% of professors mentioned the non-transferability of research findings. Later work by Joram et al. (2020) identified the generalizability/transferability of findings as a theme when asking teachers about their perceived barriers to research use. Cain's (2017) interviews uncovered a similar theme, leading him to conclude that "teachers suggested that findings from research could not be generalised to all pupils or all academic subjects" (p. 16). As will be shown below, issues of external validity are heavily prioritized in practitioners' own standards of research quality.

One clarification must be made before introducing the standards educators use when vetting the quality of research. It is important to recognize that, although educators and researchers view the nature and sources of educational knowledge differently, educators' conceptualization of research actually converges with the traditional conceptualizations of researchers. This has been repeatedly shown in Borg's (2008, 2009) work with English teachers around the world, where they were presented with ten scenarios describing types of teaching-related inquiry. In a series of surveys and follow-up interviews, Borg (2008, 2009) found that teachers' conceptions of research were aligned with "conventional scientific ideas" (2009, p. 367). In both studies, the scenario referencing quantitative data and statistics was most likely to be rated as "definitely research" by respondents. The idea of research being necessarily quantitative has been found in other work with educators (e.g., Finnigan, Daly, & Che, 2013), though some literature refutes this as well (Kennedy, 1999; Niaz, 2009). Another feature identified by Borg was that educators may have viewed this scenario as "research" because it

referenced publication in an academic journal. Hence, there is evidence that practitioners, though hesitant about the relevance and quality of educational research, still perceive the act of research similarly to many researchers.

**Standards Used by Educators to Determine Research Quality**

Both educators' epistemological stances, as well as their conceptions of research, inform how they assess research quality. Beginning with the latter, Borg's (2008, 2009) work draws explicit connections between his participants' conceptualizations of educational research and the characteristics they emphasize as indicators of high-quality research. Three of the five characteristics teachers reported as most important when assessing research quality were specific research practices that, in their eyes, beget quality. The first characteristic is the testing of a hypothesis, which 80% of respondents rating it as "important" or "very important" to their perceptions of research (Borg, 2009). In follow-up interviews, one of Borg's (2009) teachers clarified the following:

> "Well you don't want to make your research and results fit your hypothesis just to make it look as though you have an excellent result and it is a neat and tidy piece of research." (p. 368)

This comment is reminiscent to findings in both Boardman et al.'s (2005) and Penuel et al.'s (2017) work—both of which found that educators worried about how findings can be massaged by researchers to support whatever conclusion(s) they want. Accordingly, it appears that some educators find the a priori establishment of hypotheses to underlie credible research.

A second characteristic endorsed by a majority of Borg's (2009) sample was the control of variables. Again, this response seems to be unique to Borg's work, but it segues well into findings of a similar nature in work by Cain (2016). Cain (2016)

interviewed teachers after asking them to review three journal articles related to the effective teaching of gifted and talented students. He found that teachers were savvy in their critiques of those articles, and in some cases, their critiques mirrored his own scholarly concerns. For instance, one teacher identified—albeit with less technical language—how one of the studies made unwarranted conclusions about "academic gains" without assessing, or controlling for, students' academic abilities in a pretest. Another suggested that the measurement strategy used may not have been appropriate given the students' ability level. That teacher was, in effect, concerned about a possible ceiling effect. Taken together, this work shows that some teachers are able to detect methodological issues in educational research when they are given the time to do so. It also suggests that educators may use some of the very same criteria as researchers when evaluating a study's quality.

A third characteristic of quality endorsed by 72% of teachers in one study (Joram, 2007) and 67% in another (Borg, 2009) concerned study sample size. According to participants, sample size was one of the factors they used when deciding what scenarios constituted "research" at an earlier part of the study (as mentioned above). When asked why they rated one scenario higher than the others, a Turkish teacher commented "The sample—500 people…what more could you want!" (Borg, 2009, p. 366). This finding affirmed what had, at that time, already been noted by Ratcliffe and colleagues (2005), who explored science educators' perceptions of evidence-based practices. Similar to the experiences of Cain (2016), Ratcliffe et al's (2005) sample also called into question the credibility of research evidence based on a number of methodological factors, but in that study, sample size was the issue raised most frequently. Specifically, comments were

made about how small samples make results hard to generalize to "other schools" and/or they lead to erroneous findings because the sample might be "hand picked" (p. 14). These observations further demonstrate how some individuals within the practice community may hold standards of research that do, in fact, align with those in the research community.

Nevertheless, other work suggests—as Farley-Ripple and colleagues (2018) do—that "school-based decision-makers often prefer evidence from organizations or contexts (e.g., demographics, location, performance) similar to their own, regardless of study design" (p. 240). A comprehensive mixed-methods investigation by Finnigan and colleagues (2013) found support for this claim, and moreover, for educators' preference for "local data" (p. 490). This makes sense given the dominant epistemological viewpoint of educators, which, as vocalized by one of Finnigan et al.'s (2013) interviewees, means that "what works here may not work somewhere else and vice versa" (p. 483). This may also be why, in their survey responses, participants ranked research findings from local school/district evaluations five times more credible than research findings in "webbased [sic] clearinghouses or listservs" (Finnigan et al., 2013). Consider, also, the findings from Neal, Neal, Lawlor, Mills, and McAlindon (2018), which suggest that a critical facilitator of research use was the perceived compatibility between an educator's own context and the context in which the study was conducted. In fact, this compatibility was mentioned almost twice as often as any other theme in Neal and colleagues' (2018) interviews. As articulated by one educator in their study:

> "Well…when I look at research, I like to see if the demographics is comparable to my demographics, so I can compare it because every community is different. So, I want to see if there's some correlation between communities." (Neal et al., 2018, p. 11)

Nearly the exact same sentiment was expressed by educators in the electronic focus groups conducted by Joram and colleagues (2020). When compared to the value educators place on a study's internal validity (i.e., its sample size, its control of confounding variables, etc.), it appears that educators place at least equal—if not greater value—on external validity when assessing the quality of educational research.

**Standards Used by the WWC**

Although the WWC calls itself a "clearinghouse" for educational research, it is more than that. Whereas a "clearinghouse" evokes the idea of a repository or collection, the WWC is, in actuality, a curated selection of research—meaning that it not only collates research, but it assesses it as well. The WWC justifies the need to assess research quality because "not all education research is equal" (WWC, n.d.-g). Though somewhat implicit, the WWC is alluding to the belief that educational research is of notoriously low quality. This belief is frequently traced back to Kaestle (1993), though it was a motivator for Grover Whitehurst as well (e.g., Viadero, 2001).

In order to ensure that only "trustworthy research" (WWC, n.d.-g) makes its way to practitioners, the WWC has tasked itself with differentiating the good from the bad. It does so by reviewing studies according to a predetermined set of standards—the intricacies of which are spelled out in a *Standards Handbook*. Rather than recount those standards here, it seems more important to detail the assumptions on which they are based, and in turn, to compare those assumptions to those held by educators. One of these assumptions involves the privileging of internal validity, and moreover, the belief that standards informed by internal validity are the best way to ensure that only high-quality research is synthesized and disseminated. This privileging is clearly communicated in the

latest iteration of the WWC's *Procedures Handbook* (WWC, 2020d), which specifies that:

> "The WWC standards focus on the causal validity within the study sample—that is, internal validity—rather than the extent to which the findings might be replicated in other settings—that is, external validity." (p. 1)

These standards include, among other things, a focus on a study's sample size, sampling procedures, and control of extraneous variables. Indeed, the literature suggests that some educators value these standards when vetting research quality themselves, but other work suggests that standards related to a study's generalizability—like the diversity of its sample or setting—are of greater concern to educators. Therefore, there may be incongruence between educators' indicators of research quality and the WWC's own.

Among the standards meant to ensure internal validity is one related to study design. To only allow "well-designed studies" to factor into its conclusions (WWC, n.d.-g), the WWC defines "design" in the following way:

> "The method by which intervention and comparison groups are assigned (group design and regression discontinuity design) or the method by which an outcome measure is assessed repeatedly within and across different phases that are defined by the presence or absence of an intervention (single-case design). Designs eligible for WWC review are randomized controlled trials, quasi-experimental designs, regression discontinuity designs, and single-case designs. (WWC, n.d.-c)

Randomized control trials (RCTs) are viewed with special admiration by the WWC because, as was originally asserted by Whitehurst (2003), they are "the only sure method for determining the effectiveness of education programs and practices" (p. 6). It is unclear if educators believe similarly.

In addition, the WWC's standards forsake qualitative information, meaning that only quantitative findings are considered. This may be because its approach to research synthesis cannot accommodate qualitative data, as there are no significance tests or effect

sizes to compare across studies. WWC documents have clearly stated that "The WWC does NOT rate qualitative studies" (IES, 2018b, p. 5), or, as noted elsewhere, that:

> "Qualitative studies with comparison groups (for example, those comparing DDDM implementation across several districts) will be summarized but will not be subject to a formal WWC review" (IES, n.d-c., p. 1).

Therefore, as this second excerpt implies, even in cases when qualitative information is available in an eligible research design, it is still omitted from formal review (e.g., Woodworth, 2008). This may appeal to some educators (as suggested in Cain, 2016), but perhaps others may find qualitative data to be more trustworthy, more compelling, or simply more informative given the research questions they desire answers to.

### Relationships Between Communities

According to this dimension, both the nature and extent of interactions between the research and practice communities affect the uptake of educational research (Farley-Ripple et al., 2018). The nature of these relationships is complex, though they can loosely be classified as producer-pushed, user-pulled, and exchange-based (e.g., Lavis, Robertson, Woodside, McLeod, & Abelson, 2003; Levin, 2011). Farley-Ripple et al. (2018) acknowledge that indirect relationships also exist between researchers and practitioners, whereby a "researcher broker" or "intermediary" (for a discussion, see Neal, Neal, Kornbluh, Mills, & Lawler, 2015) helps bridge the two. In this section of the review, I describe these various classifications while also introducing research on the effectiveness of each relationship type at facilitating research use. Knowing this will be helpful when evaluating the WWC in Chapter 4.

**Unidirectional Disseminative Relationships**

As described by Lavis, Robertson, Woodside, McLeod and Abelson (2003), one-way knowledge transfer processes are the traditional way that research is translated between practitioners and researchers—not only in education, but in healthcare policy and practice as well. Commonly this is thought of as a "producer-push" process, whereby knowledge producers attempt to "push" their work to knowledge users through seemingly users-friendly means. Levin (2011) discusses how, these days, that typically involves Internet-based methods, such as clearinghouses of research and/or blogs discussing research in non-technical language. Intuitively, this can be conceptualized as a unidirectional transfer process that moves from knowledge producers to knowledge users. By definition, however, this approach also necessitates what has been called a "user-pull" process (Levin, 2011; Tseng, 2012), whereby knowledge users actively seek out research information to help guide their practice. This can also be conceptualized as a unidirectional process that moves from the knowledge user towards the knowledge producer.

This producer-push / user-pull process remains the default strategy of knowledge dissemination in the educational research arena. The ubiquity of this strategy has been noted off-hand in commentaries about the research-practice gap (e.g., Nutley, Walter, Davies, 2009; Cooper, Levin, & Campbell, 2009; Tseng, 2012) as well as in empirical work (e.g., Cooper, 2012). Despite its ubiquity, the effectiveness of the producer-push / user-pull strategy has also been questioned by scholars (e.g., Levin, 2011; Anwaruddin, 2015; Gorard et al., 2020). There are countless reasons why a strategy relying on effective dissemination by researchers as well as active searching by educators is likely to

fail. On the "push" side of the equation, literature has identified both individualized and institutionalized reasons why researchers may struggle to disseminate their work effectively. One reason resides in researchers' concerns about their own abilities to draft plain-language reports of their research; in fact, researchers reported in interviews with Vanderlinke and van Braak (2010) that they were more comfortable writing in field-specific jargon. In other places, such as in Ball's (2012) presidential address to the American Educational Research Association, the reluctance of researchers to employ "personal voice" when sharing research may contribute to its perceived lack of relevance to practitioners. Likewise, as is also noted by Ball (2012), there are institutionalized reasons why the producer-push strategy is shortsighted. Indeed, the academic community does not incentivize public dissemination nearly as highly as it does scholarly dissemination, which researchers recognize as a barrier to "pushing" practice-minded research products to educators (e.g., instructional materials; Burkhardt & Schoenfeld, 2003). There is also reason to believe that research methods that may be especially compelling to educators are the same methods that the dominant voices in educational research have been reluctant to embrace (e.g., case studies; Stake, 2005). Given this, as well as the aforementioned literature demonstrating teachers' own reluctance to embrace educational research, there is ample evidence to problematize the usefulness of a producer-push dissemination strategy.

Even when teachers do, in fact, decide to consult educational research, there are individual- and institutional-level reasons why they may struggle to "pull" that work into practice (for a discussion, see Anwaruddin, 2015). One notable individual-level barrier—namely, educators' information literacy—was initially investigated by Williams and

Coles (2007a, 2007b) and subsequently explored by others (for a review, see Henderson & Corry, 2020). This work implicates the user-pull strategy by showing that, though teachers are most hesitant about their abilities to evaluate research findings and communicate those findings to colleagues, they also lack confidence in their abilities to seek out research in the first place. Earlier work by Cousins and Walker (2000) found that teachers' self-perceived ability to consume educational research was associated with their beliefs about its usefulness; notably, of the five dependent variables in that analysis, teachers ranked their research consumption abilities lowest.

Although teacher educational programs and in-service trainings are paying greater attention to teachers' data literacy (e.g., Henderson & Corry, 2020), other school- and district-level factors function as hurdles to the user-pull process. One factor, which will be described in greater detail below, is that schools and districts seldom allocate time for teachers to seek out research. As such, reviews of literature continue to cite teachers' lack of time as an institutional factor impeding research use (Hemsley-Brown & Sharp, 2003; Dagenais et al., 2012; Cain, 2016). Thus, in its most fundamental sense, the producer-push / user-pull dissemination approach does little to connect the research and practice "islands"; if anything, it allows both to exist in isolation, which, according to the Two-Communities perspective, will fail to bring about greater research use.

**Bidirectional Disseminative Relationships**

What, then, may be a better alternative to these unidirectional relationships? Lavis and colleagues (2003) allude to earlier work suggesting that the hallmark of effective knowledge transfer is interaction, and more specifically, "…interaction between the clinician and an 'expert' who has been trained in the principles of academic detailing" (p.

226). Notably, Lavis et al. (2003) were hardly the first to recognize the transfer potential of research-practitioner interactions. Early educational theorists, such as John Dewey and Paulo Freire, maintained that the only way to bridge the educational research-practice divide was to invite practitioners across the bridge (Dewey, 1929; Freire, 1968). Subsequently, social scientists have advanced methodologies (e.g., research-practice partnerships; Coburn & Penuel, 2016) and approaches (participatory evaluation; Cousins & Earl, 1992) premised on the power of research-practitioner interaction. Though these approaches may have been more focused on issues of equity rather than issues of dissemination, subsequent work finds that interaction can, in fact, facilitate knowledge transfer (e.g., Cousins & Walker, 2000; Farrell et al., 2018).

Of course, one prerequisite to successful research-practitioner interaction is practitioners' desire to do so, especially considering the institutional barriers noted above. Literature suggests that both sides are willing to engage in these partnerships, especially as the popularity of research-practice partnerships in education, as well as in other fields, continues to grow (for a review, see Coburn & Penuel, 2016). For example, not only did 89% of the respondents surveyed by Barton and Tindle (2019) report having opinions about the questions/topics that educational researchers pursue, but over half indicated a desire to be involved in the research themselves; this was especially true for convening attendees, of which 77% reported wanting to be involved in research. Admittedly, executing these types of collaborations is not easy (e.g., Agan et al., 2020), but the challenges may be worth enduring given evidence on how effective these collaborations are at facilitating research use.

Indeed, reviews of the existing literature (Dagenais et al., 2012; Schaik et al., 2018) finds research-practice partnerships to be a useful way of encouraging research use—with Dagenais and colleagues (2012) going as far as to conclude that "sustained interactivity between researchers and practitioners *guarantees* the use of the results produced by such common effort" (p. 300; emphasis added). Although recent reviews (i.e., Gorard et al., 2020) are more measured when discussing the possible benefits of these partnerships, extant evidence (albeit somewhat dated) does link interactivity between the research and practitioner communities with greater research use. For example, Huberman's (1990) multiyear case study of a Swiss educational research initiative found interactions between knowledge producers and knowledge users not only fostered subsequent research use, but it also helped breed partnerships that transcended the project itself. When asking knowledge users about the factors that contributed to their understanding and use of findings, his participants reported "establishing 'personal' contacts with researchers" as the most important factor. In addition, Cousins and Simon's (1996) study of policy-induced partnerships in Canada supported Huberman's conclusions by finding that partnerships were positively associated with self-reported research utilization. Even so, Coburn and Penuel (2016) review of research-practice partnerships across fields contends that it is still an "open question" as to how strong this association is, though IES-sponsored work by these same authors is encouraging (Farrell et al., 2018).

A second, closely related literature discussed by Coburn and Penuel (2016), which examines how these partnerships impact the research-related attitudes of educators, is both larger and more conclusive. For example, early work by Cousins and

Walker (2000) found teachers' participation in education research to be the strongest predictor of five different research-related attitudes, including perceptions of research's utility and relevance to practice. Science educators in Ratcliffe et al.'s (2005) interviews asserted that greater collaboration between practitioners and researchers could help change the negative attitudes that some educators had of educational research. Cantalini-Williams and colleagues (2015) found elementary school teachers' participation on a Collaborative Inquiry Team with university researchers not only contributed to more positive attitudes about educational research, but it also improved teachers' self-perceived abilities to conduct research on their own; these same attitudinal changes have been reported by district-level administrators engaged in research-practice partnerships (Farrell et al., 2018). Finally, participants of an Australian "Academic Partners" program acknowledged the intricacies of research-practice partnerships while also noting their benefits (Beveridge, Mockler, & Gore, 2018). Teachers, as well as the researchers they partnered with, expressed appreciation for the "two-way learning" (p. 32) that occurred during the partnerships. Whereas teachers felt like the program humanized academics and their researchers, researchers learned to recognize the value of teachers' insider knowledge. Taken together, these findings support the predictions of Farley-Ripple et al's (2018) model. By encouraging "cross-cultural" communication, research-practice partnerships help converge the prevailing perspectives of the research and practice communities.

**Indirect Disseminative Relationships**

Given this project's focus on the What Works Clearinghouse, it is necessary to also acknowledge the ways in which exchanges between the research and practice

communities can be mediated by individuals (e.g., a district- or school-level data champion; Gorard et al., 2020), initiatives (e.g., Best Evidence Encyclopedia; Center for Data-Driven Reform in Education, 2013), or organizations (e.g., think tanks; Cooper, 2013). Although scholarship continues to be preoccupied with "individual framings" (Nutley, 2009, p. 556) of research use, there is some attention being paid to the ways these "research brokers" attempt to bring educational research into practice. That work has culminated in several important insights, including the observation that brokering efforts are playing an increasingly important role as efforts to bridge the research-practice gap intensify. This importance has been reported across time and context. Rich's (2005) analysis of U.S. think tanks, which often play the part of research broker, found that their number had quadrupled between the 1970s and 2000s. Similarly, Cooper's (2010, 2012, 2013) work has long-noted the growing number (and influence) of knowledge mobilization intermediaries in the Canadian education landscape, with nearly half (43%) of the intermediaries in her sample having been founded since 2000 (Cooper, 2012). Recent work in the United Kingdom recognizes the growing popularity of another type of broker: research champions. Research champions are teachers who are responsible for occupying the "third space" between research and practice at their schools (Burn, Conway, Edwards, & Harries, 2020; Gorard et al., 2020). Even though it functions external of schools and their districts, the What Works Clearinghouse can be understood as occupying a similar role (Farley-Ripple el al., 2018).

Perhaps in response to the increasing popularity of research brokering efforts, other literature has attempted to typologize these efforts and draw distinctions between them (Ward, House, & Hamer, 2009; Cooper, 2012; Neal et al., 2015, 2019).

Fundamental to this work is the realization that not all knowledge brokers function identically. Neal and colleagues' (2015, 2019) work is especially informative here because it (a) implies an allegiance to Two-Communities Theory and (b) extends a previously proposed typology (e.g., Gould & Fernandez, 1989) to the education sector. Gould and Fernandez (1989) first identified five distinct types of brokers—all of which can be grouped into two larger classifications. One of these classifications refers to broker types that facilitate knowledge transfer from one practice community (e.g., school administrators in one district) to another practice community (e.g., school administrators in another district). This transfer process can be mediated by a fellow practitioner (e.g., a superintendent, which constitutes *coordinator brokerage)* or from a member of the research community (e.g., an educational psychologist, which constitutes *itinerant brokerage*). The other classification refers to brokerage bridging the research and practice communities. This type of transfer can be mediated by a practitioner (e.g., a research champion, which constitutes *gatekeeper brokerage*), a researcher (which constitutes *representative brokerage*) or by a member of some other subgroup (e.g., an outreach coordinator from an education-oriented nonprofit, which constitutes *liaison brokerage*). In interviews with employees from two public school district, Neal and colleagues (2015) found gatekeeping to be the most commonly reported form of brokerage.

Unfortunately, and as conceded elsewhere (e.g., Cooper, 2012; 2013), little is known about the effectiveness of these brokerage types at impacting research utilization in education. Neal and colleagues (2015, 2019) do offer observations about what types of brokerage might be most effective for knowledge transfer. For instance, in Neal et al.'s (2015) study, only a small fraction (i.e., 18%) of all brokerage chains mentioned by their

respondents could be coded as successful bridging of the research and practice communities. In other words, brokerage of information between practice communities was far more common than brokerage involving both the research and practice subgroups. Subsequent work (Neal, Neal, Mills, Lawlor, McAlindon, 2019) analyzed the social networks of administrators recruited through a statewide random sampling procedure. They were especially interested in how these administrators obtained information about instructional programs that might be implemented in their schools/districts. Results suggested that educators were five times less likely to receive information from a researcher if they relied on itinerant brokerage; alternatively, they were twice as likely to obtain information from a researcher in cases of representative or liaison brokerage. In response, they concluded that "[i]nterventions designed to narrow the research-practice gap could focus on identifying and leveraging individuals poised to serve as brokers in a liaison capacity" (Neal et al., 2019, p. 48), even though liaison brokerage was reported with relative rarity by their sample.

**Relationships Maintained by the WWC**

Encouragingly, liaison brokerage is exactly the type of brokerage the WWC is designed to facilitate. The WWC does not conduct research itself; instead, it acts as an intermediary between those who conduct research and those who will (hopefully) use it in practice. Indeed, in their example of liaison brokerage, Neal et al. (2015) referred to a foundation that "synthesizes" and "distributes" research (p. 5)—both of which the WWC does as well. Recognizing the WWC as a liaison broker means recognizing that it is well-positioned to bridge the research and practice communities.

Just because it is well-positioned, however, does not mean that it is functioning effectively. In fact, similar to how Neal and colleagues (2015) differentiate between the *research-to-practice* gap and the *research-practice* gap, it is important to differentiate between the WWC's role at brokering the former rather than the latter. The idea of a research-to-practice gap implies a unidirectional flow of research information—namely, from the research community to the practice community. Even though these unidirectional disseminative relationships have been theorized as less effective than bidirectional ones (see above), the WWC tends to function more unidirectionally than bidirectionally. Of course, there are exceptions. For example, the WWC offers a service called the WWC Help Desk. This service allows users to ask questions about the WWC review process, receive help with navigating its resources, or offer suggestions—all through the use of an online contact form. Another example is how the WWC develops its practice guides, and more specifically, who is included when doing so. As stated in its Procedures Handbook (WWC, 2020d), practitioners are invited to participate in the process:

> "Practice guides are developed under the guidance of a panel composed of at least six members. Each panel is chaired by a nationally recognized researcher with expertise in the topic. The panel consists of at least four researchers who have diverse expertise in the relevant content area and/or relevant methodological expertise, along with at least two practitioners who have backgrounds that allow them to offer guidance about implementation of the recommendations." (p. C-2)

Although, in this example, representatives from the research community still outnumber those from the practice community, there are certainly cases where the WWC facilitates "exchange" (Farley-Ripple et al., 2018, p. 241) between these communities.

At the same time, there are plenty of instances where members of the practice community are rarely included or solicited for feedback, if at all. As conveyed above, the

procedure used to identify topics for future WWC review protocols does consider "public nominations" from the WWC Help Desk, but it also considers "scans of the literature or of research funded by the U.S. Department of Education" (WWC, 2020d, p. A-2). In this way, the federal government has an opportunity to simultaneously control its own research agenda and the research reviewed by the WWC. There are other examples where the exclusion of practitioners is even more stark, such as in how the review protocol takes shape once the topic area has been decided upon. When describing this process in its Procedures Handbook, the WWC (2020d) specifies that:

> "The review team leadership—including a lead methodologist and content experts as described in appendix C—makes decisions about key parameters, such as eligible population groups, types of interventions, study characteristics, and outcomes of interest" (p. 5)

Elsewhere, the WWC states that its review protocols are "[d]eveloped with substantive experts" (WWC, n.d.-e). Together this suggests that, even in cases where the practice community helps nominate a research topic area for study, it is ultimately up to "experts"—and more cogently, members of the research community—to identify the student subgroups and outcomes of interest. This then constitutes a unidirectional disseminative relationship, not a bidirectional one.

### Structures/Processes/Incentives

Finally, this review recognizes—just as Farley-Ripple and colleagues' (2018) do—that "a range of conditions influence use, including organizational structure, culture, and leadership" (p. 240). These factors can both promote and discourage practitioners' use of educational research. Accordingly, this section includes an overview of (a) tangible structures that facilitate or hinder research use, and (b) the institutional/cultural norms that impact use. In nearly all of cases, these factors exist outside the jurisdiction of

initiatives like the WWC, and as such, little can be said about how the WWC contributes to them. Even so, this section of the review is critical because it communicates the extent to which educational research use is an institutional-level issue, not only a practitioner-level one.

**Models of Institutional Influences on Research Use**

Before introducing these institutional factors, it is worth showing how models of the research-practice gap, as well as subsequent suggestions for closing it, have recognized how external factors influence practitioners' research uptake. Though Nutley and colleagues (2008, 2009) were not the first to acknowledge this, they remain some of the first scholars to clearly articulate its significance. Consider, for example, the following proclamation:

> "For progress to be made, there is a need to move beyond individualized framings of research use in order to capture what using research might mean within wider organizations and systems. In line with this, there is increasing interest in the vision of evidence-based practice encapsulated in the embedded research and organizational excellence models." (Nutley et al., 2009, p. 556)

As suggested here, Nutley et al.'s work differentiates between three models of evidence-based practice (a) *the research-based practitioner model*, the (b) *embedded research model*, and the (c) *organizational excellence model*. *The research-based practitioner model* represents the commonly conceptualized model of educational research use, whereby an individual practitioner is expected to seek out research literature and use it to inform their practice. This model can be likened to the user-pull process (e.g., Lavis et al., 2003) discussed above. Second, the *embedded research model* moves away from the idea of research-based practice as an individual-level responsibility—namely, a responsibility on the part of the educator. Instead, it recognizes that research makes its

way to practice through a messy, indirect pathway shaped by systems and policies external of the individual. Nutley et al.'s (2008) third model, *the organization excellence model*, closely resembles *the embedded-research model*. It differs slightly, however, in that it focuses on the role that external organizations, such as third-party research brokers, play at translating research-based insights. One of the primary tensions between these models is how they view the role of the practitioner as a research user. Whereas the research-based practitioner model assumes that practitioners have a degree of individual autonomy, the embedded research and organizational excellence models acknowledge the structural constraints faced by practitioners, and thus they try to circumvent them (in the case of the embedded research model) or ameliorate them altogether (in the case of the organizational excellence model).

Ever since Nutley and colleagues (2008, 2009) proposed these models, reviews of the empirical literature have found the research-based practitioner model to be shortsighted. For example, Dagenais et al. (2012) identified ten school-level characteristics, such as a school's commitment to organizational learning or its prioritization of research-related professional development activities, that impacted research use. Cain (2016) summarized work suggesting that teachers' use of research is sometimes motivated by extrinsic factors, such as a school or district's promotion offerings. Schaik and colleagues' (2018) review culminated in their identification of four levels that impacted educators' academic knowledge use—one of which is referred to as the "school-organizational level." Ultimately, these reviews justify Farley-Ripple et al's (2018) assertion that:

> "[because] contextual factors related to structures, processes, and incentives influence research use, it is important to understand when and to what degree

> these factors increase or reduce the gap between research and practice communities." (p. 240)

**Structural Influencers of Research Use**

The number of structural factors identified in the literature as facilitators or barriers to research use are too numerous to review here. Thus, I focus on three of the most commonly identified ones, including (a) the time allotted for research engagement, (b) the offering of research-related professional development, and (c) the agency afforded to teachers to implement evidence-based practices or programs. As briefly noted earlier, teachers' lack of time is oft-cited in existing reviews as a structural factor impacting research use (Hemsley-Brown & Sharp, 2003; Dagenais, 2012; Cain, 2016). Williams and Coles (2007) study of teachers' information literacy asked participants to rank-order a list of possible barriers to accessing educational research; lack of time was the most frequently endorsed barrier to access, with three-quarters of participants rating it as a "moderate" or "heavy" barrier. In another survey (i.e., Lysenko et al., 2014), having "available time to read a journal" was found to be a self-reported factor in educators' research use—even more so for teachers than for administrators. Qualitative work has found this to be the case well. Canadian teachers participating in Martinovic and colleagues' (2012) online focus groups reported time constraints as a limitation to their engagement with research-based practices, and a survey administered as part of that study suggested similarly. Likewise, recent work by Joram and colleagues (2020) concluded that "time is still the greatest factor limiting this [research] engagement" (p. 8). Unique to that work, however, was one participant's experience of having "dedicated time for discussing research…", and as such, they felt "…very supported in that regard" (p. 8).

Given the evidence, though, teachers having sufficient time to engage with research appears to be a rarity rather than a regularity.

A second structural factor noted in the literature is the availability of professional development opportunities offered at the school and/or district level (see Schaik, 2018). Much of this work asks educators if they believed research-related professional development could encourage greater research use in the classroom, and in response, many believe that it would. For instance, teachers, administrators, and individuals serving in intermediary roles told Vanderlinde and van Braak (2010) that professional learning communities (PLCs) would be an effective way of bringing research to practice. Others, such as the teachers interviewed by Behrstock-Sherratt, Drill, & Miller (2011), believed that professional development (PD) sessions could help bridge the research-practice gap. Based on survey results and follow-up interviews, Lysenko et al. (2015) also concluded that "on-going research-based professional development" (p. 51) could help support practitioners' research use. Comparably little work has explored the extent to which research-based PD does, in fact, facilitate use, and the findings in that literature are mixed. Promising findings emerged from Dresner and Worley's (2006) study of a PD initiative that facilitated relationships between teachers and researchers. Not only did teachers report value in those partnerships after the initiative concluded, but interviews five years later found that many teachers still believed the PD to be impactful. Alternatively, however, subsequent work conducted by Berhstock-Sherratt and colleagues (i.e., Drill et al., 2013) found that teachers' participation in a research-focused PD opportunity did not lead to greater research use. Even so, there is still some evidence that research-related PD can improve educators' attitudes about educational research, or,

as suggested by Williams and Coles (2003, 2007a), enhance their abilities to access that research.

A third structural factor associated with teachers' research use is the degree of agency they have in making decisions about what or how they teach. Leat, Reid, and Lofthouse (2015) identified teacher agency as one of five themes in what is known about educators' engagement with research in practice. They observed that a lack of agency may be one reason why teachers do not use research "even where conditions may appear to be superficially conducive to research engagement" (p. 277). Indeed, Joram and colleagues' (2020) findings speak exactly to this point. In that work, several teachers explained how they had little decision-making power—that they were "passive recipients" (p. 6) of decisions made by administrators. One teacher explained how their attempts at innovating their school's curricula were dismissed by administration. Another noted how, even if they wanted to change their teaching strategies in accordance with educational research, they would not be able to:

> "There seems to be an implicit assumption that, with research informing decisions, teachers are free to experiment in their classes. While this may be true in a limited capacity, teachers are often limited by the initiative of their district, school or division." (Joram et al., 2020, p. 8)

At the same time, a different teacher shared how he was given a great deal of latitude to experiment in the classroom, suggesting that teachers who are allowed agency may be more likely to implement research-based practices.

**Cultural Influencers of Research Use**

These structural influencers of research use are both contributors to and byproducts of the larger institutional culture they are embedded in. To acknowledge structural influencers of research use without attending to cultural influencers would be a

mistake; both are important (for a discussion, see Levin, 2011). Though measuring these cultural factors and their impact on research use presents a challenge, researchers have attempted to do exactly that. Findings suggest that educators operate within cultures that can either encourage or discourage research use, and in turn, the influence of these cultures on research use is significant (e.g., Hemsley-Brown & Sharp, 2003; Dagenais et al., 2012; Schaik et al., 2018). Some of this work has merely described what educators perceived their institutional research culture to be. For instance, Borg (2008, 2009) asked respondents in his survey to rate their agreement with a number of statements related to their institutional culture. Findings from his 2008 study demonstrated that, in general, teachers did not report their institutional culture as being very research-friendly. As an example, only 33% of teachers agreed that "teachers at their school talk about research", and just over 20% agreed that "management encourages teachers to do research" (Borg, 2008, p. 7). Focus groups conducted by Martinovic et al. (2012) and Joram et al., (2020) further suggest that many teachers reside in cultures resistant to inquiry and experimentation.

Other work has gone one step further by examining if and to what extent a research-positive school culture predicts attitudes towards and/or use of research. Cousins and Walker (2000) found small, yet positive associations between educators' ratings of their school's organizational learning capacity and their perceptions of research usefulness and relevance. Penuel et al.'s (2017) survey of school- and district-level administrators found that two institutional characteristics—namely, administrators' self-reported frequency of research-related discussions and their perceptions of their institutional culture as one where "research is seen as useful source of information" (p.

5)—were positively related to research use. Uniquely, this work examined these relationships across instrumental, conceptual, symbolic, and imposed forms of research use. They found the reported frequencies of discussing research were associated most strongly with instrumental use. In contrast, perceptions of institutional research culture were associated most strongly with symbolic usage. Lysenko et al. (2014) found a similar pattern in their majority-teacher sample, with perceptions of a school's openness to research and change positively predicting instrumental, conceptual, and symbolic use; notably, however, the size of these relationships was far smaller than they were in Penuel et al.'s (2017) work. Even so, this work still demonstrates that teachers research use is, to an extent, at the mercy of cultural norms about research that permeate their schools and districts.

**WWC's Approach to These Influencers**

The WWC has acknowledged that educators face barriers to their research use. In their *Procedures Handbook*, they concede that:

> "…it can be difficult, time consuming, and costly for decisionmakers to access and draw conclusions from relevant studies about the effectiveness of these interventions." (p. 1)

Although no specific barriers are mentioned, they can be inferred. First, the idea that it can be difficult for decisionmakers to access and draw conclusions from research seems to allude to one of the WWC's underlying assumptions—namely, that educators are ill-prepared to consume educational research (see Chapter 4). Drawing conclusions from research, at least in its raw form, can be difficult for those without a research background (e.g., Williams & Coles, 2007a). Thus, the WWC identifies and assesses research on behalf of educators. Second, the idea that it can be time-consuming to access and draw

conclusions from research alludes to this assumption as well (i.e., that it might take longer for practitioners to do so than researchers), but it also acknowledges the general lack of time educators have to engage with research, The WWC tries to address this by developing resources that can be engaged with quickly and efficiently. Third, the idea that it can be costly to engage with research may be a reference to paywalls maintained by most traditional research journals, which studies have found impede educators' research use (e.g., Sherratt & Miller, 2011); the WWC circumvents this barrier by offering its resources freely. Accordingly, the WWC is, in some ways, a direct response to factors that hinder educators' research use.

Even so, not all barriers identified in the literature can be avoided by research synthesis efforts like the WWC. As an example, the WWC has little effect on the institutional cultures in which educators operate. In a similar way, the WWC does not affect the structural influencers at the school- and/or district-level that are associated with research use. The agency afforded to teachers to try new things, such as the recommendations of a WWC practice guide, is not determined by the WWC, nor is the amount of time allocated for teachers' research engagement. In sum, there are myriad "contextual factors…[that] increase or reduce the gap between research and practice communities" (Farley-Ripple et al., 2018, p. 240) that exist outside of the WWC's control.

## Takeaways

Despite prevailing concerns about both the quantity and quality of literature examining educators' engagement with research (e.g., Gorard et al., 2020), a substantial amount can still be said on the topic. Here I reviewed literature elucidating educators'

research interests, their preferred research access points, their standards when assessing research quality, the types of relationships that facilitate educators' research use, and the types of external factors that facilitate (or hinder) that use. Given the heft of this review, a series of charts summarizing key takeaways from each of the chapter's five sections are available in the subsequent pages (see Tables 2.1-2.5). These charts will serve as a useful touchpoint in Chapter 4, as contribution analysis invites previously published literature to be used as an evidence source when evaluating a program's theory of change (Mayne, 2008, 2012). In the following chapter, this—as well as the other features of contribution analysis—will be explained in detail.

## Chapter 2 Tables and Figures

Table 2.1 - Chapter 2 Takeaways for the "Problems Addressed by Research" Section

| DIMENSION | TOPIC | TAKEAWAY | REPRESENTATIVE LITERATURE |
|---|---|---|---|
| Problems Addressed by Research | Educators Attitudes About Educational Research | Educators' attitudes about research predict the extent of their research use. | Lysenko et al., 2014; Penuel et al., 2017 |
| | | Administrators generally see value in educational research. | Biddle & Saha, 2006; Penuel et al., 2018 |
| | | Teachers report ambivalence—and in some cases, contestation—towards educational research. | Lysenko et al., 2015; Cain, 2017 |
| | Reasons for Teachers' Skepticism Towards Research Use | Teachers' aversion to risk may contribute to this skepticism. | Le Fevre, 2014; Joram et al., 2020 |
| | | Skepticism may also result from teachers experiencing "non-congruence" between research and their own classroom experiences. | Behrstock-Sherratt, Drill, & Miller, 2011; Cain, 2017 |
| | | Teachers' belief in the uniqueness of their students/setting may contribute to skepticism of research conducted elsewhere. | Cain, 2017; Joram et al., 2020 |

Table 2.1 (continued) - Chapter 2 Takeaways for the "Problems Addressed by Research" Section

| | | | |
|---|---|---|---|
| Problems Addressed by Research | Educators' Research Interests | Administrators may be more interested in school- and teacher-level topic areas; teachers, on the other hand, prefer student-level ones. | Penuel et al., 2018; Barton & Tindle, 2019 |
| | | Teachers are interested in research on various indicators of student success—not only achievement, but also motivation and engagement. | Everton et al., 2000; Barton & Tindle, 2019 |
| | | Teachers are also interested in research related to diversity, inclusion, and multiculturalism. | Ion & Iucu; 2014; Barton & Tindle, 2019 |

Table 2.2 - Chapter 2 Takeaways for the "Usefulness of Research Products" Section

| DIMENSION | TOPIC | TAKEAWAY | REPRESENTATIVE LITERATURE |
|---|---|---|---|
| Usefulness of Research Products | The Nature of Educators' Research Use | Despite desires to use research instrumentally, educators often use research in other ways. | Dagenais et al., 2012; Cain, 2016 |
| | | Imposed use is a common reality for administrators. | Coburn & Talbert, 2006; Penuel et al., 2017, 2018 |
| | | Teachers tend to use research conceptually. | Lysenko et al., 2015; Cain, 2015 |
| | | Administrators' research use tends to outpace teachers' use. | Hemsley-Brown & Sharp, 2003; Lysenko et al., 2015 |
| | Educators' Research Access Points | Educators rely on a diversity of resources to access educational research. | Dagenais et al, 2012; Barton & Tindle, 2019 |
| | | Administrators may be more likely to use traditional access points (e.g., journals), whereas teachers may be more likely to use newer types of access points (e.g., social media). | Lysenko et al., 2014; Barton & Tindle, 2019 |
| | | Both groups, however, are especially likely to use access points that involve interpersonal interaction—like discussing research with a colleague. | Drill et al., 2012; Penuel et al., 2017 |

Table 2.2 (continued) - Chapter 2 Takeaways for the "Usefulness of Research Products" Section

| | | | |
|---|---|---|---|
| Usefulness of Research Products | Characteristics of Favored Access Points | An access point's applicability to practice is an important characteristic to teachers. | Lysenko et al., 2015; Barton & Tindle, 2019 |
| | | Access points that predigest raw research findings into non-technical summaries appeal to educators. | Williams & Coles, 2007a; Vanderlinde & van Braak, 2010 |
| | | Both teachers and administrators prefer access points that feature up-to-date, timely information. | Behrstock-Sherratt et al., 2012; Dagenais et al., 2012 |

Table 2.3 - Chapter 2 Takeaways for the "Nature/Quality of Research" Section

| DIMENSION | TOPIC | TAKEAWAY | REPRESENTATIVE LITERATURE |
|---|---|---|---|
| Nature/Quality of Research | Credibility of Educational Research | Administrators generally find research to be credible, though they still recognize that research can be biased. | Coburn & Talbert, 2006; Penuel et al., 2017 |
| | | Many teachers question the credibility of educational research, believing that it often lacks methodological quality. | Boardman et al., 2005; Lysenko et al., 2014, 2015 |
| | Epistemological Differences | Teachers prefer experiential sources of knowledges to theoretical ones—meaning that they trust knowledge accumulated through practice. | Joram, 2007; Bråten & Ferguson, 2015 |
| | | Teachers generally place greater value on the perceived idiosyncrasies of students than researchers do. | Cain, 2017; Joram et al., 2020 |
| | | Likewise, teachers are more skeptical about the generalizability of research knowledge than researchers. | Joram, 2007; Cain, 2017 |
| | | Nevertheless, some educators conceptualize research similarly to how it is conceptualized in the traditional scientific canon (i.e., the collection of quantitative data; the use of statistics). | Borg, 2008, 2009 |

Table 2.3 (continued) - Chapter 2 Takeaways for the "Nature/Quality of Research" Section

| | | | |
|---|---|---|---|
| Nature/Quality of Research | Standards Used by Educators to Determine Research Quality | In some cases, educators judge research quality according to the same indicators used by the research community (e.g., hypothesis-testing, control of variables, sample size). | Ratcliffe et al., 2005; Borg, 2009; Joram, 2009; Cain, 2016 |
| | | Alternatively, many educators still place greater value on characteristics of external validity than internal validity. | Finnigan et al., 2013; Neal et al., 2018; Joram et al., 2020 |

Table 2.4 - Chapter 2 Takeaways for the "Relationships Between Communities" Section

| DIMENSION | TOPIC | TAKEAWAY | REPRESENTATIVE LITERATURE |
|---|---|---|---|
| Relationships Between Communities | Unidirectional Disseminative Relationships | Traditional disseminative relationships involve the passing of knowledge from researchers to users ("producer-push"); in turn, users may seek out this knowledge ("user-pull") | Lavis et al., 2003; Levin, 2011; Tseng, 2012 |
| | | This type of relationship remains more common than other ones. | Nutley et al., 2009; Cooper et al., 2012 |
| | | Even so, evidence suggest that this type of relationship may be ineffective—either because of characteristics related to the research itself (e.g., jargon) or external barriers to users' access (e.g., lack of time). | Levin, 2011; Anwaruddin, 2015; Gorard et al., 2020 |
| | Bidirectional Disseminative Relationships | Bidirectional relationships are characterized by interactions between researchers and knowledge users (e.g., research-practice partnerships). | Lavis et al., 2003 |
| | | Research-practitioner partnerships may elicit research use, though more evidence is needed. | Coburn & Penuel, 2016 |
| | | Research-practitioner partnerships commonly leave practitioners with more positive attitudes about educational research. | Cousins & Walker, 2000; Cantalini-Williams et al., 2014: Farrell et al., 2018 |

Table 2.4 (continued) - Chapter 2 Takeaways for the “Relationships Between Communities” Section

| | | | |
|---|---|---|---|
| | Indirect Disseminative Relationships | Sometimes the dissemination of research is “brokered” through a third party; this has become increasingly common. | Ward et al., 2009; Neal et al., 2015 |
| | | Certain types of brokerage, such as those that use a member from some other subgroup to connect researchers and practitioners, may be the most effective type of brokerage. | Neal et al., 2019 |

Table 2.5 - Chapter 2 Takeaways for the "Structures/Processes/Incentives" Section

| DIMENSION | TOPIC | TAKEAWAY | REPRESENTATIVE LITERATURE |
|---|---|---|---|
| Structures/Processes/Incentives | Models of Institutional Influences on Research Use | Models of evidence-based practice indicate that institutional factors are key facilitators of (or barriers to) practitioners' research use. | Nutley et al., 2008, 2009 |
| | | Subsequent reviews of the literature indicate that these factors are, indeed, influential. | Dagenais et al., 2012; Schaik et al., 2018 |
| | Structural Influencers of Research Use | The time allotted for practitioners' engagement with research is a frequently-noted influencer of use. | Williams & Coles, 2007b; Martinovic et al., 2012; Joram et al., 2020 |
| | | Educators believe that research-related professional development opportunities could help elicit use, though more research is needed. | Dresner & Worley, 2006; Vanderlinde & van Braak, 2010; Drill et al., 2013 |
| | | The agency afforded to teachers is another factor that influences their use of research-based evidence. | Leat et al., 2015; Joram et al., 2020 |
| | Cultural Influencers of Research Use | Educators operate within cultures that either encourage or discourage research use. | Levin et al., 2011; Schaik et al., 2018 |
| | | Research has found educators' perceptions of their organizational culture to be associated with research-related attitudes and use. | Cousins & Walker, 2000; Penuel et al., 2017 |

# CHAPTER 3 – METHODS

Is the What Works Clearinghouse (WWC) working? If not, are its underlying assumptions to blame? And if so, how can they be rectified in ways that lead to program improvement? In this project, I pursue answers to each of these questions using a methodology called contribution analysis (Mayne, 2008, 2012a, 2019; Better Evaluation, 2020). Not only is contribution analysis flexible enough to address each of these questions, but, as described in Chapter 1, it satisfies the methodological appropriateness principle of evaluative thinking (ET). Indeed, as I hope to show both here and in the subsequent chapter, the process of executing a contribution analysis is, in itself, a form of ET. Put another way, the execution of a contribution analysis entails exactly what Vo (2013) specifies in her definition of ET, which she described as:

> "…the process by which one marshals evaluative data and evidence to construct arguments that allow one to arrive at contextualized value judgments in a transparent fashion" (p. 107).

In the following pages, I first introduce contribution analysis and justify its appropriateness for the current project. After that, I describe how it was used to answer my questions about the WWC.

## Contribution Analysis: An Overview and Justification

Contribution analysis is an analytical technique consistent with the theory-based evaluation tradition (for a discussion of the relationship between the two, see Leeuw, 2012). Theory-based evaluation emerged in the 1970s and 1980s as an alternative to the "experimental paradigm" of program evaluation, which dominated the field up until that point (e.g., Rossi & Wright, 1984). As implied by its name, the "experimental paradigm" encouraged program evaluations to utilize random sampling procedures and experimental

research designs whenever possible. Early proponents of theory-based evaluation (e.g., Chen & Rossi, 1983, 1989) sought an alternative to the experimental paradigm because they believed it to be reductionistic. Specifically, they noticed that when an evaluation adhered to the experimental paradigm, it was "not necessary to understand how a social program works in order to estimate its net effects" (Chen & Rossi, 1983, p. 284). In this way, the evaluation becomes less informative than it could have otherwise been. It is less informative because, even if a program is found to work, the evaluator may not know which aspects of the program to celebrate; alternatively, if a program is found to be deficient, the evaluator may not know which aspects of the program to recommend revising. As such, the theory-based evaluation tradition, which has grown in popularity since the pioneering days of Chen and Rossi, seeks to make evaluations more informative by investigating the "black box" (Chen & Rossi, 1983, p. 291) between a program's inputs and impacts.

Contribution analysis is one way of doing so. Its aim is to determine if, how, and why a program has contributed to its intended result, and it does so by scrutinizing the program's *theory of change*. Theories of change are visual articulations of how a program's activities are expected to contribute to an intended impact (e.g., Rogers, 2008, 2014). Much has been written about theories of change, such as how they should be developed, what they should include, and how they should be assessed (see Stein & Valters, 2012 and Breuer, Lee, De Silva, & Lund, 2016 for reviews). For our purposes, it is enough to think of theories of change as program logic models with additional components. One of these components, which is central to theories of change but often omitted from logic models, is a program's *causal assumptions*. Causal assumptions are

the underlying events or conditions that are needed for a program's underlying logic to progress as planned. For example, a food pantry program designed to mitigate college student food insecurity makes several assumptions about how its activities (e.g., making free or reduced-price food available to beneficiaries) contributes to its intended impact (e.g., reduced student food insecurity). Specifically, it makes assumptions about the program's reach (e.g., students are aware of the pantry's existence), outputs (e.g., students like/want the available food enough to visit the pantry), outcomes (e.g., a student will visit the pantry repeatedly so that they become food secure), and others. By explicating these assumptions along with other program components, a well-articulated theory of change acts as a compass during complex evaluations—pointing the evaluator in directions where the program's logic or underlying assumptions warrant additional exploration.

Because of its ability to add clarity to particularly complex evaluations, contribution analysis has been both well-received (Mayne, 2011; Patton, 2012) and well-utilized (e.g., Delahais & Toulemonde, 2012) in the evaluation profession. This is also why contribution analysis is well-suited for the current project. Indeed, one strength of contribution analysis is its ability to generate credible contribution claims in settings where experimental evaluation designs are untenable (Mayne, 2012). The What Works Clearinghouse exists in such a setting. To evaluate the WWC using a traditional counterfactual notion of causality—which, as explained by Mayne (2019), would require determining "what would have happened without the intervention" (p. 174)—is impractical given the resources needed to conduct a true experiment. Instead, contribution analysis proceeds according to a generative notion of causality. This means

that the evaluator is asked to "infer" causality based on how "reasonable" each step in a program's theory of change seems based on available empirical evidence (Mayne, 2019). Dybdal, Nielsen, & Lemire (2011) explain this well in the following passage:

> "Hence the challenge is not so much one of establishing, and providing proof, that the program caused the outcomes (i.e., demonstrating attribution), but rather of providing the best possible evidence that the intervention contributed to the outcomes of interest (i.e., demonstrating contribution)." (p. 32)

Not only is this approach more practical, but it expands the evaluation's scope from "if" a program worked to "how" or "why" it might work based on the program theory.

As mentioned above, by relying so heavily on a program's theory of change, contribution analysis also requires that a program's underlying assumptions be systematically identified and tested. Its attention to these assumptions is a second reason why contribution analysis is well-suited for this project. Indeed, critics of the "what works" movement in education often construct their arguments by identifying, unpacking, and renouncing the assumptions on which it is based (e.g., Biesta, 2007, 2010); rarely, however, are these assumptions evaluated based on the available empirical evidence. Contribution analysis requires the evaluator to do exactly that. Additionally, just as attending to programmatic assumptions is important for evaluative purposes, it is important for theoretical purposes as well. As explained in Chapter 1, Farley-Ripple and colleagues (2018; see also CRUE, n.d.) propose a theoretical framework that attributes the research-practice gap to differences in the assumptions held by the research and practice communities. Because contribution analysis encourages evaluators to outline programmatic assumptions when developing theories of change, conducting a contribution analysis for the purposes of this project requires that I cross-examine the

WWC's underlying assumptions with those of educators. Doing so is, in effect, a test of Farley-Ripple et al.'s framework.

A third reason why contribution analysis is well-suited for this project is that it produces information helpful for program redesign. To reiterate, my goal is to couple a summative evaluation of the WWC's impact with formative recommendations for program improvement. Articulating the WWC's theory of change is vital to this goal, as it can help pinpoint breakdowns in the program's theory of change. Pinpointing specific breakdowns will, in turn, help me generate more targeted feedback for the WWC.

## Using Contribution Analysis in the Current Evaluation

To conduct a contribution analysis of the WWC, I followed the six-step procedure developed by Mayne (2008, 2012b, 2019). This approach has been used in numerous evaluations to date, including those examining research-to-practice knowledge transfer (e.g., Morton, 2015; Riley, Kernoghan, Stockton, Montague, Yessis, & Willis, 2018). Although Mayne has made few changes to this procedure over time, evaluators have begun taking liberties when adapting his procedure to their projects (e.g., Budhwani & McDavid, 2017); in fact, Mayne (2019) embraces these modifications of contribution analysis as "good practice" (p. 271). Accordingly, though I have tried to stay faithful to Mayne's six steps, I will acknowledge times when my approach strays from his. Most notably, Mayne's (2012b) prudent suggestions for including program stakeholders at various stages of the analysis were not used here. This was due to issues of practicality, as well as the fact that this project is not formally sponsored by the Institute of Education Sciences (IES).

Below I overview this project's adherence to each of these six steps, noting their data sources and, when applicable, how those sources were analyzed. The step-by-step nature of contribution analysis does not lend itself well to a separated discussion of methods and results, so some methodological information will be shared in Chapter 4 as well. Ultimately, the following pages are meant to alert the reader to the sequence of steps—as well as the nature of those steps—in which results are discussed in Chapter 4.

**Step 1: Establish the Specific Cause/Effect Issue to Be Addressed**

Before developing a program's theory of change, the evaluator must first explore the nature and extent of the program's expected contribution (Mayne, 2011, 2012b). In other words, they must determine what it would look like for the program to "work." This step is important to developing credible contribution claims, yet it is often neglected (Mayne, 2019). Indeed, it is only after a program's expected contribution is identified that the evaluator can begin hashing out its program theory. Step 1 was approached with two goals in mind. The first was to explore the *nature* of the WWC's expected contribution. The second, somewhat more challenging goal was to discover the *extent* of the WWC's expected contribution—that is, when the WWC's contribution would be large enough to be considered "meaningful." The procedures used to determine both the nature and extent of the WWC's contribution are detailed below.

***Exploring the Nature of the WWC's Expected Contribution***

Because this project is not being commissioned by the IES, a participatory approach to identifying the WWC's expected contribution was not used. Rather than asking stakeholders to reflect on their goals for the program, I collated and analyzed three sets of evidence to determine how both the IES and the WWC have understood the

WWC's roles and responsibilities over time. One set of evidence comes from the IES's original conceptualization of the WWC prior to its founding. In other words, how did policymakers envision the WWC's contribution before it was ever made public? To answer this question, I examined comments made by the IES's founding director—Grover Whitehurst—in both the years leading up to the WWC as well as immediately after its founding. Those comments were extracted from:

**Congressional hearings.** In 2002, Whitehurst participated in two congressional hearings regarding the reauthorization of the Office of Education Research and Improvement (OERI). Because the IES was designed to replace OERI as part of the *Education Sciences Reform Act of 2002*, Whitehurst provided testimonies in both House and Senate hearings to outline his vision for the IES and answer questions posed by lawmakers. Those hearings took place in front of the House Subcommittee on Educational Reform (part of a larger committee on Education and the Workplace) and Senate Committee on Health, Education, Labor, and Pensions. I acquired transcripts from these testimonies and reviewed them for comments about the WWC. Representative statements are included in Table 4.1.

**Interviews.** Shortly after adopting his post as Director of IES, Whitehurst promoted the WWC in an interview with *THE Journal: Technological Horizons in Education* (Mageau, 2004). *THE Journal* is a practitioner-friendly journal dedicated to

> "…informing and educating K-12 senior-level district and school administrators, technologists, and tech-savvy educators within districts, schools, and classrooms to improve and advance the learning process through the use of technology" (THE Journal, 2020).

In this specific interview, Whitehurst is asked questions about the WWC, its standards, and its role as part of No Child Left Behind legislation. His comments in this interview

were thought to be a useful supplement to those made in the Congressional hearings, as they were ultimately intended to be heard by practitioners—not policymakers.

**Lectures.** In his first two years as IES director, Whitehurst also gave several lectures in academic settings. Those included his invited address at the 2003 annual meeting of the American Educational Research Association, as well as his 2004 lecture at Northwestern University's Institute for Policy Research. Transcripts of both lectures were available for review. Though both focused on the IES's mission more broadly, Whitehurst does name the WWC as a part of that mission in both occasions. My analysis of WWC-related statements from these lectures, as well as those made in the *THE* interview and congressional hearings, is presented in Chapter 4.

Because Whitehurst's comments were made early in the WWC's existence, a second set of evidence examined how the WWC's desired contributions may have changed over time. Specifically, it sought insights into how the WWC has self-conceptualized its mission over the past two decades. To do so, two sources of evidence were analyzed, which included:

**Standards and Procedures Handbooks.** Since 2008, the WWC has formalized its research review standards, as well as its procedures for implementing those standards, in a series of handbooks. These handbooks have been archived and made available for public consumption on its website (WWC, n.d.-d). Six iterations of these handbooks have been published to date, and for this analysis, I reviewed all six for information about the WWC's intended purpose and/or mission. Version 1 was sparse and only featured information about how the WWC vets educational research, so it is not included in Table 4.2, which includes excerpts from the five remaining handbooks.

**"What We Do" Video.** Because the handbooks were largely technical in nature, they appeared more suitable for researchers interested in the intricacies of the WWC's review process—not intended beneficiaries. It was therefore important to find a complementary data source showcasing how the WWC presents itself to its intended beneficiaries. I found it in a short, animated video entitled "What We Do" (IES, 2018c), which overviews the WWC's mission and its various resources for educators. The video was transcribed using the native capabilities of YouTube.com, and that transcript was later analyzed for information about the WWC's intended contribution.

Given the extent to which the WWC was described in these sources as an aspect of the IES's larger mission (which is discussed in Chapter 4), a third set of evidence seemed necessary. Specifically, now that it is nearly two decades old, how is the WWC viewed by the IES? Evidence related to this line of inquiry was found in three places:

**IES Director Blog Posts.** Mark Schneider, the current IES director, intermittently writes blog posts with IES-related updates. I reviewed these posts and found three with mentions of the WWC. Those included a 2018 post entitled "Changes are Coming to the WWC," which, of the three, focused most on the WWC. Others included a 2019 post summarizing his first year in office entitled "First Year Accomplishments", and remarks from January 2020 entitled "A New Year's Update from the IES Director". Excerpts from these posts are available in Table 4.3.

**IES Director's Biennial Report.** The IES Director is also required, on a biennial basis, to submit a report to Congress outlining IES activities and contracts awarded over $100,000. The latest available report overviews the 2017 and 2018 fiscal years, and it is posted on the IES's "About Us" webpage for public consumption. It also includes

updates on IES initiatives, such as the WWC, which is why this report proved useful in better understanding how the IES currently understands the WWC's contributions.

**IES "About Us" Webpage.** This webpage also had information about the IES's own mission as well as its conceptualization of the WWC as part of that mission.

### *Exploring the Extent of the WWC's Expected Contribution*

As alluded to above, contribution analysis does not merely require that the evaluator identify the expected nature of a program's contribution. For judgements of that contribution to be accurate, the evaluator also needs to determine the extent—or, in other word, the size—of that expected contribution. To determine if the WWC has made a meaningful impact, a predetermined benchmark from which to measure that impact is needed. Although the WWC currently exists without publicized benchmark(s) of program success (an observation also noted by the Government Accountability Agency; GAO, 2010), a benchmark did, at one time, exist. In an evaluation of the IES conducted by Baldwin and colleagues (2008), the authors allude to a benchmark that IES established as part of the Bush Administration's *Program Assessment Rating Tool* (PART). PART is a now-defunct federal initiative that required the Office of Management and Budget to rate all federal programs on their effectiveness. An aspect of PART was the establishment of performance indicators from which programs could be judged. According to Baldwin and colleagues (2008), one of IES's PART goals was specific to the WWC:

> "The complexities of increasing utilization are acknowledged in the IES PART long-term outcome measure that focuses on the percentage of decisionmakers surveyed in 2013–2014 who indicate they consult the What Works Clearinghouse prior to making decision(s) on reading, writing, math, science or teacher quality interventions. The target set for 2013–2014 is 25 percent." (p. 108)

This same benchmark has been noted in other sources as well (GAO, 2010). As such, it grounded my own understanding of how the IES envisioned the extent of the WWC's contribution.

**Step 2: Developing the Theory of Change**

Whereas Step 1 explored *what* the WWC intends to contribute, Step 2 explores *how* the WWC pursues its contribution. This is done by articulating its theory of change. As noted earlier, a theory of change is a visual model of how a program is intended to work. Although they are commonly used for program design and management, theories of change are also useful when assessing a program's impact (Mayne, 2015, 2017). This is because they articulate each step in a program's causal chain—from its activities to its outcomes to its impact (as well as the assumptions in between). Doing so helps keep evaluations of complex programs focused and manageable. It also helps identify weak points in a program's theory that are, in turn, good candidates for program redesign. Given the WWC's complexity as well as my goal of providing feedback for program improvement, developing a robust theory of change was of the upmost importance.

Although developing a well-articulated theory of change might be considered a precursor to contribution analysis, it is more than that—especially for programs who have existed without one. Theories of change are, as implied by Mayne (2017), an important *result* of the contribution analysis process; hence, I both discuss it and depict it (see Figure 4.1) in Chapter 4. Here I overview the general procedure used to develop the WWC's theory of change, as well as some of the limitations of that procedure.

### *Procedure*

In his walkthroughs of contribution analysis, Mayne (2015, 2017, 2019) has repeatedly shared "generic" theory of change templates to assist with Step 2. An example is included in Figure 3.1. At the same time, however, Mayne has struggled to show evaluators *how* to develop theories of change on their own; he acknowledges this shortcoming (2019), as do others (e.g., Leeuw, 2012). Given this, I developed the WWC's theory of change using a hybrid approach. Specifically, I applied Leeuw's "policy-scientific method" of reconstructing program theory (1991, 2003) to the theory of change template suggested by Mayne (2015, 2017, 2019). Leeuw suggested that, in situations where stakeholders have inadequately articulated their program's theory of change (or not even tried to do so), evaluators can use artifacts, such as program documents and interview transcripts, to better understand "what the goals are of the policy or program under review" (2003, p. 7). His approach has been used in projects like this one, such as when Ehren and colleagues (2005) evaluated the impact of the *Dutch Educational Supervision Act*. The remainder of Leeuw's method involves the conversion of information gleaned from artifacts into a theory of change model. I did the opposite. Rather than build a theory of change from the extracted evidence, I systematically searched for evidence from which to "fill in" the various components of Mayne's template. The main components of Mayne's template—including both the program's impact pathway and its underlying assumptions—are described below.

**Impact Pathway.** A program's impact pathway can be thought of as its logic model. It explicates the sequence of steps comprising the program's causal pathway—from activities to outputs to outcomes to impacts. In Figures 3.1 and 4.1, the impact

pathway is centered and demarcated by bolded boxes and arrows. To articulate the steps in this pathway, I relied on many of the same evidence sources listed in Step 1, including (a) Grover Whitehurst's commentaries, (b) the WWC Procedures and Standards Handbooks, and (c) IES documentation, including its webpages and biannual reports.

**Assumptions.** The second component integral to Mayne's (2015, 2017, 2019) theory of change is a program's underlying assumptions. One type of underlying assumption is a *causal assumption*, which represents an event and/or condition "likely necessary" (Mayne, 2019, p. 172) for one step of the impact pathway to proceed to the next. I have labelled these in Figure 4.1 and represented them with dashed boxes. As an example, for the WWC's outputs to reach its intended beneficiaries, several conditions must be met, including beneficiaries being aware that they exist. The second type of assumptions are called *rationale assumptions*, which are the underlying premises on which a program is founded; in Figure 4.1, they are represented with dotted boxes. Although Mayne (2015) discusses this class of assumptions in less detail, I have included them in my theory of change given their recent popularity in evaluation-related publications (e.g., Nkwake, 2013; Nkwake & Morrow, 2016). Evidence of these assumptions were found in some of the same congressional hearings participated in by Whitehurst (Hearing Before the Senate Committee, 2002b, Hearing Before the House Subcommittee, 2002c), which have been analyzed in more detail by Zoellner (2010). Importantly, Mayne (2015) concedes that including each and every assumption can muddy a theory of change (e.g., including that "the sun must rise" as a necessary assumption is pointless and distracting). Instead, theories of change should prioritize

those most salient assumptions that, when left unrealized, keep the impact pathway from progressing as planned. I attempted to heed that suggestion here.

***Limitations***

Several qualifications should be recognized before proceeding to Step 3. First, and as noted before, this theory of change was not developed in consultation with stakeholders. This evaluation is not being commissioned by IES but rather completed as part of an independent project. I did, however, reach out to the WWC Help Desk to inquire if a theory of change had already been developed; if so, I had planned to use it. In response to my inquiry, I received the following message:

> "Unfortunately the Clearinghouse does not have a logic model or theory of change. Models and theories are typical for mission-driven organizations (e.g., foundations). The Clearinghouse reviews evidence about what works, but since we are part of independent and nonpartisan federal entity, we don't have a theory of change per se." (personal communication)

Interestingly, as I have shown in Tables 4.2 and 4.3, both the IES and the WWC have explicit missions that, among other things, outline their intended contributions. Even so, I proceeded by drafting the theory of change myself. Because this did not occur alongside stakeholders, it was important that I developed each component of the model using the same evidence sources featured in Step 1.

A second point is that my theory of change is not intended to be comprehensive. Rather, theories of change are meant to be useful to the program and/or its evaluation. In the context of complex programs, useful theories of change often sacrifice detail for clarity (Maine, 2015). Depending on their intended purpose, theories of change can be developed according to varying levels of granularity, which Mayne (2019) refers to as *narrative* (e.g., brief description without attention to assumptions), *overview* (e.g., more

detailed with attention to some assumptions, but not causal assumptions), and *nested* (e.g., specific to one causal pathway and all causal assumptions are explicated). This project's theory of change is best thought of as a nested theory of change, which captures the main causal pathway of the WWC without attending to other IES initiatives such as the ERIC or the RELs—both of which are also meant to help with research dissemination.

Third, this theory of change depicts the WWC's impact pathway and underlying assumptions when beneficiaries engage with it voluntarily. The reality, however, might not be that simple. Weiss's (2005, 2008) notion of imposed research use recognizes the following:

> "Within the past decade, a number of government agencies have adopted an innovative strategy to impose the use of evaluation evidence on local policy makers and practitioners. This "imposed use" (Weiss et al., 2005) requires applicants for federal program funds to show that the program they wish to run has been scientifically evaluated and found successful." (Weiss, 2008, p. 30)

Imposed use is a reality for the WWC, especially given the past incentives offered by the *Investing in Innovation Fund* (i3) to implement interventions that meet the WWC's standards of evidence (for an overview, see Boulay et al., 2018). I will discuss the issue of imposed use more thoroughly in Chapter 5. For now, it is enough to recognize that instances of imposed use make several assumptions in this theory of change (such as educators' perceptions of the WWC's relevance or credibility) irrelevant if they are, indeed, forced to use it.

Fourth and finally, this theory of change was based on a generic behavior change model created by Mayne (2015, 2019). Most interventions aim to change the behavior of their beneficiaries. The WWC, for example, intends to contribute to the IES's larger goal

of "increas[ing] use of research and data in education decision-making" (IES, n.d.-a)—which is, of course, a behavioral outcome. Its theory of change must account for that. Luckily, Mayne (2015) developed a generic behavior change model that is "intuitive and is based on a synthesis of empirical evidence on behaviour change" (p. 179). As such, I used this specific model as my template when developing the WWC's theory of change.

## Step 3: Gathering Existing Evidence

The third step in any contribution analysis involves the gathering and collating of evidence pertinent to each theory of change component—both its impact pathway and its underlying assumptions. This evidence is then used to assess the realization (or not) of each step in the impact pathway and the validity of each underlying assumption (Mayne, 2012b). Critically, this step is only the first iteration of data collection efforts. Data collection is meant to occur repeatedly until there is enough evidence to fully evaluate the theory of change. Step 5 is often thought of as the step when primary data collection occurs, whereas Step 3 is usually dedicated to the gathering of secondary (i.e., pre-existing) evidence. The secondary evidence relevant to this project—which came from (a) IES/WWC sources, (b) previous evaluations, (c) reviews of empirical literature, and (d) recent large-scale surveys—is described in the following sections.

### *IES/WWC Sources*

An array of evidence was available directly through the IES and WWC websites. Not only did these sources become key indicators of the WWC's activities and outputs, but they also provided insight into the types of metrics the IES considers important when publicizing the WWC and its functions.

**WWC Published Outputs.** The WWC publishes virtually all of its outputs—such as its individual study reviews, intervention reports, and practice guides—on its website. They also archive previous versions of reports and protocols that have since been updated. Oftentimes these publications are posted with contextualizing information, and as such, I was able to compute the following frequencies by counting the number, type, and publishing date of various outputs. All tallies were made in May 2020, and they are summarized in Table 3.1. They include:

***# of review protocols developed/updated over time.*** Review protocols are the formal documents used to guide every WWC review. They ensure that the reviews proceed systematically by defining certain parameters (e.g., outcomes, sample characteristics, design standards, etc.) to guide each review as it progresses (WWC, n.d.-e). These protocols continue to be developed as new reviews are initiated, and they are occasionally updated to better reflect the WWC's changing standards as well as changes in the literature. Both the development and updating of review protocols indicates that the WWC is functioning as intended; in other words, both are evidence of the WWC's activities occurring. To determine the rates at which these protocols were developed/updated over time, I examined each of the 44 protocols posted on the dedicated WWC webpage (WWC, n.d.-e) and logged each protocol's publication and/or update date in an Excel spreadsheet. From there, I was able to determine how many protocols had been published and/or reviewed during each year of the WWC's existence (see Table 3.1 for details).

***# of practice guides published/revised.*** Practice guides are also an important WWC output, especially for practitioners. These guides convert evidence into

recommendations for addressing challenges in their classrooms and schools. Similar to review protocols, previously published practice guides are occasionally revised. Thus, I used the publication/revision rate of WWC practice guides as an indicator of output production. I was able to determine these rates by logging the year at which each of the 24 guides were published and/or revised using information available on the "Search Publications" WWC webpage (WWC, n.d.-i). It should be noted that two practice guides, *Preventing Dropout in Secondary Schools* and *Effective Literacy and English Language Instruction for English Learners in the Elementary Grades*, had received substantive updates several years after their original publication date. However, because these guides are listed separately on the WWC website from their updated versions, they were considered separately here.

***# of intervention reports published over time.*** Intervention reports synthesize evidence from studies meeting the WWC standards and summarize that evidence in a short statement about an intervention's effects. The publication rate of these intervention reports was used as evidence of the WWC's output production. To determine these rates, I again extracted the publication year from each of the 593 intervention reports posted on their "Search Publications" WWC webpage (WWC, n.d.-i) and copied them to an Excel spreadsheet (see Table 3.1). The 69 intervention reports archived by the WWC were not considered because their updated versions were included in the 593 total.

# of ***instructional videos paired with practice guides.*** Many of the WWC practice guides are paired with supplemental resources, such as archived webinars, infographics, and supplemental reading. Recently, the WWC has also begun offering instructional videos showing WWC practice guide recommendations being implemented in the

classroom. These videos are especially relevant for teachers, many of whom are eager for demonstrations of research-based practices in action (e.g., Barton & Tindle, 2019). Given their potential appeal to teachers, I tallied the number of instructional videos offered alongside each of the WWC's practice guides.

***# of practitioners on practice guide development teams.*** Practice guides are developed by a team of at least six individuals. Although the majority of team members are "nationally-recognized experts" in a given subject matter, WWC Procedures Handbooks have, since Version 3, also stated that the development team includes "two practitioners who have backgrounds that allow them to offer guidance about implementation of the recommendations" (WWC, 2020d, p. C-2). Including practitioners on these teams is an important step towards building their credibility among users, as previous work suggests that educators are more likely to "buy-in" to research information when it is shared and/or tested by a fellow educator (e.g., Drill et al., 2012). Inclusion of practitioners in the practice guide teams could then, in turn, serve as an indicator of the WWC's potential relevance and/or credibility among beneficiaries. I explored the extent to which the inclusion of practitioners was occurring by investigating the composition of each practice guide development team, which is posted on the WWC website. Unfortunately, the WWC never operationalizes who they consider to be a "practitioner" when recruiting for their practice guide teams. For my purposes, I identified "practitioners" as individuals whose affiliations (which were also listed on the website) included a state-level education agency, district-level office, or specific school. I was unable to identify practitioners in the two protocols regarding postsecondary

interventions, as both experts and practitioners would have, in those cases, been affiliated with universities.

***Strength of evidence used to support practice guides.*** Additionally, the evidence used to support each recommendation in a WWC practice guide is rated as "minimal", "moderate", or "strong" (for an overview, see WWC, 2017b). Ratings of minimal indicate that there is "[n]o consistent evidence that demonstrates the practices' positive effects, because it has not been studied or there is weak or conflicting evidence of effectiveness" (WWC, 2017a). Rating of moderate indicate that there is "[a]mbiguous evidence that the improvement in student outcomes is the direct result of the practices or whether the findings can be replicated with a diverse population of students" (WWC, 2017a). Ratings of strong indicate that there is "Consistent evidence that the practices improve student outcomes for a diverse population of students" (WWC, 2017a). I examined each practice guide's recommendations and tallied their ratings to calculate the percentage of recommendations rated in each of the three categories across all published guides.

**WWC Data Extraction Tool.** As part of the WWC's commitment to transparency, it continuously publishes a downloadable dataset with information resulting from its individual study reviews (WWC, 2020b). This "extraction tool" includes data across three units of analysis: (a) information at the "finding" level, which includes data related to every finding featured in reviewed studies, (b) information at the "studies" level, which includes data related to the individual study reviews conducted by WWC reviewers, and (3) information at the "intervention report" level, which includes data related to each outcome reviewed for inclusion in the WWC's intervention reports. For

this project's purposes, only data from the "studies" and "intervention report" levels were analyzed. I downloaded the data on May 5th, 2020, and transferred it to Excel for cleaning. Upon downloading the dataset, users are provided with a codebook explaining each variable, which proved helpful when deciding which indicators would be pertinent to the contribution analysis.

While cleaning and interpreting the extractable data, several unique features of the dataset became clear—two of which are worth clarifying here. First, while examining data extracted at the "studies" level, I noticed that the number of cases in the dataset (i.e., 15,124) did not match the number of individual studies (i.e., 10,872) catalogued on the WWC website. This may be because, in the extractable data, each case seemed to represent the *act* of reviewing a study for a specific review protocol; it did not represent an individual study itself. Distinguishing the two is important when remembering that an individual study may be reviewed more than once (i.e., as part of multiple review protocols). As an example, several individual studies were reviewed for both the *Primary Mathematics and Secondary Mathematics* protocols. Given this, as well as the possibility that some studies were reviewed by the WWC but never entered into the extractable data, the discrepancy between the reviews of individual studies logged on the WWC website and the individual study reviews featured in the extractable data were reconcilable.

A second point of clarification is related to the "intervention report" level. Whereas the WWC website had 593 intervention reports catalogued, the number of cases in the "intervention reports" dataset totaled 811. This is because each case in the extractable data represents a specific "outcome" featured in an intervention report, and each WWC intervention report can feature multiple outcomes. For example, the

intervention report summarizing research on the *Knowledge is Power Program (KIPP)* charter school model reports effects related to five outcomes—namely, math achievement, language arts achievement, science achievement, social studies achievement, and students' progression into college. It should also be noted that, compared to the 593 reports listed on the website, only 537 unique interventions were included in the extractable data. The reason for this discrepancy is unknown.

Once the data were cleaned, I used the R statistical system (R Core Team, 2020), along with the RCommander graphical user interface package (Fox & Bouchet-Valat, 2020), to calculate several indicators of interest. The R coding syntax used to do so is included in Appendix A. Those indicators include:

***# of individual study reviews conducted/posted.*** Data were available to compute the number of individual study reviews the WWC has conducted/posted during each year of its existence. This computation could, in turn, be used as an indicator of the WWC's activities as well as its outputs; as recognized by the WWC (n.d.-g), individual study reviews are "the foundation of all WWC review products." To compute this indicator, I used data at the "Studies" level of analysis. The "ReviewedDate" and "Posting_Date" variables indicated the specific date the study was reviewed and/or posted, respectively. These variables were transformed so that they only included the review or posting year. Frequencies were then calculated on those variables to determine how many studies reviews were conducted each year. Unfortunately, no posting dates appeared to be logged prior to October of 2017.

***Size of backlog for individual study reviews.*** Given previous findings of an output backlog (GAO, 2010), I used these same variables to calculate the size of the

backlog between the time that individual study reviews were conducted and the time they were posted on the WWC website. This was done by transforming the dates listed in the "ReviewedDate" and "Posting_Date" variables into numeric objects and then calculating the difference between them.

***# of studies meeting WWC standards included in each intervention report outcome.*** To identify the number of studies summarized in each intervention report outcome, I generated both a histogram and a frequency table on the "NumStudiesMeetingStandards" variable, which is part of the "Intervention Report" level of analysis. This provided a number of studies eligible for review synthesized in each of the WWC's intervention report outcomes.

***Average sample size across intervention report outcomes.*** In the intervention-level data, there was also an indicator of sample size ("Sample_Size_Intervention") for each of the intervention report outcomes. I calculated both the mean and median of this variable for all entries with at least one eligible study.

***Demographic data from intervention report outcomes.*** The intervention-level data also included several dummy variables indicating the demographic variables considered as part of the outcome. Among these demographic characteristics are indicators of ethnicity, free/reduced lunch status, English language learners, gender, school type (i.e., public, private, charter, or parochial), and school setting (i.e., rural, suburban, and urban). Frequencies were conducted to determine the breakdown of these characteristics across all intervention report outcomes with at least one eligible study.

***Rating breakdown of individual studies.*** In the "Studies" dataset, each individual study is assigned a rating based on its adherence to WWC standards. Studies considered

as part of an intervention report's findings must be rated as "Meeting Standards Without Reservations" or "Meeting Standards with Reservations." Along with these ratings, studies in the dataset also had ratings such as "Does Not Meet WWC Standards" or "Ineligible for Review" or "Not Rated." The frequencies at which each of these ratings appeared in the dataset were calculated.

***Topic areas of individual study reviews.*** The "Studies" dataset also included dummy variables indicating each individual study review's relevance to the twelve WWC topic areas (e.g., Literacy, Charter Schools, Kindergarten to 12$^{th}$ Grade). To determine the number of study reviews relevant to each topic area, I sorted the dataset to only include studies meeting WWC standards (with or without reservations) and then conducted frequency analyses on the twelve dummy variables. Each study review could be categorized as belonging to more than one topic area.

***Outcome domains featured in intervention reports.*** The "Intervention Reports" dataset also includes a variable indicating the specific domain(s) (e.g., math achievement, credit accumulation, teacher retention) targeted by each intervention report outcome. This indicator was used to determine the types of outcomes receiving the most attention in WWC intervention reports.

***Protocols referenced in intervention reports / individual study reviews.*** In both the "Studies" and "Intervention Reports" datasets, a variable called "Protocol" indicates under what review protocol the study or intervention report outcome in question was reviewed. These protocols are related to certain topic and/or outcome domains. By calculating the frequency at which each review protocol was listed, I was able to further

understand which topics and/or outcome domains were most and least prevalent in the WWC's outputs.

***Effectiveness ratings of intervention report outcomes.*** Based on its review of the evidence, each intervention report outcome includes a rating indicating the extent to which the program or practice "worked" at improving that outcome. The "Intervention Reports" dataset includes a variable specifying these ratings for each of the intervention report outcomes with at least one eligible study. These ratings include "Positive Effects," "Potentially Positive Effects," "No Discernable Effects," "Potentially Negative Effects", and "Negative Effects." The frequency at which each of these effectiveness ratings appeared in the dataset was calculated as well.

**List of WWC Certified Reviewers.** The WWC maintains a public list of its certified reviewers (WWC, 2020a), which can be exported as an Excel file. I exported the data on October 20th, 2020. 284 individual reviewers were named in the dataset, along with 48 review organizations. The available data included reviewer contact information, organization affiliations (for individual reviewers), and a field specifying which WWC Review Standards each individual/organization was certified in. Unfortunately, the "version" number of each certification—which serves as a proxy indicator for how up-to-date each certification is (i.e., Version 1.0 of the Review Standards were published in May 2008 whereas Version 4.0 was published in October 2017)—was only available for individual reviewers, so only their data were analyzed. Frequency analyses and descriptive statistics were conducted in Excel.

**Year in Review Posts.** Since 2016, the WWC has published an annual update summarizing its accomplishments for the year (WWC, 2016b, 2017d, 2018, 2019). Links

to these summaries are often placed on the website's landing page. Unfortunately, the types of indicators included in these summaries (i.e., number of studies reviewed, number of reports published, number of webinars conducted, number of Help Desk questions answered, number of website visits, etc.) differs across years, which makes it challenging to assess changes over time. Even so, these posts were helpful because they hinted at the types of performance measures that the WWC finds valuable enough to report publicly.

**IES Biennial Reports to Congress.** As required by the *Education Sciences Reform Act* of 2002, IES is required to submit a biennial report to Congress that details its spending decisions as well as the activities of its four Centers. I reviewed these reports (IES, 2013, 2017, 2018a) per the suggestion of the WWC Help Desk Team, who responded to one of my inquiries with the following suggestion:

> "In 2018 a different team assumed control of the website and communication facets of the WWC. Since then we've been able to get more analytics data and have tracked visits/downloads more consistently. There are a few metrics from earlier years in the last IES Director's biennial report." (personal communication)

Although these reports only included a few useful metrics, they did help—like the Year in Review posts (e.g., WWC, 2019)—contextualize the types of WWC activities and outputs that matter to IES. These reports also provided some information on the WWC's funding levels over time.

**USAspending.gov.** The USAspending website is the U.S. Government's official source of spending data. I used this website to develop a more nuanced snapshot of WWC funding levels over time. Although IES is required to list all grants and contracts in its biennial reports, it only includes contracts in excess of $100,000. Furthermore, these reports fluctuate in how much detail they provide (e.g., the report prepared for years 2013-2014 and 2015-2016 does not provide information about each contract's purpose;

IES, 2017). It was therefore important to consult USAspending for additional details. To find all WWC-related contracts, I used the "Keyword Search" function to filter out all contracts with the keywords "What Works Clearinghouse" included. That search yielded 89 transactions from 24 distinct contracts; these transactions occurred between September 2004 to March 2020. I then used the "Award History" log included in each contract to determine the amount of money awarded during each year it was in operation. See Figure 4.2 for a visualization of funding levels over time.

### *Previous Evaluations*

To my knowledge, the WWC has only received one dedicated evaluation from an independent evaluator, and that was conducted a decade ago (i.e., GAO, 2010). Even so, that evaluation proved to be a rich source of evidence, as did a few others.

**Baldwin et al.'s Evaluation of IES (2008).** In 2007, the National Board for Education Sciences (NBES) commissioned an external evaluation of the "effectiveness of IES in carrying out its priorities and mission using primarily preexisting data sources" (p. iii). The evaluation contract was awarded to Synergy Enterprise Incorporated and its subcontractor, the Center for Evaluation and Education Policy (CEEP). The evaluation focused on the IES generally, however it did cover the WWC in some detail. Although, at that time, the WWC was only just beginning to produce outputs, the evaluation features several indicators useful for my theory of change, including findings from a survey of website visitors' professional roles and reasons for visiting.

**GAO Evaluation of WWC (2010).** In 2009, the Government Accountability Office undertook an evaluation of the WWC as required by the *Omnibus Appropriations Act* (2009). Specifically, the GAO examined (a) the rigor of the WWC's review process

relative to "accepted standards in research evaluation" (p. i), (b) the WWC's output production over time, and (c) the usefulness of the Clearinghouse's dissemination efforts. Several of the evaluation's findings were applicable to this contribution analysis, including survey results from state- and district-level officials about their knowledge and use of the WWC. Findings from a less thorough, but still informative survey conducted with educators and school-level administrators were available as well.

**WWC User Feedback Campaign (2016).** The WWC conducted their own user feedback campaign from 2014 to 2016. The campaign included in-person focus groups, virtual focus groups, and a pop-up feedback survey featured on the WWC's website. Findings from this internal evaluation are briefly summarized on the WWC's website (e.g., WWC, n.d.-f), which included (a) a demographic breakdown of respondents from the pop-up survey, and (b) a summary of focus group participants' opinions about the WWC's website. Unfortunately, few additional details were provided.

**Barton and Tindle (2019).** Most recently, the IES has partnered with the EdTech Evidence Exchange (formerly the Jefferson Education Exchange) to better understand educators' information needs and desires for the national educational research agenda. The EdTech Evidence Exchange is a nonprofit organization that aims to "help educators make better-informed decisions about education technology based on their specific instructional environments" (Barton & Tindle, 2019, p. 2). In October and November of 2018, EdTech Evidence Exchange and IES organized a series of "convenings" in order to better understand how educators engage with educational research. As part of the three convenings, a total of 1,297 educators were surveyed about their attitudes towards, usage of, and access to educational research. Some of the survey was actually modelled after

Lysenko et al.'s (2015) work, which is further detailed below. Uniquely, the survey also asked several questions about awareness and use of IES programs, including the WWC. Responses to these items were especially helpful given how recently they were collected.

In their report summarizing the results of the survey, Barton and Tindle (2019) differentiate between two groups of educators in their sample. The first is referred to as "association members," and they make up the bulk of the sample (i.e., 1,153 out of 1,297 educator; 89%). The second is referred to as "convening attendees", and they make up the remainder. Association members are educators who received the survey because of their affiliations with certain professional organizations. Convening attendees, on the other hand, were asked to complete the survey as part of their participation in one of the convenings. Both groups include teachers and administrators, but importantly, several characteristics differentiated the two groups. Specifically, convening attendees reported spending considerably more time with students (i.e., 70% reported that they spent 31-40 hours per week with students) when compared to association members (i.e., 48% reported spending 10 hours or less with students). In addition, 80% of convening attendees identified themselves as "PreK-12 teachers", whereas only 41% of association members identified themselves as such. Together these findings suggest that the convening attendee subsample may be treated as a proxy for classroom teachers, whereas the association member subsample may be a proxy for administrators. Although not perfect proxies, comparing these subgroups will be helpful in examining how the WWC's theory of change might be more or less defensible depending on the type of educator in question.

### *Reviews of Empirical Literature*

As recommended by Mayne (2008, 2012b), synthesizing findings from the research literature can serve as a useful evidence source during contribution analysis. Conveniently, there have already been several reviews of the literature on how educational research information is used in practice settings. Accordingly, I used the following reviews as evidence sources in this contribution analysis:

**Hemsley-Brown and Sharp (2003).** This work reviewed existing literature in both education and medicine about practitioners' use of research-evidence in practice, and moreover, what approaches to research dissemination facilitated use. The search included medical literature in response to Hargreaves' (1996) oft-cited remarks about differences in evidence use between the education and medical fields. The review examined English-language work published between 1988 and 2001. It identified 183 empirical, theoretical, and argument papers, but only 21 were deemed sufficiently pertinent to the review. Of those 21, only 6 were peer-reviewed articles related to educators' research use. Although this review is limited by its size and datedness, some of its emergent themes—including the importance of research/practitioner collaborations and the significance of organizational barriers/facilitators to research use—are reiterated in subsequent reviews. Additionally, since it reviewed literature published prior to 2001, it is especially applicable to the rationale assumptions underlying the WWC's creation.

**Dagenais, Lysenko, Abrami, Bernard, Ramde, and Janosz (2012).** The scope of this review overlaps with Hemsley-Brown and Sharp's (2003), though not completely. Specifically, it forsakes theoretical literature and commentaries to only review empirical research about (a) the use of research by educators, and (b) predictors of that use. They

examined all English- and French-language work published between 1990 and 2010. Notably, they excluded any work that examined educators' research use within the context of special education. Their review yielded 1,326 initial citations, with only 27 papers ultimately meeting their criteria. Ultimately, five determinants of research use were extracted—including localized research projects, high-quality communication between researchers and practitioners, practitioners' involvement in research, practitioners' teaching experience, and organizational support.

**Cain (2016).** This review acknowledged those conducted by Hemsley-Brown and Sharp (2003) and Dagenais et al. (2012) while also trying to extend them. Both its scope and approach were unique. With regard to the former, Cain mused that "what is largely absent from the debate is a consideration of what actually happens when teachers encounter research – what they access, how they make sense of it and what they do with it" (p. 619). As such, he focused on literature pertinent to those issues. With regard to his approach, Cain followed a narrative review technique (e.g., Hammersley, 2001a) in which he also acknowledged his own positionality (i.e., his concerns about how the "what works" movement in education is undermining teachers' values). Similar to previous reviews, Cain's review only yielded 32 articles meeting his review criteria. Based on this evidence, he concluded that few teachers actually engage with research, and those who do may be "using" it as a basis for reflection rather than as an instrument to guide decision-making.

**Schaik, Volman, Admiraal, and Schenke (2018).** This review conducted by Schaik and colleagues focused on literature explicating the "barriers and conditions for teachers' academic knowledge utilisation" (p. 50). Its search criteria included all peer-

reviewed English-language publications published between 2001 and 2016, which yielded 447 papers for review. 66 were deemed worthy of analysis. Analysis included the construction of a framework to organize the various barriers and conditions identified, which were organized into four levels: (a) the research knowledge level, (b) the communication level, (c) the school organization level, and (d) the individual teacher level. They concluded that many of the same barriers identified by Hemsley-Brown and Sharp (2003) are still being recognized in the contemporary literature. Their conclusion is also similar, in that they propose partnerships between schools and research institutes as a possible way of circumventing these barriers.

**Gorard, See, and Siddiqui (2020).** This review is currently awaiting publication in *Review of Education*, though an "early view" version was acquired through interlibrary loan. Gorard et al.'s review is unique in that is places very strict methodological requirements on the studies included in its review. Specifically, it only examined studies with counterfactual designs so that causal claims can be made about how educational research is translated into use. 33 studies were considered to be of acceptable quality.

### *Recent Large-Scale Surveys*

Several large-scale surveys of educators' research use have been conducted in recent years. Since several of these surveys never made their way into subsequent literature reviews (e.g., Schaik et al., 2018; Gorard et al., 2020), findings from these surveys were also treated as their own evidence source.

**Penuel et al. (2017).** This survey was conducted by the National Center for Research in Policy and Practice, an IES-funded research center. The survey underwent an extensive development process before it was administered, which involved psychometric

testing and a series of forty cognitive interviews with educators. Respondents were a nationally representative sample of 773 school- and district-level administrators from 45 U.S. states. 485 different school districts were represented. Survey items asked about respondents' attitudes towards educational research, the sources they used to access it, and how they used it in practice. One item asked about their specific use of IES resources as research access points, including the WWC. This item will be especially useful given that similar estimates of WWC use, such as those offered by Baldwin et al. (2008) and the Government Accountability Office (2010), are outdated.

**Lysenko and colleagues (2014, 2016).** Lysenko, Abrami, Bernard, Dagenais, and Janosz (2014, 2016) developed and administered the Questionnaire about the Use of Research-Based Information (QURBI) to better understand how educators, specifically, use research-based information. The 2016 study surveyed 1,153 Canadian educators. Along with the 44 Likert-type items making up the QURBI, 6 open-ended questions were asked as well. The 2014 study was larger, and analyses were presented in greater detail. Of the 2,425 Canadian educators studied, most (1,979; 82%) identified as teachers. Even so, Lysenko et al. (2014) parsed analyses out based on educators' professional roles, such that teachers' responses could be compared to the responses of administrators and education professionals (e.g., school psychologists).

**Step 4: Assemble and Assess the Contribution Claim**

In Step 4, the evidence gathered in the previous step was used to generate a *contribution claim*. As its name implies, a contribution claim is a conclusion about whether or not an intervention has realized its intended contribution. Yet, as recently put forward by Mayne (2019), the most useful contribution claims are more than a simple

yes/no conclusion about an intervention's impact. They include claims about "how" and "why" an intervention made a meaningful difference, as well as for "whom" (p. 175). Together these claims are assembled from a larger *contribution story,* which walks step-by-step through an intervention's theory of change to determine if, based on the evidence gathered in Step 3, its component and/or assumptions that are being realized (or not). The telling of this contribution story, as well as the contribution claim(s) that can be assembled from it, are the principal results of a contribution analysis; thus, they are addressed Chapter 4.

**Step 5: Gathering Additional Evidence**

As suggested earlier, contribution analysis is most successful when it proceeds as "an iterative process" (Mayne, 2008, p. 3). So, after assembling an initial story of an evaluand's contribution, it is recommended that the evaluator consider additional evidence sources that might enhance the credibility of the contribution claims (Mayne, 2008, 2012b, 2019). This usually means collecting original data, since existing evidence should have been considered in Step 4. Once additional evidence is considered, an updated contribution story can be presented in Step 6. In Step 5, evidence sources should also be chosen based on any emergent holes in the contribution story. In other words, if certain claims could not be made or were made according to minimal evidence, this second round of data collection should prioritize evidence that helps fill those holes. Accordingly, in Step 5, I identified two additional sources of evidence that would bolster the contribution story. The first was transcripts of previously conducted focus groups by a WWC contractor. The second involved developing and administering a survey to further

understand how preservice teachers engage with the WWC and its assumptions. Both are detailed below.

### *WWC Focus Group Transcripts*

**Background.** Time constraints did not allow for conducting my own interviews or focus groups with WWC users; however, I was able to secure existing focus group data from another source and analyze that data for themes related to the WWC's theory of change. Specifically, in 2016, both virtual and in-person focus groups were conducted by the Mathematica Policy Institute (one of the primary WWC contractors between 2007 and 2017) as part of a user feedback campaign. In an internal memo, Mathematica described the purpose of the feedback campaign as follows:

> "In 2015 and 2016, the What Works Clearinghouse (WWC) will release new products on the WWC website including videos, practice guide summaries, special features pages, and the redesigned Find What Works tool. This new content is tailored to specific stakeholders, especially practitioners, with a focus on explaining how to use WWC products when making key decisions in education. Mathematica developed a plan for obtaining feedback on these products' utility, customer satisfaction, and ideas for improvements."

I was aware of these focus groups because of a post on the WWC website summarizing select findings from its user feedback campaign (WWC, n.d.-f). Knowing that additional analysis of these data could assist with my contribution analysis, I submitted a Freedom of Information Act request to the U.S. Department of Education in March 2020. In return, I received anonymized focus group data and accompanying memos between Mathematica and the Institute of Education Sciences (IES).

#### Virtual Focus Groups.

***Participants.*** To recruit participants for their virtual focus groups, Mathematica used preexisting contact networks maintained by the WWC/IES. Specifically, emails

were sent to potential respondents using contact information gathered through: (1) the WWC email listserv, (2) WWC Help Desk inquiries, (3) WWC blogs/articles, (4) visits to WWC-sponsored conference/event booths, (5) WWC webinars, and (6) IES email lists. Announcements were also posted on social media and the WWC's website. 215 individuals expressed some interest in participating in the focus groups, but not all were included. Instead, a process of prioritization and stratified random sampling occurred. For example, Mathematica prioritized the recruitment of respondents who identified as teachers, administrators, state/district-level education officials, and to a lesser extent, researchers. They also deliberately selected their eventual focus group participants using a sampling procedure ensuring that participants from a diversity of geographic regions were represented. These efforts culminated in four groups of stakeholders—yielding 28 total participants. Characteristics of the four groups are included in Table 3.2.

***Procedure.*** Each stakeholder group participated in their own discussion, which was facilitated by a Mathematica employee. The focus groups were conducted using an asynchronous virtual focus group platform called QualBoard®. The platform allowed for three-day long conversations in which participants proceeded through seven distinct modules at their own pace; participants could ask questions and respond to each other's comments. Module 1 asked about participants' background knowledge of the WWC and their engagement with it to date. Modules 2-5 introduced participants to newly designed WWC products and asked for feedback. Module 6 asked participants for ideas about how the WWC could further disseminate educational research and who those dissemination efforts should be catered to. Module 7 concluded the focus group. Upon completion,

respondents received a $30 Amazon gift card as compensation. For a breakdown of the questions asked in the virtual focus group, see Appendix B.

**In-Person Focus Groups.**

***Participants.*** Mathematica also facilitated four in-person focus groups with researchers, practitioners, and other stakeholders associated with Regional Education Laboratories (RELs) across the country (Mathematica, 2016). 25 participants were recruited through their respective RELs. Unfortunately, far less information was provided about these participants in the documents released by the U.S. Department of Education, so a breakdown of their demographic characteristics could not be generated.

***Procedure.*** Focus groups were facilitated by one or two Mathematica employees. Each conversation lasted approximately 60 minutes and proceeded similarly to the virtual focus groups. After overviewing the consent document, facilitators asked each participant to introduce themselves and their "day job(s)." Next, participants were asked about their familiarity with the WWC and their perceptions about its intended audiences. After that, participants were introduced to new WWC products and asked for feedback. Finally, before concluding the focus group, the facilitator asked participants to comment on the WWC's existing dissemination practices. A copy of the in-person focus group protocol is also available in Appendix B.

**Analysis.** Transcripts of all eight focus groups were analyzed using a deductive coding procedure (Hayes, 1997; Hsieh & Shannon, 2005). In a deductive thematic analysis, researchers utilize existing theory to identify a pre-determined coding schema or "codebook." That codebook is then used to sort excerpts of qualitative data into categories from which themes can be extracted and rationalized using exemplars. This

approach differs from a purely inductive one (e.g., Glaser & Strauss, 2017) because the codebook is developed prior to—rather than during—data analysis. Although inductive approaches are "conventional" in instances of content analysis (Hsieh & Shannon, 2005), deductive approaches are helpful in evaluation contexts where a logic model or theory of change already exists (Hayes, 1997). Given the theory-driven nature of this evaluation, as well as the fact that a theory of change was explicated in Step 3, a deductive approach was deemed suitable for my purposes. It also conforms with methodological decisions made in studies similar to this one (e.g., Neal et al., 2018).

Accordingly, before coding began, an a priori codebook was developed to directly align with certain components of the WWC's theory of change. As a reminder, Steps 5 and 6 of contribution analysis are intended to strengthen the contribution story developed in Step 4. Thus, special attention was paid to evidence related to the WWC's causal assumptions, as these were implicated in the contribution story (see Chapter 4) as a reason why the WWC may be failing to achieve its desired impact. Likewise, the codebook was also designed to focus on teachers' experiences—a decision which is rationalized in Chapter 4 as well. A copy of the codebook is included in Table 3.3. During analysis, focus group excerpts were categorized according to one or more of the available codes. All excerpts within each category were then reexamined to identify common themes among them. Those themes, along with example excerpts, are presented in Step 6, which describes how the extant contribution story was updated in light of new evidence.

### *Preservice Teacher Survey*

**Background.** To better understand early-career teachers' engagement with the WWC and its underlying assumptions, I developed a short survey. This survey was unique for two reasons. First, it did not inform teachers of the project's true purpose until after completion. Previous surveys of educators' WWC engagement (e.g., GAO, 2010) have made respondents aware of its evaluative purpose, which could have, in turn, allowed for social desirability effects. In this survey, respondents were told that the survey inquired into their engagement with educational research generally, not with the WWC specifically. All respondents who finished the survey were informed of its true purpose and given an opportunity to withhold their data. A second unique feature of this survey was that it targeted preservice teachers—a subgroup underrepresented in previous investigations into educators' research use and use of the WWC specifically. These unique features ensured that the survey would provide additional evidence for my contribution story while also making a novel scholarly contribution.

**Participants.** Survey data were collected from completers of a Teacher Education Program (TEP) at a large, public university in the southeastern United States. Completers refer to students who had completed coursework but were still participating in their teaching practicum/seminar. To recruit participants, I used a purposive sampling strategy that piggybacked onto a larger project conducted by the university's College of Education. Specifically, because part of the College's project already involved the administration of an annual student survey for accreditation and quality improvement purposes, I was granted permission to add additional items specific to my project.

Importantly, I received approval from both the College of Education, as well as the university's institutional review board, before adding my items to the existing survey.

105 respondents approved of having their data analyzed for research purposes. Respondents mostly identified as female ($n$ = 80; 76.2%), White ($n$ = 89; 84.8%), and continuing-generation college students ($n$ = 81, 77.1%). Over half ($n$ = 57, 54.3%) were born and educated in the same state that the university was located in. Respondents' demographics, though somewhat homogenous, were representative of the College of Education's more generally (University of Kentucky, 2020).

**Instrumentation.** To assess their engagement with the What Works Clearinghouse (WWC) and its underlying assumptions, participants were asked to complete three blocks of survey questions. The first block asked respondents to describe how they understood the terms "research" and "evidence" in the context of educational decision-making. Respondents were provided with a text box to enter their descriptions. The second block included twelve closed-response items about respondents' perspectives about educational research, which were developed to align with various causal assumptions in the WWC's theory of change (see Table 3.4). The items were framed as statements, and participants were asked to rate how much they agreed or disagreed with each statement on a four-point Likert scale. The third block of questions included seven questions about respondents' engagement with the WWC. Specifically, they asked about respondents' familiarity with and use of the WWC, as well as the degree of this familiarity and use.

Great care was taken to ensure the integrity of these items—both in terms of their alignment to my contribution analysis as well as their accordance with best practices in

survey design. To ensure alignment, I developed a survey matrix as suggested by Sampson, Nelson, & Bradley (2021). Put simply, a survey matrix "outlines the survey and allows the developers to document evidence of its validity before administration" (Sampson et al., 2021, p. 20). A portion of this matrix, which shows how each item aligns with one or more of the WWC's causal assumptions and any applicable literature, is included in Table 3.4. Survey items, along with response frequencies for each, are included in Tables 3.5-3.7. I also pilot-tested my survey with graduate students enrolled in a Survey Methods course. Those students—who were learning about survey design best practices at the time of the pilot—were asked to critique the survey items for issues with language and bias. Items were revised accordingly.

**Procedure.** Participants were emailed a survey invitation in Spring 2020. Upon completion of the accreditation survey, participants were then invited to complete an additional set of questions for research purposes. Instructions clarified that their participation was strictly voluntary, and moreover, that their responses would remain anonymous. For those choosing to complete the additional survey items, they received a debriefing statement upon completion. Along with discussing the purpose of my project, this statement included a link to the What Works Clearinghouse website so that interested respondents could learn more.

**Analysis Plan.** Results from this survey were intended to clarify the contribution story built from my first set of evidence. As such, analysis largely focused on building a descriptive understanding of teachers' engagement with the WWC and its underlying assumptions.

***Block 1.*** Data from the open-ended responses in Block 1 were analyzed according to Braun and Clarke's (2006) guidelines for thematic analysis. Specifically, after becoming acquainted with the data, I identified themes and subthemes emerging across the responses. This approach has been used in other studies of teachers' research engagement (e.g., Joram et al., 2020). I have chosen to use an inductive coding procedure for these responses because the topics addressed (i.e., respondents' conceptualization of "research" and "evidence") were not implicated in the contribution question or the theory of change. Thus, a more exploratory approach was justified. Though not related to the WWC's causal assumptions, these responses were informative because of their pertinence to one of the WWC's rationale assumptions. That assumption, which has been considered elsewhere (e.g., Joram, 2007; Biesta, 2010; Tseng, 2012), is that educators' conceptualizations of "research" and "evidence" are congruent with the WWC's own.

**Block 2.** Data from each of the twelve, Likert-type items in Block 2 were examined using frequency analyses conducted using R (R Core Team, 2020). Attention was paid to how responses looked different based on the causal assumptions addressed. For example, literature (e.g., Hemsley-Brown & Sharp, 2003; Cain, 2016) suggests that items 1-3, which target rationale assumptions, might receive more agreeable ratings than items 4-12, which target causal assumptions. Comparing frequencies across the items was important to both corroborating the contribution story as well as developing specific recommendations for improving the WWC.

**Block 3.** Data from Block 3 were examined using frequency analyses as well. Importantly, this is how previous work (e.g., GAO, 2010; Penuel et al., 2017, Barton & Tindle, 2019) has analyzed respondents' engagement with the WWC, and as such,

analyzing my data similarly allowed for comparison between my findings and the extant evidence. The open-response items concluding this block required coding. To do so, I utilized the same deductive procedure described above. The a priori codebooks for these items are specified in Tables 3.8 and 3.9, respectively. In each codebook, the codes were deliberately aligned with certain assumptions articulated in the WWC's theory of change.

**Step 6: Updating the Contribution Story**

Finally, Step 6 involved a revision to the contribution story developed in Step 4. That revision was made according to the data sources examined in Step 5. Both this revision, as well as the original contribution claims, are detailed in the next chapter.

## Chapter 3 Tables and Figures

Table 3.1 - Counts of Indicators Over the WWC's Lifespan

| Indicator | 2003 | '04 | '05 | '06 | '07 | '08 | '09 | '10 | '11 | '12 | '13 | '14 | '15 | '16 | '17 | '18 | '19 | '20 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Study reviews conducted | 0 | 6 | 188 | 103 | 1146 | 610 | 937 | 2373 | 184 | 1426 | 1625 | 607 | 991 | 3044 | 1596 | 122 | 127 | 39 |
| Study reviews posted | - | - | - | - | - | - | - | - | - | - | - | - | - | - | 141 | 436 | 56 | 127 |
| Review protocols developed | 0 | 0 | 1 | 5 | 1 | 1 | 5 | 4 | 4 | 2 | 2 | 6 | 1 | 6 | 0 | 1 | 5 | 0 |
| Review protocols updated | 0 | 0 | 0 | 0 | 0 | 0 | 6 | 2 | 0 | 9 | 1 | 4 | 6 | 2 | 4 | 2 | 6 | 2 |
| Intervention reports developed | 0 | 0 | 55 | 24 | 303 | 13 | 17 | 36 | 6 | 25 | 27 | 8 | 15 | 27 | 16 | 9 | 4 | 8 |
| Practice guides developed | 0 | 0 | 0 | 0 | 3 | 4 | 5 | 2 | 0 | 2 | 1 | 1 | 1 | 3 | 1 | 0 | 1 | 0 |
| Practice guides revised | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 2 | 3 | 0 |

*Note.* Gaps in counts for the "Study reviews posted" indicator were due to the unavailability of data for those years. Specifically, entries included in the "Posting_Date" variable of the WWC extractable data did not begin until October 2017.

Table 3.2 - Participant Characteristics from the Virtual Focus Groups

| | Group 1 (Researchers & Academics) | Group 2 (Teachers) | Group 3 (School/District Leaders) | Group 4 (State Admin & Parents) |
|---|---|---|---|---|
| **# of Participants** | 10 | 6 | 4 | 8 |
| **Professional Position** | | | | |
| Teacher (K-12) | 1 | 6 | 2 | 1 |
| Researcher | 8 | 1 | 0 | 0 |
| Faculty (Higher Ed) | 4 | 0 | 0 | 2 |
| Postsecondary Admin | 0 | 0 | 0 | 3 |
| State-Level Admin | 0 | 0 | 0 | 3 |
| District-Level Admin | 0 | 0 | 2 | 0 |
| School-Level Admin | 0 | 0 | 1 | 1 |
| Other School Staff | 0 | 0 | 1 | 0 |
| Parents | 0 | 0 | 0 | 3 |
| **Familiarity with WWC** | | | | |
| Familiar | 8 | 2 | 2 | 6 |

Table 3.2 (continued) - Participant Characteristics from the Virtual Focus Groups

| | | | | |
|---|---|---|---|---|
| Somewhat Familiar | 2 | 3 | 2 | 2 |
| Not Familiar | 0 | 1 | 0 | 0 |
| **Geographic Region** | | | | |
| Midwest | 3 | 3 | 0 | 1 |
| South | 2 | 2 | 1 | 1 |
| Northeast | 2 | 0 | 1 | 2 |
| West | 1 | 1 | 1 | 2 |
| Unknown | 2 | 0 | 1 | 2 |

*Note.* Focus group participants could select multiple categories when reporting their "Professional Position", which is why the summed number of selected categories is often greater than the total number of participants in each focus group.

Table 3.3 - A Priori Codebook for Focus Group Data

| Assumption | Code | Definition |
|---|---|---|
| Beneficiaries are aware of outputs. | Awareness | This code was used when the respondent referenced teachers' awareness of the WWC—either an abundance of awareness and/or a lack of awareness. This code was also appropriate when the respondent offers suggestions for expanding educators' awareness of the WWC. |
| | Audience | This code was used when the respondent commented on the WWC's current audience—either suggesting that it is sufficient or that it doesn't include all relevant groups. This code was also appropriate when respondents offered suggestions for expanding the WWC's audience, especially among teachers. |
| Beneficiaries are willing to seek them out. | Research-seeking attitudes/behaviors | This code was used when the respondent alluded to teachers' research-seeking practices—that is, whether (or not) they seek research to inform their practices in the classroom. This code was also appropriate when teachers' attitudes about educational research were addressed. |
| | Format needs/preferences | This code was appropriate when the respondent alluded to educators' resource needs and/or format preferences within the context of the WWC. They might have indicated that the WWC's current dissemination formats were appropriate, or that they could be expanded to better meet educators' needs. This code was also appropriate when the respondent suggests ways in which the WWC could better respond to educators' format preferences. |
| | Access points | This code was used when the access points educators use to discover research-related resources were mentioned. It is also used when respondents suggest other access points that could be useful for dissemination. |

Table 3.3 (continued) - A Priori Codebook for Focus Group Data

| | | |
|---|---|---|
| Beneficiaries are able to seek them out. | Barriers to access | This code was used when the respondent refers to conditions external of the WWC that impede educators' access to the WWC. |
| Resources are perceived as relevant. | Timeliness of resources | This code was used when the timeliness of WWC resources are addressed. They may either be perceived as current or as lagging behind the state of the field. |
| | Application to practice | This code was used when the WWC and/or its resources were said to be actionable in a classroom context or inapplicable to practice. |
| | Information needs/preferences | This code was used when the respondent alluded to educators' information needs in the context of the WWC. They might indicate that the WWC is meeting these needs, or that other types of information might be needed. This code was also appropriate when the respondent suggested ways in which the WWC could better respond to educators' information needs. |
| Resources are perceived as trustworthy | Perceptions of validity/reliability | This code was used when teachers' trust of the WWC (and research featured therein) was mentioned. |
| Instrumental use is possible. | Forms of use | This code was used when teachers' research use was mentioned by teachers themselves or by their colleagues. Forms of use may include instrumental use as well as other types. |
| | Barriers to use | This code was used when barriers to teachers' research use were mentioned, including (but not limited to) their lack of agency or difficulties with implementing evidence-based practices. |

Table 3.4 - Survey Matrix for Block 2 Items

| ITEM | | ALIGNMENT | | RATIONALE | |
|---|---|---|---|---|---|
| # | Text | Component | Subcomponent | WWC/IES | Relevant Literature |
| 1 | I find most educational research to be irrelevant to practice. | Diagnostic Assumptions | The supply of research is itself deficient. | Whitehurst, 2003 | Borg, 2009; Williams and Coles, 2007a; |
| 2 | It is challenging for me to determine whether or not research is of high-quality. | | Educators are poorly prepared to consume research. | WWC, 2020d | Williams and Coles, 2007a; 2007b |
| 3 | I rarely come across research findings that are presented in an easy-to-understand way. | | The supply of research is itself deficient. | WWC, 2020d | Williams and Coles, 2007a; 2007b |
| 4 | The most relevant research identifies if an intervention "works", as opposed to "how" or "why" it works. | Capacity Change Assumptions | Resources are perceived as relevant. (scope) | Whitehurst, 2003 | Harkin, 2005 |
| 5 | Researching the desired benefits of an intervention is more important than researching its unintended consequences. | | Resources are perceived as relevant. (scope) | WWC glossary entry for "outcomes" | Le Fevre, 2014; Pareja Roblin et al., 2018 |
| 6 | Interventions that improve students' test scores matter more than interventions focusing on other outcomes. | | Resources are perceived as relevant. (topic areas) | WWC *Find What Works* filter | Everton, Galton, & Pell, 2010; Ion and Iucu, 2014 |

Table 3.4 (continued) - Survey Matrix for Block 2 Items

| | | | | | |
|---|---|---|---|---|---|
| 7 | When determining if an intervention works or not, researchers should collect quantitative data--not qualitative data. | Capacity Change Assumptions | Resources are perceived as trustworthy. (data type) | WWC, 2020d | Kennedy, 1999; Naiz, 2009 |
| 8 | Controlled experiments--as opposed to other types of research--are the best way to determine if an intervention works or not. | | Resources are perceived as trustworthy. (design type) | WWC, 2020d | Cowen et al., 2017 |
| 9 | When I judge the quality of a research study, I care more about its methods and design than the context in which it was conducted. | | Resources are perceived as trustworthy. (context) | WWC, 2020d | Dagenais et al., 2012 |
| 10 | The best way to share research with educators is by publishing free, easy-to-understand reports. | Reach Assumptions | Beneficiaries are willing to seek outputs out. | WWC, 2020d | Levin, 1993; Cooper, 2012 |
| 11 | Any educator who wants to can implement an evidence-based practice in their classroom. | Behavior Change Assumptions | Instrumental use is possible. (organizational barriers) | Implicit | Hammersley, 2005 |
| 12 | It is reasonable to ask educators to implement research-based practices without altering them. | | Instrumental use is possible. (fidelity) | Whitehurst, 2004 | Biesta, 2007 |

Table 3.5 - Response Frequencies for Survey Block 1

***Were you born and primarily educated (kindergarten through high school) in Kentucky?***

| | Count | Percentage |
|---|---|---|
| Yes | 57 | 54% |
| No | 48 | 46% |

***What is your race/ethnicity?***

| | Count | Percentage |
|---|---|---|
| Asian | 1 | 1% |
| Black / African American | 2 | 2% |
| Hispanic / Latinx | 2 | 2% |
| White | 89 | 85% |
| Other | 3 | 3% |
| More than One Race | 8 | 7% |

***What is your gender?***

| | Count | Percentage |
|---|---|---|
| Female | 80 | 76% |
| Gender Non-Conforming | 1 | 1% |
| Male | 24 | 23% |

***Are you considered a first-generation student?***

| | Count | Percentage |
|---|---|---|
| Yes | 24 | 23% |
| No | 81 | 77% |

Table 3.6 - Response Frequencies for Survey Block 2

| | Strongly disagree | Disagree | Agree | Strongly agree |
|---|---|---|---|---|
| I find most educational research to be irrelevant to educational practice. | 20 (20%) | 62 (61%) | 15 (15%) | 4 (4%) |
| It is challenging for me to determine whether or not research is of high-quality. | 12 (12%) | 56 (55%) | 29 (29%) | 4 (4%) |
| I rarely come across research findings that are presented in an easy-to-understand way. | 1 (1%) | 46 (47%) | 43 (43%) | 9 (9%) |
| I find educational research to be most relevant when it studies if an intervention "works", as opposed to "how" or "why" it works. | 8 (8%) | 50 (50%) | 36 (36%) | 6 (6%) |
| Researching the desired benefits of an intervention is more important than researching its unintended consequences. | 9 (9%) | 63 (63%) | 25 (25%) | 3 (3%) |
| Interventions that improve students' test scores matter more than interventions focusing on other outcomes. | 31 (31%) | 46 (46%) | 17 (17%) | 6 (6%) |
| When determining if an intervention works or not, researchers should collect quantitative data--not qualitative data. | 7 (7%) | 45 (46%) | 40 (40%) | 7 (7%) |
| Controlled experiments--as opposed to other types of research--are the best way to determine if an intervention works or not. | 4 (4%) | 39 (39%) | 46 (46%) | 11 (11%) |
| When I judge the quality of a research study, I care more about its methods and design than the context in which it was conducted. | 3 (3%) | 53 (53%) | 39 (39%) | 5 (5%) |
| The best way to share research with educators is by publishing free, easy-to-understand reports. | 0 (0%) | 9 (8%) | 57 (56%) | 36 (36%) |
| Any educator who wants to can implement an evidence-based practice in their classroom. | 0 (0%) | 14 (14%) | 65 (64%) | 22 (22%) |
| It is reasonable to ask educators to implement research-based practices without altering them. | 8 (8%) | 43 (43%) | 44 (44%) | 4 (4%) |

*Note.* Not every respondent answered every item.

Table 3.7 - Response Frequencies for Survey Block 3

***Have you ever heard of the What Works Clearinghouse (WWC)?***

| | Count | Percentage |
|---|---|---|
| Yes | 21 | 20% |
| No | 84 | 80% |

***Have you ever used the What Works Clearinghouse (WWC)?***

| | Count | Percentage |
|---|---|---|
| Yes | 16 | 76% |
| No | 5 | 24% |

***How frequently do you use the What Works Clearinghouse (WWC)?***

| | Count | Percentage |
|---|---|---|
| Less than once a year | 1 | 6% |
| Once a year | 4 | 25% |
| 2-3 times a year | 4 | 25% |
| Once a month | 5 | 31% |
| 2-3 times a month | 2 | 13% |
| Once a week | 0 | 0% |
| 2-3 times a week | 0 | 0% |

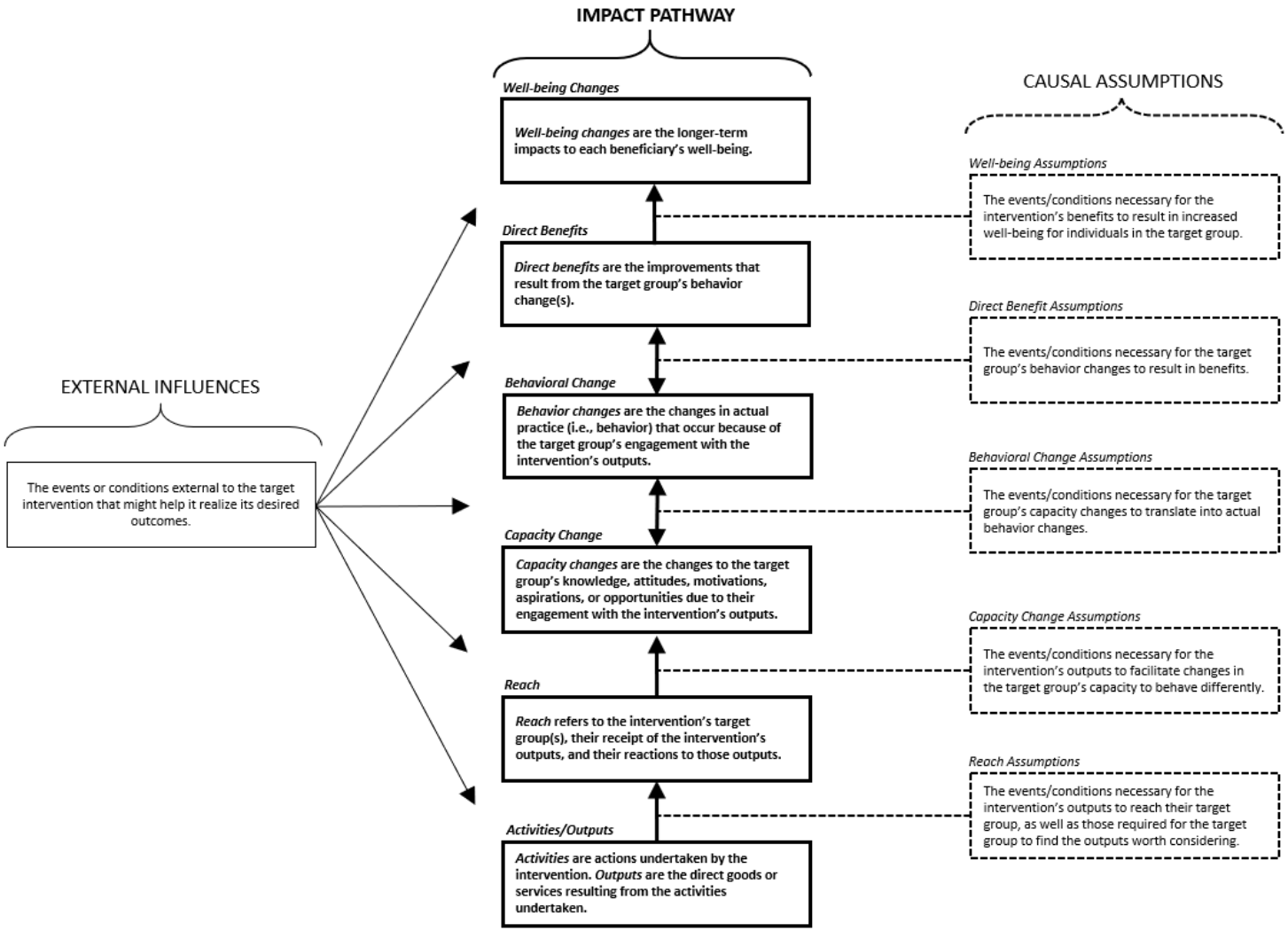
Table 3.8 - Codebook for Item 22 Regarding Why Respondents Have Heard of the WWC but Not Used It

| Assumption | Code | Definition |
|---|---|---|
| Information/resource is perceived as relevant. | Relevance of information | This code was used when the respondent suggested that the WWC disseminates information that is irrelevant. This might be due to its scope, the standards of evidence through which it is vetted, or the topic areas that it addresses. |
| | Relevance of resource | This code was used when the respondent suggested the manner through which the WWC disseminates information is irrelevant. This might be because it is perceived as inconclusive, inapplicable to practice, or less relevant than other types of research access points. |
| Information/resource is perceived as trustworthy. | Trustworthiness of information | This code was used when the respondent suggested that they were skeptical of the information disseminated by the WWC. This skepticism might be because the information doesn't meet certain standards of rigor, doesn't account for student and/or contextual characteristics, or does not align with respondents' own experiences as an educator. |
| | Trustworthiness of resource | This code was used when the respondent suggested that they were skeptical of the dissemination method itself. This might be because researchers, rather than other teachers, seem to be sharing the information. |
| Instrumental use is possible. | Barriers to instrumental use | This code was used when the respondent alluded to barriers that prevent them from using research even when they are motivated to do so. These barriers might include certain organizational characteristics that discourage research use or make it difficult to do so. |

Table 3.9 - Codebook for Item 23 Regarding If Respondents Use the WWC as Their Primary Source of Research Information

| Assumption | Code | Definition |
|---|---|---|
| Information/resource is perceived as relevant. | Relevance of information | This code was used when the respondent suggested that WWC is not their primary source of research because the information it disseminates is irrelevant. This might be due to its scope, the standards of evidence through which it is vetted, or the topic areas that it addresses. |
| | Relevance of resource | This code was used when the respondent suggested that the WWC is not their primary source of research because its dissemination method itself is deemed irrelevant. This might be because it is perceived as inconclusive, inapplicable to practice, or less relevant than other types of research access points. |
| Information/resource is perceived as trustworthy. | Trustworthiness of information | This code was used when the respondent suggested that the WWC is not their primary source of research because they are skeptical of the information it features. This skepticism might be because the information doesn't meet certain standards of rigor, doesn't account for student and/or contextual characteristics, or does not align with respondents' own experiences as an educator. |
| | Trustworthiness of resource | This code was used when the respondent suggested that the WWC is not their primary source of research because they are skeptical of the dissemination method itself. This might be because researchers, rather than other teachers, seem to be sharing the information. |

Figure 3.1 – Depiction of Mayne's Generic Theory of Change Model for Behavior. Adapted from Mayne (2015, 2017, 2019)



# CHAPTER 4 – RESULTS

When an evaluation adheres to Mayne's (2008, 2012, 2019) six steps of contribution analysis, it systematically builds an evidence-based argument about the impact of an evaluand. Accordingly, in this chapter, I show how insights about the WWC's contribution were built over the course of this analysis—not just at its finale. I do this by discussing each step in turn. For example, in Step 1, I demonstrate how the contribution analysis resulted in an articulation of the WWC's intended contribution. In Step 2, I demonstrate how the contribution analysis resulted in a verifiable theory of change, one that explicates how the WWC seeks to make its contribution. In Step 4, I demonstrate how the contribution analysis scrutinized that theory of change by developing a contribution story, and in turn, assembling interim contribution claims. And finally, in Step 6, I demonstrate how the analysis of additional evidence resulted in a revised contribution story and strengthened contribution claim.

## Step 1: Establish the Specific Cause/Effect Issue to Be Addressed

Prior to articulating a program's theory of change, it is important to closely examine the specific contribution sought by the program, as well as the expected size of that contribution (Mayne, 2011, 2012b). Once the expected contribution is identified, it serves as the capstone towards which the program's theory of change strives.

### Identifying the Nature of the WWC's Expected Contribution

#### *The IES's Original Conceptualization of the WWC*

My first goal was to understand how IES originally conceived of the WWC and its intended contributions. To do so, I examined how the IES's founding director, Grover Whitehurst, initially discussed the WWC in congressional hearings, interviews, and

lectures (see Table 4.1). Several themes are reiterated by Whitehurst across these data sources. The first, which was most explicit in the earlier congressional hearings, is his belief that the WWC represents a novel government-sponsored dissemination strategy for educational research. In his testimony to the House Subcommittee on Education Reform, Whitehurst draws a divide between extant government-sponsored dissemination modes (i.e., the Education Research and Information Clearinghouse; ERIC) and his vision of the WWC:

> "Again, that is exactly our intent with the What Works Clearinghouse effort. Currently, at least for some topics, and school size, the one you mentioned, is one of those topics with a fairly large amount of literature. If you go to our current dissemination effort, which is the Education Research and Information Clearinghouse, ERIC, and click on class size, the problem is that you will generate hundreds of hits. Some of those lead to articles or scientific papers or summaries that accurately and informatively describe the research in that area. Many others do not. The descriptions, when they are of high quality, in many cases, are not framed in a way that would be particularly useful to the school superintendent, and it is just very difficult." (Hearing before the House Subcommittee, 2002c p. 17-18)

In sum, the WWC was expected to be an improvement over ERIC by ensuring that (a) all disseminated research is vetted for quality, and (b) that research is presented in an engaging, user-friendly way. Whitehurst's conversation with the Senate Committee proceeds similarly, in that he again speaks of the WWC as an innovation. He introduces the WWC as a dissemination approach that:

> "for *the first time* [will be] a place that people can turn for evidence with respect to educational products and programs and approaches…" (Hearing before the Senate Committee. 2002d, emphasis mine).

Taken together, his comments suggest that the WWC was intended to contribute something more to current research dissemination strategies such as ERIC.

But what, exactly, was that contribution expected to be? Clues exist in Whitehurst's discussion with the House Subcommittee, as well as in his subsequent lectures and interviews. Across these three sources, Whitehurst reiterates three related points about the WWC's intended contribution. The first, as just mentioned, is that it will ensure only high-quality educational research is disseminated to educators and policymakers. For example, although some of the hits in a typical ERIC search include trustworthy scientific literature, he claimed that "many do not" (Hearing before the House Subcommittee, 2002c). The WWC's systematic review process was meant to change this. The second expected contribution from the WWC was that it would be "user-friendly" and "accessible." He makes this most clear in a lecture given to the Institute for Policy Research at Northwestern University. In it, Whitehurst lays out his goals for the newly-established Institute of Education Sciences (IES). After acknowledging that "rigorous research by itself will not transform education into an evidence-based field," he discusses the WWC's place in making research knowledge "clear, user-friendly, and easily accessible" (Whitehurst, 2004, p 15). This is yet another contribution expected of the WWC. Third and relatedly, Whitehurst implies that the WWC is meant to be useful for education decision-making. This logic is somewhat implicit, but it does come across in one of his congressional testimonies, in which he states that:

> "…it is very important that the U.S. Department of Education provide that information to practitioners and school superintendents and educators in the form that you are mentioning, and so that it is user-friendly, pre-adjusted, understandable, and useful in decision-making." (Hearing before the House Subcommittee, 2002c)

He follows similar logic in his *THE* interview, in which he responds to a question about the WWC's purpose by calling it "an instrument that is meant to be *used by people*…that is relevant to the decisions they have to make" (p. 2, emphasis added). Whitehurst's comments indicate that the WWC is intended to deliver educational research in a way that certifies its quality, ensures its accessibility, and aids in the decision-making process.

### *The WWC's Own Conceptualization Over Time*

Though Whitehurst's original expectations for the WWC were informative, they occurred very early on in the WWC's lifespan. Given that the WWC recently celebrated its 15th anniversary (Schneider, 2018), I wanted to examine how the WWC has come to define its own intended contribution and how this may have shifted over time. Evidence gathered from the WWC's Standards and Procedures Handbooks, as well as some of the WWC's other published resources, was used to do so (see Table 4.2). Two points of interest emerged from my review. First, although the standards and procedures featured in the handbooks had changed over time, the WWC's mission has remained stable. Since 2008, every handbook has communicated the WWC's mission in similar terms—namely, as being a "central and trusted source of scientific evidence for what works in education." This mission can be connected to Whitehurst's original comments. For example, in order to become a trusted source of scientific evidence, the WWC vets research for quality. Similarly, in order to become a central source of scientific evidence, the WWC disseminates research in user-friendly ways.

Interestingly, this mission of being "central and trusted" does not make an explicit connection back to Whitehurst's hope that the WWC would be used for decision-making.

This connection is made in another place. The recently produced "What Works Video" (WWC, 2018c) describes the WWC in the following way:

> "An investment of the Institute of Education Sciences in the U.S. Department of Education, the WWC is a resource that helps teachers, researchers, administrators, and policymakers *make evidence-based decisions*." (emphasis mine)

Not only does this description allude to Whitehurst's hope, but it also situates the WWC within the Institute of Education Sciences more broadly. Realizing this, we can return back to the handbook passages and notice the same thing. The WWC's mission is always framed as part of the IES's larger mission. This is notable because of how the IES considers the WWC in its own documentation, which I will now describe briefly.

***The IES's Conceptualization of the WWC***

Considering that the WWC is frequently referenced in relation to the IES's larger goals, I next examined how current IES sources characterize the WWC and its role within the Institute. As shown in Table 4.3, nearly all of these sources reaffirmed the IES-WWC relationship referenced in the handbooks. What's more, these IES sources spoke of the WWC's centrality within the IES's larger mission, referring to the WWC as a "flagship product" and its outputs as "marquee" and "some of the most important products of the Institute." Elsewhere, the IES recognizes the WWC as part of its goal to "…increase use of data and research in education decision making" (IES, n.d.-a). In other words, it may be inappropriate to consider the WWC's intended contribution outside of the IES's own. The WWC's mission of becoming a "central and trusted" source of scientific evidence is perhaps best thought of as *how* it fits in to the IES's larger goal—namely, its goal of promoting evidence-based decision-making in education.

**Determining the Extent of the Contribution Expected**

After clarifying the nature of the WWC's expected contribution, it was necessary to also clarify the expected extent of that contribution. This is important because the contribution analysis must differentiate between a program's observed impact and its intended one. More specifically, to what extent does the WWC need to elicit research use among practitioners in order for its contribution to be considered "meaningful" or "important" (e.g., Mayne, 2011, p. 1)? As explained in the previous chapter, in order to answer this question, a predetermined benchmark is needed. The last time a benchmark was explicitly (or at least publicly) established was as part of the IES's PART indicators, which set the following target:

> "…the percentage of decisionmakers surveyed in 2013–2014 who indicate they consult the What Works Clearinghouse prior to making decision(s) on reading, writing, math, science or teacher quality interventions. The target set for 2013–2014 is 25 percent" (Baldwin et al., 2008, p. 108)

Although a survey of decisionmakers never occurred, the established benchmark is still helpful in that it that provides a rough indication of how "central" the WWC was expected to be by 2014.

Of course, the 2013-2014 fiscal year has long since passed. Therefore, using a simple calculation, I took the liberty of updating this benchmark for the year 2020. I did this by dividing the target set for 2013-2014—that is, 25% usage by education decisionmakers—by the number of years the WWC would have existed at that time (i.e., 12 years, if we consider the WWC's initial conception as occurring in 2002). Given that estimated rate of yearly increase (i.e., 2.08% per year), as well as the fact that 6 years have passed since the 2013-2014 target year, we can reasonably expect that an additional 12.5% of educators (i.e., 6*2.08) would need to report consulting the WWC in 2020 in

order for it to stay on track with this updated benchmark. Accordingly, I concluded that the expected size of the WWC's current contribution would be that 37.5% (i.e., 25% + 12.5%) of educators report using it to inform their decision-making.

## Step 2: Developing the Theory of Change

In many ways, Step 2 is the most critical step in Mayne's (2008, 2012b, 2019) instructions. The centrality of a program's theory of change in the contribution analysis (CA) is firmly stated in Mayne's (2012) explanation of the procedure:

> "CA is based on the existence of, or more usually, the development of a postulated theory of change for the intervention being examined. The analysis examines and tests this theory against logic and the evidence available from results observed and the various assumptions behind the theory of change, and examines other influencing factors. The analysis either confirms – verifies – the postulated theory of change or suggests revisions in the theory where the reality appears otherwise." (p. 271)

Notable to this definition is the acknowledgement that theories of change often have to be developed as part of a contribution analysis; seldom do they exist beforehand. Hence, a verifiable theory of change for the WWC resulted from this contribution analysis, and, as such, its development is explained below. The complete theory of change in depicted in Figure 4.1.

### *Developing the Impact Pathway*

The initial step of developing a theory of change involves establishing the program's impact pathway. An impact pathway is the backbone for any theory of change. Like a logic model, it illustrates the sequence of steps comprising the program's causal pathway—from activities to impacts. In my model, the impact pathway is demarcated by bolded boxes and arrows. As acknowledged elsewhere (Lemire, Whynot, & Montague, 2019), impact pathways and logic models often oversimplify a program's causal pathway

by suggesting that it proceeds linearly. To account for this in my model, I include bidirectional arrows in some steps, which is meant to acknowledge the possibility of feedback loops between steps (e.g., capacity changes and behavioral changes feeding one another).

**Activities.** Activities refer to the actions taken by a program to bring about its intended results. The WWC has been very consistent in how it describes its activities. Beginning with Version 2.1 of the WWC Handbook (2011), the WWC has repeatedly described itself as doing the following:

> "…*identifying* existing research on education interventions, *assessing* the quality of this research, and *summarizing* and *disseminating* the evidence from studies that meet WWC standards." (Handbook V4.1, 2019; emphasis added)

In addition, these activities are characteristic of systematic review efforts more generally (Boaz, Ashby, & Young, 2002; Schlooser, 2006). As such, they were specified as the WWC's main activities in this theory of change.

**Outputs.** Outputs refer to the tangible goods or services produced by a program's activities to help progress it towards its intended results. The following four outputs, which were also addressed in Chapter 2, are included in the model:

(1) The "Find What Works" database, which "allows users to identify programs, policies, and practices that have been shown to improve student outcomes",

(2) WWC Practice Guides, which "provide practical recommendations for policy and practice changes",

(3) WWC Intervention Reports, which act as "a summary of findings of the highest-quality research on a given program, policy, practice, or product in education", and

(4) WWC Reviews of Individual Studies, which serve as the basis for these other outputs.

Admittedly, the WWC produces many other outputs, including webinars, videos, and occasional conference presentations. These other outputs were excluded for two reasons. First, they often address issues of research production rather than research use, which means they fall outside the type of contribution examined in this analysis (i.e., increasing rates of evidence-based decision-making per the IES's mission). For example, webinars hosted in 2019 included topics such as "Missing Data in Group Design Studies" and "Documenting Study Context in WWC Reviews of Group Study Designs," suggesting that they are intended for researchers more so than educators. Second and relatedly, I wanted this theory of change to include outputs commonly spotlighted by the WWC, especially when it interfaces with beneficiaries. The WWC's "What We Do" video proved to be a helpful resource from which to identify these outputs. In that video, each of the four outputs included in my model were highlighted, which is why they were included in turn.

**Reach.** Reach is the step in a program's causal chain where its outputs are received by intended beneficiaries. Reach is often overlooked in logic models (e.g., Mayne, 2015). This is unfortunate given that failures in program implementation are a common reason for program ineffectiveness (for a discussion, see Love, 2004). So, in order for the WWC to make its intended contribution, its outputs must reach its intended beneficiaries. But who, exactly, is the WWC intending to reach? Generally speaking, the WWC casts a wide net when discussing its intended users. As noted in the "What We Do" video, these users not only include educators but also researchers, policymakers, and parents (IES, 2018c). Nevertheless, for these purposes, I have chosen to emphasize that the WWC seeks to reach teachers and administrators. Emphasizing these groups was

necessary given the available evidence. To be clear, sufficient evidence exists regarding policymakers' engagement with educational research, but very little is known about their engagement with the WWC specifically. Even less work has examined parents' research engagement and use of the WWC. Because Step 4 and 6 hinge on evidence pertinent to the theory of change, I restricted my focus to intended users from which evidence was available—namely, teachers and administrators. As noted in Chapter 2, there was also a theoretical reason for doing so, as Farley-Ripple and colleagues' (2018) model restricts its definition of the "practice community" to "all school and district practitioners: school district administrators, principals, interventionists, and teachers" (p. 237).

**Capacity Change.** Capacity change refers to cognitive changes—such as changes in attitudes, knowledge, and aspirations—resulting from beneficiaries' engagement with the program outputs. According to Mayne (2015, 2019), these cognitive changes are necessary before eventual behavioral changes can occur. More specifically, Mayne references the work of Michie, Atkins, and West (2014), which suggests that capacity changes culminate in the motivation to change one's behavior. This is why the capacity change component in my theory of change requires that "[b]eneficiaries become *motivated* to use WWC resources to inform decision-making."

**Behavioral Change.** Behavior change refers to the changes in practice that result from beneficiaries engaging with the program's outputs. The WWC's intended behavior change is that beneficiaries will more regularly consult the WWC before making educational decisions. This is clearly identified in the IES's own benchmark for the Clearinghouse, which as a reminder, intended for a certain percentage of educators (i.e., 25%) to "consult the What Works Clearinghouse prior to making decision(s)" (as cited in

Baldwin et al., 2008, p. 108). An identical sentiment was expressed by Whitehurst in his *T.H.E. Journal* interview, in which he explained that the goal of the WWC is, quite simply, to be used in decision-making:

> "The work of the Clearinghouse is to provide an instrument that can be used by people, such as readers of T.H.E. Journal, which will provide them with such information, as is available, that's relevant to the decisions they have to make when they purchase technology, of [sic] a curriculum of [sic] a professional development model." (Mageau, 2004, p. 33-47)

### *Identifying Underlying Assumptions*

Impact pathways are necessary to theories of change, yet they are not sufficient. Theories of change require the identification and integration of assumptions at each step of the pathway. Two types of assumptions are commonly included. The first are *causal link assumptions*, which represent the underlying events and/or conditions "likely necessary" (Mayne, 2019, p. 172) for each step in the pathway to proceed as intended. The second type of assumptions are called *rationale assumptions*, which represent the underlying premises on which a program is founded. Given the nature of rationale assumptions, I discuss them first.

**Rationale Assumptions.** To identify the rationale assumptions underlying the WWC, I used a typology offered by Nkwake and Morrow (2016), which specify three types of rationale assumptions. Each type is explored below. As evidence for their existence, I again draw from statements made by policymakers and scholars during congressional hearings regarding the state of educational research in the early 2000s. For a more comprehensive analysis of these hearings, consult Zoellner (2010).

***Normative Assumptions.*** Normative assumptions are "the considerations right from before an intervention is devised, that there exists a problem and (or opportunity)

that deserves a response—that there is a discrepancy between the reality and what is ideal" (Nkwake & Morrow, 2006, p. 99). This discrepancy is alluded to in a series of comments made by Michael Castle, who served as the House Education Reform Subcommittee chairman in the years leading up to the creation of the IES. In his opening statement during the hearing on *The Reauthorization of the Office of Education Research and Improvement*, Castle vocalized his ideal reality, asserting that "I want quality education research, not fads or anecdotes, to inform educators' decisions on the best way to improve student learning and narrow achievement gaps" (Hearing Before the House Subcommittee, 2002d). Elsewhere, Castle clarifies the problem further, stating that:

> "Today, schools invest untold time and resources in one education fad after another. Without sound science to back program claims, teachers and school administrators are forced to use guesswork to determine the best classroom practices in students and students' achievement often suffers." (p. 79, as cited in Zoellner, 2010)

According to these statements, the presumed problem is clear: education reform is failing because it is being based off of "fads or anecdotes," not rigorous educational research.

Grover Whitehurst, who eventually led the charge in creating the WWC, shared these sentiments. In fact, he refers to the same problem as Castle when voicing his goals as the inaugural IES director. For instance, in an address given at the American Educational Research Association's (AERA) annual convention, he expressed concern about the gap between the findings of educational research and the actions of practitioners in the classroom. He referred to the problem as "a gulf between the bench and the trench" (2003, p. 5). He later concludes his address by describing his ideal educational landscape, stating that "I have a vision of a day when any educator or policy maker will want to know what the research says before making an important decision" (p.

13). It is for these reasons that the normative assumptions established in my theory of change (see Figure 4.1) allude both an unideal reality (i.e., the existence of a research-practice gap), and idealized future (i.e., that educational practice should be guided by educational research).

***Diagnostic Assumptions.*** Diagnostic assumptions are "stakeholders' perceptions of the major and minor causes of the core problems" (Nkwake & Morrow, 2016, p. 100). So, prior to creating the WWC, what did policymakers view as the primary causes of the research-practice gap? Later in Whitehurst's (2003) AERA address, while describing his ideal reality for educational research and practice, Whitehurst alludes to three shortcomings in educational research, claiming that he "has a vision of a day when…the research will be there. It will be rigorous. It will be relevant. It will be disseminated and accessed through tools that make it useable" (p. 13) Each of these issues—rigor, relevance, and usability—is addressed in the congressional hearings leading to the creation of the IES and WWC.

Take, for example, the issue of rigor. Whitehurst's testimony about the lack of high-quality research in databases like ERIC speaks to his concerns about the quality of educational research. G. Reid Lyon, who served as chief of the Child Development and Behavior Branch at NIH at the time of the hearings, takes a related yet distinct perspective. According to him, the research-practice gap is not strictly a byproduct of poor research quality, but rather the result of research users not being able to vet it themselves. His testimony included the following:

> "At the present time our teachers, from our studies of teachers, really don't understand what to look at and research to judge what is fluff.... They really don't know what to look for, and that is how they get stuck in these continuing fads that

> come down the pike, most of which don't have any research or any good research." (Hearing Before the House Subcommittee, 2000c)

Hence, the normative assumption regarding rigor simultaneously holds that the research-practice gap is the byproduct of educational research's poor quality, and moreover, educators' inability to distinguish the good from the bad.

Another normative assumption that appears in these testimonies is that most educational research is irrelevant to educators. Whitehurst affirms this sentiment in his 2004 lecture to Northwestern University's *Institute of Policy Research*, where he claimed that "there is a mismatch between what education decision makers want from the education research and what the education research community is providing" (p. 13). This point is reiterated in the congressional hearings. In an earlier hearing, Robert Slavin, a prominent educational researcher who continues to direct the *Center for Research and Reform in Education* at John Hopkins University, confessed that "very, very seldom do we set up research to go out and solve a problem…[t]hat has been the missing element in the portfolio of research until recently" (Hearing Before the House Subcommittee, 2000c). Slavin's words imply that educational research is, in general, only distally related to the immediate problems faced by educators. Furthermore, during one of the 2002 hearings, Douglas Christensen, the commissioner of the Nebraska Department of Education, testifies against the relevance of educational research even more directly:

> "I would like to simply say is that coming from a small state, a rural state in the heartland, we couldn't feel anymore disconnected from the national research issues and research agenda than if we were located on a manned space station. We just don't feel connected to it" (Hearing Before the House Subcommittee, 2002b)

As such, irrelevance of research was assumed to be another cause of the research-practice divide.

Third and relatedly, the inaccessibility of educational research is frequently cited as a normative assumption—that is, as a contributor to the research-practice gap. This is because, at least in the eyes of Whitehurst, that even when research is relevant and of high-quality, it is still "not framed in a way that would be particularly useful to the school superintendent" (Hearing Before the House Subcommittee, 2002c). Even as a policymaker, Castle was aware of these dissemination issues. He believed that the lack of research dissemination resulted from poor communication, claiming that "I know there are websites and there are various other ways of disseminating it, but I just am not convinced. When I talk to people at home they don't even know what I'm talking about." (Hearing Before the House Subcommittee, 2002d). Thus, any solution to the problem would need to address poor dissemination efforts as well.

***Prescriptive Assumptions.*** But what would this solution look like? Perspective assumptions "represent stakeholders' beliefs of what could be the best ways to address the problem or need" (Nkwake & Morrow, 2016, p. 100). The glaring, though largely implicit assumption underlying these congressional hearings is that government intervention is needed to help bring educational research to bear to practice. As explained in Chapter 1, this assumption has guided education policy since the mid-twentieth century (for a review, see Vinovskis, 1998). This assumption was also reiterated by Whitehurst and others as the IES was taking form. In an interview with *EdWeek* after his Senate nomination process, Whitehurst suggested that the federal government does, in fact, have a role to play in bridging the research-practice gap. Specifically, he explained how he has "often talked about how the federal government is not doing as good a job as it should be doing in marshaling researchers to respond to the practical needs of the field"

(Viadero, 2001). Policymakers agreed. For example, Michael Castle also spoke of federal intervention:

> "Education research is broken in our country, and Congress must work to make it more useful, more independent of political influence, and less bureaucratic than the current system…"

Likewise, Castle's colleague, representative Bill Goodling, who cosponsored the *Scientifically Based Educational Research, Evaluation, Statistics, and Information Act of 2000* with Michael Castle, suggested something similar:

> "[the] federal government can play an important role in gathering information, conducting research and disseminating information on education practices that work to improve America's schools." (Hearing before the Senate and House Committees, 1999)

Thus, one of the prescriptive assumptions motivating the creation of the WWC is the belief that increased government involvement would be a promising solution to the problem.

But why create the WWC specifically? As established in Chapter 1, increased government involvement in educational research has taken myriad forms. An explicit rationalization of government-sponsored systematic research review does not appear in these hearings, but it is certainly alluded to. For instance, Whitehurst's brief mention of the WWC in the House Subcommittee Hearing (2002a) is notable in the way he differentiates it from ERIC, which at that time, was the main government-operated repository for educational research. He suggests that one shortcoming of ERIC is that it allows low-quality research to be disseminated. This comment suggests that Whitehurst was seeking a solution to the issues of relevance, rigor, and accessibility noted above, and in turn, a systematic review initiative like the WWC could address these concerns in turn.

Given Whitehurst's confidence in the systematic review process, it was included as a prescriptive assumption as well.

**Causal Link Assumptions.** Now that the model's rationale assumptions have been established, attention can be turned to the *causal link assumptions*. Causal link assumptions explicate the necessary events or preconditions that must be in place for one step of the impact pathway to move to the next. I again relied on Mayne's (2015) generic behavior change model for help with identifying assumptions at each step, and, when needed, I rationalized specific assumptions using some of the same literature reviewed in Chapter 2.

***Output Assumptions.*** Unlike Mayne's model, which begins identifying assumptions at the "Reach" level of the theory of change, this theory of change specifies one output assumption. Indeed, there are necessary conditions that must be in place for the WWC's activities to result in outputs. One obvious but easily overlooked assumption is that WWC has enough capacity to produce its outputs. In other words, the WWC must have the staff and resources needed to summarize its systematic review findings in reports and webpages for beneficiaries. Including this assumption is justified based on observations from a previous WWC evaluation. Specifically, in their 2010 audit of the WWC, the U.S. Government Accountability Office (GAO) detected "a substantial backlog in its product review process" (p. 1). Because the IES review process of WWC outputs was not keeping pace with the amount of outputs produced by WWC contractors, completed intervention reports were taking, on average, 50 days to reach publication. Given this finding, I included an output assumption specifying "the capacity to develop and publish outputs."

***Reach Assumptions.*** Reach assumptions refer to likely necessary events or conditions for the WWC's outputs to reach its intended beneficiaries. Because these WWC outputs are disseminated passively according to a producer-push model (e.g., Lavis et al., 2003; Levin, 2013), the extent of their reach is contingent on several assumptions. The first is that beneficiaries are aware of the outputs; if not, then they will not seek them out. Awareness, however, does not ensure reach—especially in passive dissemination situations. Therefore, I included two more assumptions of reach. The first is that educators are willing to seek out the outputs, which must occur even if they are aware of them. The second is that educators are able to seek them out, which is justifiable given the range of organizational factors that can facilitate or hinder research use (e.g., Schaik et al., 2018).

***Capacity Change Assumptions.*** Capacity change assumptions refer to likely necessary conditions and/or events if the outputs, once reached, are to elicit the intended cognitive changes in beneficiaries. In the context of this model, these assumptions are the conditions necessary for the beneficiaries to become motivated to make decisions in accordance with the WWC's outputs. In my model, two conditions are necessary, and both align with the WWC's intention of becoming a "central" and "trusted" source of educational research. Indeed, becoming a more "central" and "trusted" source of education research depends on how central and trustworthy the WWC's beneficiaries perceive it to be. So, the first capacity change assumption is that the WWC's resources are perceived as relevant. Relevant resources are not only those that address topics of relevance, but they also include information that is perceived to be useable. The second capacity change assumption is that the WWC's resources are perceived as trustworthy.

This assumption is somewhat contingent on the first, as research suggests that educators' perceptions about the credibility of research are tied to their perceptions of its relevance (e.g., Cain, 2016). Even so, my theory of change contends that both of these assumptions must be realized before behavior change will willingly occur.

***Behavioral Change Assumptions.*** Just because a beneficiary becomes motivated to make decisions in accordance with WWC resources does not mean that they will. Accordingly, there are conditions likely necessary for this motivation to translate into instrumental research use, and these conditions are referred to as behavior change assumptions. One assumption specified in this theory of change is that instrumental use is possible. Depending on the level of authority possessed by the beneficiary, changes to their behavior may be contingent on facilitators/barriers in their practice environment (for a discussion, see Tseng, 2012). Therefore, the intended leap from capacity change to behavior change relies on this assumption.

***Intended Size of Contribution*.** One final condition must be met in order for any resulting behavioral change to constitute a meaningful contribution. As discussed above, for any behavioral change to constitute a meaningful contribution, it must meet or surpass the WWC's adjusted benchmark of at least 37.5% of educators consulting the WWC prior to making an educational decision (e.g., Baldwin et al., 2008).

### Step 3: Gathering Existing Evidence

After developing the theory of change, contribution analysis proceeds with the evaluator gathering and collating evidence—such as "previous measurement, past evaluations, and relevant literature" (Mayne, 2019, p. 272)—from which to test the

program's theory. All evidence sources used to evaluate the theory of change were previously outlined in Chapter 3.

## Step 4: Assemble and Assess the Contribution Claim

Step 4 involves what might traditionally be thought of as the "results" of the contribution analysis. Not only does this step call for a detailed evaluation of the WWC's theory of change (i.e., the telling of a "contribution story"), but it culminates in a "contribution claim" about the intervention's observed impact relative to its expected one. This claim is rarely definitive; often, it is probabilistic in nature (Mayne, 2019). Nevertheless, the contribution claim should still be compelling enough that "a reasonable person would agree from the evidence and argument that the program has made an important contribution [or not]" (Mayne, 2011, p. 62). I lay out both the contribution story and its associated contribution claims below.

### *Assessing the Rationale Assumptions*

I first examined the WWC's rationale assumptions, which refer to the often-implicit premises on which it was founded. Evidence confirming and/or refuting these assumptions was primarily drawn from the empirical literature. Given that these assumptions fueled the WWC's creation in the first place, I paid special attention to Hemsley-Brown and Sharp's (2003) literature review, which featured research published between 1988 and 2001—that is, research that would have been available at the time of the WWC's conception. I attend to more recent reviews as well. At times, evidence from sources that were not previously reviewed—including existing evaluations of IES initiatives and recent large-scale surveys of educators' research use—are considered too. Ultimately, the evidence inconsistently supports the WWC's rationale assumptions. First,

its normative assumptions are more closely aligned to the realities of teachers than they are to administrators. Second, although some evidence justifies its diagnostic assumptions, these assumptions largely ignore factors external to educational research or educators themselves (e.g., structural barriers to research use). And third, its prescriptive assumptions about how best to address the research-practice gap are largely misinformed.

**Normative Assumptions.** As demonstrated in Step 2, congressional testimony offered by Grover Whitehurst, as well as responses from several prominent policymakers, suggest that the WWC was developed in response to the belief that a research-practice gap exists in education. Despite the pervasiveness of their belief, at the time of their concerns, there was actually little empirical evidence verifying the existence of a gap. Hemsley-Brown and Sharp's (2003) review only found six empirical articles addressing the nature of research use by administrators and teacher, and those articles revealed a more nuanced research-practice landscape than Whitehurst and his colleagues initially believed. Specifically, those articles revealed that, while research was seldom used by classroom teachers, administrators actually appeared to be "regular, thoughtful users of research knowledge" (Hemsley-Brown & Sharp, 2004, p. 9). Hemsley-Brown and Sharp's observation is echoed in recent large-scale surveys, which suggest that administrators' research use tends to outpace use by teachers. For example, Penuel and colleagues (2017) survey of administrators found that roughly 80% reported using research "frequently" or "all of the time" for instrumental purposes; likewise, 60% said that their imposed use of research occurred "frequently" or "all the time." Direct comparisons between teachers and administrators, such as those reported by Lysenko et al. (2014), found administrators to report more regular research usage, whereas teachers'

usage was deemed "sporadic" (p. 14). Therefore, although it may still be desirable for administrators to consult educational research more than they already do, the evidence is consistent in showing that the research-practice gap is much narrower among them than it is among teachers. In this way, the normative assumption that a research-practice gap exists is justifiable among teachers but not among administrators.

The normative assumption that practice *should* be based off research is challenging to confirm or deny empirically, as it is essentially a value judgement. The obvious hope underlying this assumption, which is also communicated in the congressional hearings (House Subcommittee on Education Reform, 2002a; Senate Committee on Education, 2002a), is that educational outcomes (especially student achievement) could be improved by research-informed practice. Both then and now, evidence is scarce that bringing research evidence to bear in practice actually manifests in downstream improvements for students and/or teachers. In fact, Cain (2016) recently concluded that "[l]ittle is known about the effect of teachers' research use on their teaching and their students" (p. 623). Along similar lines, after reviewing the literature, Gorard et al. (2020) determined that:

> "There is no evidence that end-user outcomes (such as student test scores in education) were improved by use of evidence, and there is some evidence that they are not." (p. 18)

If this is indeed the case, then perhaps the narrowing of the research-practice gap is less desirable than policymakers initially assumed. Even so, more evidence is needed to conclude that closing the research-practice gap could, in fact, be determinantal to practice.

**Diagnostic Assumptions.** The WWC's theory of change also specifies diagnostic assumptions about the root causes of the research-practice gap. Not only was the quality, relevance, and accessibility of educational research questioned in these initial congressional hearings (House Subcommittee on Education Reform, 2002a; Senate Committee on Education, 2002a), but concerns were also expressed about educators' lack of preparedness to access and/or assess research themselves. Empirical support for these concerns is far less consistent, however. Although Hemsley-Brown and Sharp's (2003) review did draw a distinction between teachers' and administrators' perceptions of research—with teachers finding it "irrelevant, unhelpful, and too theoretical" (p. 15) while administrators had "a generally positive view of research" (p. 9)—more recent work finds that both teachers and administrators can harbor a range of perceptions.

Take teachers, for instance. Reviews of the literature (e.g., Schaik et al., 2018) still find that "[t]eachers criticise research knowledge [for] being unapproachable, inaccessible, difficult and incomprehensible" (p. 54). Teachers' criticisms of educational research were noted in qualitative work by Cain (2016) and Joram (2020) as well. Even so, teachers in Lysenko and colleagues' (2014) survey generally reported "neutral" attitudes about the relevance and trustworthiness of educational research, as well as its applicability to practice. What's more, teachers in the survey administered by the EdTech Evidence Exchange (Barton & Tindle, 2019) had slightly positive attitudes about the relevance, reliability, and usefulness of educational research. A diversity in attitudes has been reported by administrators as well. Counter to Hemsley-Brown and Sharp's (2003) conclusions, Penuel et al's (2017) survey found administrators to report "mixed" feelings about the relevance and credibility of research (p. 11). Accordingly, evidence from the

extant literature suggests that there is some support for the WWC's diagnostic assumption about the irrelevance and inapplicability of educational research, but that support is far from overwhelming. There is even some opposing evidence that educators find research to be both relevant and rigorous, which means that justification for this assumption is mixed.

The evidence is also mixed with regard to educators' abilities to be thoughtful consumers of research. Literature reviews have recognized that teachers' skills and competencies to engage with research—which other work refers to as their "information literacy" (e.g., Williams & Coles, 2007)—tend to be lacking, which in turn, impedes their research use (Dagenais et al., 2012, p. 298). Even so, in subsequent work by the same authors (e.g., Lysenko, Abrami, Bernard, Dagenais, & Janosz, 2014; Lysenko, Abrami, Bernard, & Dagenais, 2015), both practitioners and administrators have endorsed their own abilities to, among other things, "read and understand research publications" (Lysenko et al., 2015, p. 43). This finding, along with the increasing popularity of professional development trainings designed to sharpen teachers' data literacy (e.g., Ebbeler, Poortman, Schildkamp, & Pieters, 2017; Kippers, Poortman, Schildkamp, & Visscher, 2018), complicates the WWC's assumption that the research-practice gap is the result of educators being ill-prepared to consume research.

If anything, the available literature implicates factors external to both educators and educational research as the chief contributors to a research-practice gap. As detailed in Chapter 2, these factors include structural barriers, such as teachers' lack of time to engage with educational research (Martinovic et al., 2012; Joram et al., 2020) and/or the lack of agency afforded for teacher decision-making (Leat et al., 2015; Joram et al.,

2020). They also include aspects of educators' institutional cultures, such as when a school or district discourages an organizational learning culture (Cousins & Walker, 2000; Penuel et al., 2017). Models of teachers' research use (e.g., Schaik et al., 2018), as well as Farley-Ripple et al.'s (2018) framework for bridging the research-practice gap, implicate these factors as well.

**Prescriptive Assumptions.** A direct line can be drawn between the creation of the WWC and its underlying prescriptive assumptions, which suggest that a government-sponsored systematic review initiative would be the best way to address the core causes of the research-practice gap. As discussed in Chapter 1, the idea that a government-controlled initiative would be the best way to address the research-practice gap is nothing new. What is new, however, is evidence suggesting that government control may actually be counterproductive—or at least unpopular. For instance, findings from Penuel et al.'s (2017) survey of administrators found they were far less likely to consult federally-sponsored repositories of research information —like the WWC, RELs, and the National Center for Education Statistics—than people in their professional networks. Findings from evaluations sponsored by the IES suggest similarly. For example, a survey conducted as part of the Government Accountability Agency's (2010) evaluation reported that administrators were at least three times more likely to rate localized information sources (e.g., personal experience, internal research, peer conferences, and colleagues) as "useful" or "very useful" as sources for identifying useful education practices than the What Works Clearinghouse or "other federal outreach centers" (p. 60). Similarly, Barton and Tindle (2019) found "peer colleagues" to be twice as popular research access points among teachers and administrators than any of the five federally sponsored initiatives

included as response options. These findings do not necessarily mean that access points like the WWC are avoided *because* they reside under government control. However, they do suggest—just as was noted in Chapter 2—that educators have other trusted sources of research information. They also jeopardize the prescriptive assumption that government control is the most appropriate way of addressing the research-practice gap.

Likewise, these very same findings refute the WWC's second perspective assumption that systematic review is the most appropriate way of bringing research to bear in practice. Systematic review may be a useful way of addressing concerns about the quality of educational research, especially among those subscribing to the "best evidence" standards of rigor advocated by Slavin (1986). It may also appeal to educators who desire research information to be presented in predigested ways (e.g., Williams & Coles, 2007a). Even so, evidence suggests that other types of research dissemination are far more effective than systematic review. Most of the work has studied the research-seeking behaviors of administrators, not teachers. For example, when asked about their favored research access points, administrators in the GAO (2010) evaluation rated the three response options involving systematic review—including *Doing What Works*, *Child Trends*, and *RAND's Promising Practices*—as the three least useful sources of evidence. Similarly, Penuel et al. (2018) concluded that "the kinds of research district leaders find useful are not primarily peer-reviewed impact studies" (p. 540), which suggests that systematic review would, again, be of little appeal.

Reviews of the literature conclude that dissemination strategies relying on active, interpersonal exchanges are a more useful option. This has been known since Hemsley-Brown and Sharp's (2003) review, which determined that "the empirical research shows

there is no direct positive relationship between systematic dissemination of research findings and impact on policy and practice" (p. 28). Instead, they recommended greater practitioner involvement in research process and strengthened communication networks between researchers and practitioners. An identical observation was made in a new review by Gorard et al. (2020), which found that "[e]ven the best systematic syntheses of evidence often have little impact in practice" (p. 17). They, like Hemsley-Brown and Sharp, suggested the use of "a respected and trusted conduit" to actively disseminate research information (p. 1). Therefore, the assumption that a systematic review initiative like the WWC was the best way to address the research-practice gap is also refuted by the evidence.

### *Assessing the Impact Pathway / Causal Assumptions*

After assessing the WWC's rationale assumptions, the impact pathway—as well as the causal assumptions underlying it—were examined. Along with using many of the same evidence sources reference above, this section incorporates WWC-specific data sources. These include my own tallies of various WWC outputs and their characteristics, as well as analysis of a downloadable dataset managed by WWC contractors. When considered together, this evidence suggested that the WWC is struggling to reach its intended users, and for those whom it does reach, its outputs are likely perceived as irrelevant. The evidence further suggests that some educators—especially those concerned about the generalizability of research findings—may find the WWC's information to be untrustworthy.

**Activities**. The WWC describes itself as engaging in four main activities. Specifically, it identifies existing research on education interventions, assesses that

research according to its predetermined standards of quality, and summarizes findings from research meeting its quality standards; after completing these activities, conclusions are then disseminated through a variety of outputs (IES, 2018c; WWC, n.d.-g). One indicator of these activities can be found in the rate at which the WWC reviews individual studies, as these reviews are the raw material from which intervention reports and practice guides are built. Calculations from the WWC's extractable dataset suggest that its individual study review rate has proceeded as a series of peaks and valleys. As acknowledged elsewhere (e.g., GAO, 2010), the WWC's study review process got off to a slow start. 0 studies were reviewed in 2003, 6 studies were reviewed in 2004, 188 were reviewed in 2005, and 103 were reviewed in 2006 (see Table 3.1). By the late 2000s, however, it had accelerated greatly. Data indicated that in 2007 alone, the WWC quadrupled the number of individual study reviews that had occurred in its first four years (i.e., 1146 vs. 297). Between 2008 and 2017, the average number of annual study reviews was 1339. However, since 2018, the number of individual study reviews has dipped, with only 122 reviewed in 2018 and 127 reviewed in 2019—an average of 124 per year. More evidence is needed before this dip can be treated as a lapse in program activities, but it is still concerning.

A second strand of evidence pertinent to the WWC's activities can be found in the rate at which it produces and updates its review protocols. As a reminder, these protocols guide WWC review teams as they conduct their systematic reviews. Once these protocols become outdated, they are updated so that literature published since the previous review can be accommodated. The production, and subsequent update, of review protocols is a good indicator of WWC functioning because it signals that additional areas of literature

are being examined or reexamined. This activity must occur if the WWC is to keep up with the ever-expanding field of educational research and evaluation. Tallies of review protocols from the WWC's website suggest that their pace of development has followed a similar trend to the individual study reviews. The number of protocols developed during its second and third five-year contract cycles (16 and 15, respectively) double the number of protocols produced during the first five years (i.e., 7) (see Table 3.1). Despite this increase, only 5 had been developed since 2018. Though this dip is concerning, it may be the result of a shifted focus towards updating outdated protocols. Indeed, the data showed that, since 2012, the average number of updated protocols each year (4) is greater than the number of newly developed protocols (2.56). This is especially true more recently: 14 protocols have been updated since 2017 whereas only 6 protocols have been developed. These data offer mixed evidence that the WWC's activities are occurring as planned. Specifically, though protocol development has lagged recently, the WWC has increased the rates at which it updates review protocols to align with current standards and to respond to new developments in the literature.

**Output Assumptions.** The occurrence of these activities, however, does not mean that associated outputs will be generated. The WWC must have the capacity to translate findings from its reviews into resources for public consumption. To verify the assumption that this capacity exists, I examined two indicators. The first was the WWC's funding levels over time. Although increased funding does not always translate to increased capacity, it was used as an indicator nonetheless. Funding data were extracted from USAspends.org, which flagged all WWC-related contracts in my keyword search. These data, along with information reported in both the GAO evaluation and the IES

biennial reports, were triangulated to get a sense of funding patterns over time. As shown in Figure 4.2, WWC funding rose between its first and second contract cycles; in fact, GAO evaluators (2010) reported that the WWC's second five-year contract (totaling $53.3 million) essentially doubled the amount of its first contract ($26.5 million). Even so, subsequent contracts have declined in size, and year-by-year analysis using USAspends.org data confirms that funding levels have declined somewhat since 2009. This suggests that investment in the WWC may have waned slightly over the past decade, which, in turn, may indicate a similar decline in capacity.

A second indicator of capacity was considered. Specifically, I examined the available evidence for a backlog in output production. Doing so was important given the findings of previous evaluations. Specifically, the previous evaluation conducted by the GAO (2010) discovered evidence of a delay between the date at which an output was drafted by the WWC and the date it was finally published. This backlog in output production was due to a slow IES peer review procedure. At its worst, the publication of WWC outputs was delayed by an average of six months (GAO, 2010). As such, I examined the available evidence for indications of a similar backlog. Using the WWC's extractable data on their individual study reviews, I compared the date on which each review was conducted to the date on which it was posted. Data were only available for 760 individual study reviews that had occurred since 2017. Calculations suggest that the mean difference between posting and review dates was 99.5 days. To account for outliers, I also calculated the median difference, which indicated a 54-day backlog. These findings suggest that the WWC is again struggling with capacity to publish results in a timely

manner. Even so, given that this backlog used to be worse (GAO, 2010), this finding offers mixed evidence of the WWC's capacity to develop and publish outputs.

Other exported data yielded insights about the WWC's reviewers themselves. Two core insights emerged from this data. The first is that WWC certified reviewers appear to have sufficient expertise to review educational research according to the WWC standards; few, however, appear to be members of the practice community. Of the 284 individuals listed as certified reviewers in October 2020, many (177 of 284; 62%) are listed along with their organizational affiliations, and these affiliations provide clues into WWC reviewers' backgrounds. Most belong to research firms or think tanks (162 of 177); in fact, the top three affiliations listed included the American Institutes of Research (38 reviewers), SRI International (22 reviewers), and Abt. Associates (20 reviewers)—all of whom have been WWC contractors. In contrast, 14 of these reviewers had a university affiliation, and only 1 reported an affiliation with a public school district. This pattern suggests that, at least among reviewers with an organizational affiliation list, most can be assumed to be professionals with advanced degrees, yet few can be assumed to be classroom teachers or administrators.

A second observation was that, when examined more closely, the WWC's review capacity is less robust than the raw number of individual reviewers suggests. Only 104 of the 284 certified individual reviewers (37%) had an updated certification in Group Design Standards Version 4.1, which were first published in January 2020; only slightly more (123 out of 284; 43%) had certifications approved for Group Design Standards Version 4.0, which were published in October 2017. Another concerning pattern was the low number of reviewers who were approved to review other types of designs. For

example, only 4 reviewers had an updated certification for Single Case Designs, and even fewer (3 out of 284) had the latest certification for reviewing Regression Discontinuity Designs. Both the lack of reviewers with updated credentials, as well as the lack of reviewers available for single case and regression discontinuity designs, is concerning. It also implies that the WWC may have insufficient capacity to promptly translate its activities into outputs.

**Outputs.** To help inform education decision-making, the WWC disseminates its findings through a series of published resources. These include, among others, reviews of individual studies, intervention reports, and practice guides. Indications of the regularity at which these outputs are published depend on the output examined. Consider, first, the posting of individual study reviews. Extractable data from the WWC were only available beginning in 2017. That data suggests that over half of all study reviews with available data (436 of 760; 57%) were posted in 2018. Study review postings have slowed considerably since then, although postings in 2020 have, thus far, shown some improvement (see Table 3.1). Despite more modest posting rates since 2018, it's worth remembering that individual study reviews are not considered to be core WWC outputs like the intervention reports and practice guides are (WWC, n.d.-g).

As such, I next considered the publication rate of intervention reports. This indicator is especially important because, as was found in the 2010 GAO evaluation, "among school districts that use the Clearinghouse to inform decisions on effective education practices, more school districts use intervention reports relative to practice guides or quick reviews" (p. 36). Tallies of the reports' posting dates revealed a notable pattern in their publication frequency. Specifically, 64% of all intervention reports (i.e.,

382 of 593 total reports) available on the WWC website were published between 2005-2007, with 303 published in 2007 alone. Notably, this high rate of publishing coincides with the funding increases mentioned above. Between 2008-2017, however, the average publication rate was only about 19 reports per year. This rate has slowed even more lately, with only 9 and 4 reports published in 2018 and 2019, respectively. 2020 has, thus far, seen 8 intervention reports published, which is encouraging. Even so, this evidence calls into question the WWC's production of outputs, as it suggests that most intervention reports are outdated, and moreover, that fewer have been produced in recent years.

A similar trend emerged when examining publication rates of practice guides, which the current IES director Mark Schneider (2018) claims to be "some of the most downloaded documents on our website." Again, half of all practice guides (12 of 24) were initially published in a three-year span—from 2007-2009—a time when funding levels were high. The WWC has published an average of roughly 1 practice guide in each subsequent year. This decreased rate of practice guide production since the late 2000s is troubling, but it may be due to a shift in attention to updating/revising existing practice guides. Seven revisions (including two substantive updates) have occurred to date, with the first appearing in 2014. Six of these seven have occurred since 2017, which suggests that the WWC may be prioritizing updating existing guides rather than producing new ones. However, because the publication rate of key WWC resources seems to be both erratic and less prolific now than it was in the late 2000s, there is only spotty evidence that the WWC is continuing to publish outputs at rates that match or surpass previous rates.

**Reach Assumptions.** Regardless, there are still plenty of WWC outputs available for public consumption. For them to be consumed, however, requires that they actually reach their intended users. Reach is contingent on (at least) three assumptions specified in the theory of change—namely, that (a) beneficiaries are aware of the outputs, (b) beneficiaries are able to seek them out, and (c) beneficiaries are willing to seek them out.

***Awareness of outputs.*** Evidence of beneficiaries' awareness is only available from previous evaluations, and together that evidence suggest that awareness is largely dependent on the stakeholder group surveyed. Some groups of administrators are particularly aware of the WWC. For example, findings from the 2010 GAO evaluation discovered that 87% of all state-level administrators in the GAO evaluation had heard of the WWC. In contrast, less than half of all district-level administrators (i.e., 42%) and school-level administrators/principals (i.e., 35%) had. Perhaps most alarming was the finding that only 13% of teachers reported having heard of the WWC in 2010 (GAO, 2010). Importantly, this evaluation was conducted when the WWC was still somewhat young; awareness may have improved over time. According to the EdTech Evidence Exchange's latest estimate, 59% of all respondents had heard of the WWC, including 62% of association members and 36% of convening attendees (Barton & Tindle, 2019). Given that association members were more likely to report serving in non-teacher roles, this result suggests that, after an additional ten years, awareness of the WWC still seems to be contingent on educators' roles, such that administrators tend to be more aware than teachers. Even so, when considering both groups together, only 6 in 10 educators appear to have heard of the WWC. This is an underwhelming finding given that the WWC has existed for nearly two decades. Yet, because awareness appears to have increased over

time—especially among teachers—there is mixed evidence that educators are sufficiently aware of the WWC to engage with its outputs.

***Ability to seek out outputs.*** Additionally, even if beneficiaries are aware of the WWC's outputs, they may still not be able to seek them out. Some evidence from previous evaluations suggests that educators do, in fact, face barriers to accessing the WWC. For example, survey respondents in the GAO evaluation (2010) most commonly reported that they did not access the WWC more frequently because of time constraints. Likewise, and as noted in Chapter 2, the literature similarly suggests that the biggest barrier to educators' research use is their lack of time to seek out and review research. This is especially true of teachers. For example, in their review, Schaik and colleagues (2018) noted that "almost all studies point out that time is insufficiently available" for teachers to access academic research (p. 56). Similarly, Cain (2016) concluded his synthesis of the literature by stating that "this review has found that few teachers read research because most lack the time…to do so" (p. 625). Although the WWC's resources are intended to help educators "quickly" and "easily" access research (e.g., IES, 2018), models of knowledge transfer suggest that this dissemination strategy may be too passive given the countless demands on teachers' time (e.g., Levin, 2011). This was the conclusion reached by Gorard, See, and Siddiqui (2020) in their recent review, which determined that "it is not clear that simply modifying research findings into easier formats, with a passive approach to transfer, leads to any better results" (p. 27). Given this, the assumption that beneficiaries can access the WWC's resources, even when they are aware of them, seems to be ill-founded.

***Willingness to seek out outputs.*** It follows that, even if educators are able to engage with the WWC's resources, they still might not be willing to do so. They must believe that there is value in informing their practice with educational research if they are to seek it out (except in cases of imposed use; e.g., Weiss, 2005, 2008). Reviews of the literature, recent large-scale surveys, and findings from an IES-sponsored evaluation suggest, yet again, that there may be differences between teachers and administrators. For example, Penuel et al.'s survey (2017) found that district leaders and administrators reported very positive attitudes about the value of educational research. Teachers, on the other hand, voiced ambivalence about the value of research to inform practice when surveyed by Lysenko and colleagues (2014); this sentiment is echoed in recent reviews as well (e.g., Schaik et al., 2018). In contrast, positive attitudes about the usefulness of educational research were observed in the survey administered by Barton and Tindle (2019), which included both administrators and teachers. Therefore, the evidence seems to suggest that, while administrators generally see value in educational research, teachers tend to be more skeptical.

A willingness to engage with educational research does not satisfy this assumption, however. Educators who are receptive to research may still lack the willingness to pursue WWC outputs, particularly if they perceive them to be less valuable than the resources they already use. As explained above, the extant evidence suggests that many educators already have preferred "access points" to educational research, and those access points tend to involve other educators who are close to them. For instance, Penuel and colleagues (2017) found administrators' most common access points to include professional conferences and colleagues in other districts—both of which involve active

and relational forms of knowledge transfer. The EdTech Evidence Exchange survey (i.e., Barton & Tindle, 2019) supports this finding, as common access points reported by educators include professional associations, colleagues, and social media. Even dedicated evaluations of the WWC (e.g., GAO, 2010) find educators' colleagues to be a favored access point. Accordingly, the evidence suggests that the WWC may not maximize its reach among educators because it cannot replace their existing reliance on more personal, exchange-oriented access points—even if using those access points directs them to lower quality research (e.g., Gorard et al., 2020).

**Reach.** The WWC's main dissemination mode continues to be its website, which houses the *Find What Works* database as well as its other outputs. Schneider (2019) has called the WWC's website a "marquee activity" for the IES, suggesting that its use is integral to the IES's mission generally and the WWC's own mission more specifically. Though little is known about the characteristics of site visitors (e.g., teachers, administrators, parents, or researchers), some data do exist. The findings reported by Baldwin and colleagues' (2008) are promising in that they confirm the WWC's reach of intended beneficiaries. Specifically, respondents to a pop-up survey on the WWC website most commonly identified themselves as teachers (23%) and administrators (19%), with the third most common respondent being researchers (12%). A later iteration of this pop-up survey, which was administered as part of the WWC's user feedback campaign between 2014-2016, revealed a much different respondent breakdown. In that survey, respondents were able to identify themselves as belonging to multiple user groups, so the findings reported by the WWC (n.d.-f) are challenging to interpret. Even so, the general response patterns are telling. Site visitors were over four times more likely to identify

themselves as "researchers or college/university faculty of staff" (i.e., 269 out of 654 total selections) compared to "pre-K through 12th grade teachers" (i.e., 60 out of 654 total selections); only 38 identified as district-level staff and 14 as principals or assistant principals. In addition, conclusions from the feedback campaign focus groups noted that "[t]he WWC is doing a better job attracting researchers and developers to the site than practitioners" (WWC, n.d.-f). Still, more data is needed.

Findings from previous evaluations also provide indications of how frequently the WWC is used as a research access point among beneficiaries. Results from the GAO evaluation (2010) suggest that reported access rates differ based on the stakeholder group. For example, 87% of state-level officials had reported accessing the WWC at least once, and 34% of district-level officials had reported doing so. In comparison, surveys of practitioners uncovered far smaller proportions of access, with only 15% of principals / local administrators and 5% of teachers doing so. Though these findings are outdated, they still warn of poor program reach among some stakeholder groups. More recent estimates provide some evidence of improvement—but also lack thereof. In Penuel and colleagues' (2017) survey of school-, district-, and state-level administrators, 61% reported having accessed research through the WWC, though access was generally infrequent; it is unclear if these percentages differed based on administrators' roles. Regardless, reports from the EdTech Evidence Exchange survey indicated that access rates have not improved. Only 14% of the entire sample reported accessing research through the WWC, with 15% of association members and 6% of convening attendees reporting doing so (Barton & Tindle, 2019). Interestingly, when disaggregating the data based on PreK-12 roles (a disaggregation that was, unfortunately, not done for many of

the other items), they found that non-teachers were far more likely to report accessing research through the WWC than teachers were (25% vs. 9%, respectively). Together this evidence indicates that only a small percentage of educators use the WWC as a research access point, and furthermore, that access rates have remained low—if not declined some—over time. It also suggests that administrators may be more likely to use the WWC as a research access point than teachers.

**Capacity Change Assumptions.** According to the theory of change, once beneficiaries engage with the WWC's resources, we would hope that they become motivated to use information from the resources in their classrooms or schools. This motivation hinges on at least two conditions—first, that the information is perceived as relevant, and second, that the information is perceived to be trustworthy. Both conditions should not be assumed. What, then, does the evidence suggest about the relevance and/or trustworthiness of the information provided by the WWC?

***Relevance.*** In terms of relevance, three features are critical. We have known since Hemsley-Brown and Sharp's (2003) original review that research meeting educators' information needs is most likely to be used. One commonly-cited information need is timely information—that is, research related to the immediate issues confronted in practice. The matter of timeliness has been noted in other reviews (e.g., Dagenais et al., 2012) as well as in recent large-scale surveys (Penuel et al., 2017). In fact, when asked about the perceived relevance of research, more than half of Penuel and colleagues' respondents agreed with the statement "by the time research is published it is no longer useful to me" (p. 35).

Unfortunately, much of the information made available by the WWC is outdated. This is evident in two places—both of which were noted above. The first is that nearly three-quarters (i.e., 412 out of 593; 70%) of all intervention reports and half (i.e., 12 of 24; 50%) of all practice guides available to beneficiaries via its "Search Publications" webpage (WWC, n.d.-i) were published prior to 2010. Though two of these practice guides have been updated with more recent information, many have not. Second is the observed backlog between when an individual study review occurs and when it is published. This backlog has been observed previously (GAO, 2010)—as well as presently. Indeed, based on data available since 2017, the estimated size of that backlog is an average of 99.5 days; its median size is 54 days. It seems plausible that some educators, after consulting WWC resources, find them to be irrelevant because they fail to provide up-to-date information.

Along with timeliness is the issue of applicability—that is, information that can be applied in practice. In studies where educators express negative attitudes towards educational research, they often mention issues of applicability (for a review, see Schaik et al., 2018). Famously, Grover Whitehurst observed this frustration among practitioners, and in turn, concluded that they "do not want research minutia, or post-modern musings, or philosophy, or theory" (2003, p. 12), but "answers that will enhance the odds that their decisions will be successful" (2003, p. 5). This is why the WWC has adopted the "what works" scope that is has. Analysis of extractable data from its intervention reports suggests, however, that these reports may not be as applicable as was initially hoped. Calculations indicated that half of all outcomes featured in logged intervention reports were based off of zero eligible studies (i.e., 405 of 811; 50%). This means that many

existing intervention reports simply tell readers that "the WWC is unable to draw any research-based conclusions about the effectiveness or ineffectiveness [of this program]" (WWC, n.d.-c). Additionally, for the 406 intervention report outcomes in the extractable data that are based off at least one eligible study, only a small fraction (48 outcomes; 12%) concluded that the intervention in question had a "positive effect." Although 39% (i.e., 158 outcomes) are categorized as having "potentially positive effects," just as many (163 outcomes; 40%) are categorized as having "no discernible effects". As such, there are far fewer conclusions about the effectiveness of various educational interventions available in these intervention reports than may, given the number of published reports, appear to be the case.

Another way the WWC has attempted to make its outputs—and more specifically, its practice guides—more applicable is by pairing them with instructional videos demonstrating how to implement the guides' recommendations in practice. For example, beginning in April 2020, two videos were published as "additional resources" to the practice guide entitled "Foundational Skills to Support Reading for Understanding in Kindergarten Through 3rd Grade" (WWC, n.d.-a). One video shows a teacher using play-based instructional strategies to enact the practice guide's recommendations in the classroom; the other shows a three-step teaching routine to introduce first grade students to academic vocabulary per the WWC's recommendations. The nature of these videos is itself evidence-based, as findings from Barton and Tindle (2019) indicated that educators do, in fact, rate "[d]emonstrations about how to apply research recommendations" as a useful way of introducing them to research. However, after tallying up the number of instructional videos available with each practice guide, it appears that only 4 of the

WWC's 24 practice guides offered these demonstrations. Although we would not expect all practice guides to have accompanying instructional videos (e.g., those related to dropout prevention), additional videos are needed if the applicability—and in turn, the relevance—of the practice guides is to be maximized.

A third issue related to the WWC's relevance concerns beneficiaries' topics and outcomes of interest. This issue is implicated in Farley-Ripple et al.'s (2018) model, which suggests that research is most likely to be used by educators when the problems being researched are the same problems that educators find compelling. Indeed, early evidence from the GAO (2010) evaluation discovered that over half (i.e., 57%) of all district-level respondents said that they would be more likely to use the WWC if it addressed additional topic areas. According, what topic areas would we expect to be most compelling to educators, and in turn, are these topic areas currently being addressed? Barton and Tindle's (2019) work as part of the EdTech Evidence Exchange survey found that the responses of convening attendees and association members converged when asked what topic areas they would want education research to address. Three agreed-upon topic areas included (a) special education, (b) diversity, inclusion, and acceptance concerns, and (c) technology's effect on students. Given these reported interests, the WWC seems to be underprioritizing and/or neglecting several areas of research. It does maintain a "Children and Youth with Disabilities" topic area, however only 101 of the 1,384 individual study reviews meeting WWC standards (i.e., 7%) were categorized as part of that topic area; in comparison, 542 (i.e., 39%) were part of the "Literacy" topic area. The WWC has yet to create a dedicated topic area related to technology, though it did conduct a rapid review of distance learning research in response to the COVID-19

pandemic (e.g., Sahni et al., 2021); in that review, only 15 studies meeting WWC standards were considered. Finally, the WWC does not have a dedicated topic area related to diversity, equity, and inclusion. Thus, it seems plausible that, after accessing WWC outputs, some users would find them to overlook the topic areas that they are seeking research on.

What about outcomes of interest? Again, the evidence suggests that some outcomes of importance to educators are seldom featured in intervention reports. Specifically, Barton and Tindle (2019) identified several outcomes that were of interest to both of their subsamples, including (a) student achievement, (b) student engagement, and (c) social-emotional support. Outcomes related to student achievement are regularly featured in intervention reports; for example, 42 of the 406 outcomes (i.e., 10%) based on at least one eligible study were related to general mathematics achievement alone. Engagement, however, is featured far less frequently, with only 1 of the 406 outcomes (i.e., <1%) relating to "school engagement." Likewise, only 4 of the 406 outcomes (i.e., 1%) were categorized as related to "social-emotional development." Taken together, it appears as if the outcomes prioritized by the WWC are not always the same as those pursued by educators.

***Trustworthiness.*** A second condition necessary for motivation is trustworthiness. Information featured in the WWC outputs must be perceived as credible if educators are to become motivated to use it for decision-making. Two features seem to dictate how credible educators find research. The first are educators' perceptions of a study's validity. As noted in Cain's review (2016), some extant literature suggests that educators' understanding of educational research conformed with "conventional scientific ideas"

that prioritize internal validity. Specifically, educators in that literature perceived quantification, experimentation (e.g., the inclusion of control variables), and large sample sizes as necessary components of educational research. For educators with such an understanding, the WWC's review standards—which privilege these same features—might elicit trust. Looking closer, however, these same features might also elicit skepticism. Consider, for instance, the average sample size included in intervention report outcomes. Though their mean sample size is sizable (i.e., 8,651), their median sample size is far less impressive (i.e., 282). Another troubling observation, which might concern these same educators interested in matters of internal validity, is that over half of intervention report outcomes featuring at least one eligible study (248 out of 406; 62%) are based off of a single study; 83% are based on two or less.

The small number of studies included in each intervention report is attributable to, among other things, its strict standards of evidence. As noted in the GAO (2010) evaluation, of the 2,669 studies reviewed by the WWC at that time, 92% were screened out for not meeting its standards—only leaving 226 to be synthesized. To be sure, this has improved over time, albeit only slightly. Of the 15,124 study reviews recorded in the WWC's database, 10,597 (i.e., 70%) resulted in the study being deemed "ineligible for review" or "not rated;" an additional 1,935 were said to "not meet WWC standards." While such stringent standards might enhance the WWC credibility, the GAO still found that the:

> "Researchers and education professionals we interviewed suggested that the WWC produces limited information because its screening criteria are too restrictive." (p. 13)

In addition, for those educators who, like the WWC, see value in these restrictions, they may actually be disappointed to find a sizable proportion of these included studies to incompletely meet the WWC's standards. Specifically, analysis of the WWC extractable data showed that over a third (i.e., 36%; 494 out of 1,384) of study reviews that resulted in the study being approved were qualified with the rating "Meets WWC standards with reservations." This suggests that many of the conclusions reached in WWC intervention reports are themselves based on studies that "provide a lower degree of confidence that an observed effect was caused by the intervention" (WWC, n.d.-c).

Despite using such restrictive standards when screening research for its intervention reports, the WWC is far less restrictive when developing its practice guides. Specifically, as initially pointed out by the GAO (2010), the WWC relaxed its standards of evidence when incorporating evidence into practice guides—meaning that these guides "…also incorporate studies that do not have designs that are eligible for WWC review, or in some cases, are reviewed and do not meet WWC evidence standards" (p. 15). As such, the recommendations featured in the WWC's practice guides are based off of a different evidence base than the conclusions raised in its intervention reports. When examining these recommendations more closely, my analysis revealed that 44% (i.e., 53 of 120) of all recommendations are categorized as having a "minimal" evidence rating. This rating means that there is "no consistent evidence that demonstrates the practices' positive effects, because it has not been studied or there is weak or conflicting evidence of effectiveness" (WWC, 2017a). This, too, might give educators pause, especially those who subscribe to the same standards of internal validity as the WWC does when reviewing individual studies.

At the same time, the extant literature suggests that issues of external validity are also important to educators, if not more so, when they engage with research (see Dagenais et al., 2012). More precisely, educators' perceptions of a study's transferability—from the context in which was conducted to the context inhabited by the educator—is critical, especially for teachers (Leko et al., 2019). Teachers desire similarities between a study's setting and their own, as well as similarities between the study's sample and the characteristics of their own students (e.g., Cain, 2017; Joram et al., 2020). The WWC has made deliberate changes to their website to help beneficiaries find research conducted with "students like yours" (IES, 2018). Despite these changes, the research featured in WWC resources seems to be lacking in information required for teachers to make these comparisons. Though most intervention report outcomes based on eligible studies include demographic data about the genders (335 out of 406; 83%) and races/ethnicities (277 out of 406; 68%) of the students sampled, information about other demographic characteristics is missing. For instance, only about 50% (i.e., 205 out of 406) of reports feature data about students' free/reduced lunch status and fewer than that (168 out of 406; 41%) report on the number of ELL students in their sample(s). Likewise, although information about type of school in which a study was conducted (i.e., public, private, charter, or parochial) is generally available, information about that school's urbanicity (i.e., rural, suburban, or urban) is often missing. In fact, about 41% (168 out of 406) of all logged outcomes do not feature this contextualizing information. Educators may, in turn, be hesitant to trust the conclusions reached in the intervention report.

A second feature related to the educators' perceptions of trustworthiness has less to do with the research information itself. Instead, it is related to where that information

comes from. As already noted, one of educators' most common research access points is their colleagues (e.g., Barton & Tindle, 2019; GAO, 2010; Penuel et al., 2017). This may be because they are perceived as more trustworthy than sources that are less familiar with the day-to-day challenges of educational practice (e.g., Schaik et al., 2018). Perhaps for this reason, the WWC has a system in place to try and account for educators' realities in their dissemination of research information. Specifically, as part of their practice guide development teams, the WWC includes two practitioners "who have backgrounds that allow them to offer guidance about implementation of the recommendations" (WWC, 2020d, p. C-2). Based on information from the WWC website, however, it appears as if this system is not being fully realized. Although all but one of the practice guide development teams (excluding practice guides related to postsecondary outcomes; see Chapter 3) include someone in a state-, district-, or school-level administrator or teaching position, only four include *two* individuals with those affiliations. Thus, educators might be skeptical of most practice guides and their recommendations, as they may be perceived as coming from researchers rather than fellow practitioners.

**Capacity Change.** At present, I am unable to find any indicators of beneficiaries' motivation to use the information featured in WWC resources once they engage with them. If motivation can be implied based on conditions of relevance and trustworthiness mentioned above, then this motivation is seldom felt by teachers.

**Behavior Change Assumptions.** There are certainly cases in which the WWC's resources reach intended beneficiaries, who, in turn, become motivated to make decisions based on what they have learned. This motivation is expected to, in turn, result in beneficiaries' use of the WWC's resources—or, more specifically, the information

therein—to inform their decision-making. Doing so is an example of the WWC being used instrumentally—that is, to directly inform actions taken to solve a target problem. For this to occur, however, requires that at least one assumption is met.

***Instrumental use is possible*.** Even when an educator is motivated to use research evidence instrumentally, they often cannot because of various external barriers. The existence of both structural and cultural barriers to research use were reviewed in Chapter 2, including lack of time, lack of agency, and an institutional culture that discourages research use (for reviews, see Dagenais et al., 2012; Schaik et al., 2018). An interesting quality of this assumption is that it exists largely outside of the WWC's control. Although the WWC can make its resources short and digestible, it is unable to affect the amount of time educators are given to consume research. Likewise, the WWC has no control over educators' institutional cultures. Therefore, given this lack of control, as well as how prominent these structural and institutional factors are in models of research use—including the model put forward by Farley-Ripple and colleagues (2018)—this assumption is refutable.

**Behavior Change.** The WWC's ultimate goal is a goal of behavior change—that is, the goal of increasing research use in education decision-making. In particular, its goal is that educators use research-based information provided in WWC resources to inform their decision-making. Some evidence, mostly collected as part of previous evaluations, suggests that some administrators and teachers do report using the WWC's resources instrumentally, though the extent of that usage remains underwhelming. To illustrate, the GAO (2010) evaluation asked all district-level administrators who had reported accessing the WWC (34% of the entire sample) to elaborate on their usage of it. Specifically, they

were asked about the extent to which they used information in the WWC to (a) inform the professional development of teachers, (b) intervene with poorly performing schools, (c) develop school improvement plans, and (d) inform curricular decisions. Most did report using WWC information for these purposes. Even so, across these four usage options, the GAO estimated that three-quarters of respondents (i.e., 72%) had used the WWC to inform their education decision-making to a “small” or “moderate” extent; in contrast, only 18% reported usage that could be categorized as “large” or “very large”. Thus, though administrators do use the WWC in instrumental ways, they do so to a lesser extent than the IES might wish.

More recently, the EdTech Evidence Exchange survey found that usage might also differ based on the type of educator examined. When considering the entirety of the sample, 29% had reported using information from the WWC. Upon closer examination, however, 32% of association members reported using the WWC, whereas only 11% of convening attendees did so (Barton & Tindle, 2019). These percentages require some qualification, however. Earlier in the survey, only 14% of respondents had said that they used the WWC as a research access point; how, then, can reported usage be greater than reported access? The response options included in the survey are to blame. Specifically, a respondent was considered to have "used" the WWC if they responded in any of the three following ways: “3 = Someone else uses this to gather research information I use; 4 = I use this to gather research information and it is easy to use; or 5 = I use this to gather research information but it is hard to use." Given that one of these categories implicated “someone else” accessing WWC information on behalf of the respondent, these percentages are tricky to interpret.

### *Intended Size of Contribution*

It is important to consider these findings in light of the PART goal set by the IES in the early 2000s. Doing so allows us to determine if educators' usage of the WWC does, in fact, demonstrate that the WWC is having a meaningful contribution. The goal was that 25% of educators would use the WWC to inform their decision-making by 2014; by extension, I estimated that at least 37.5% of educators would need to use the WWC as part of their decision-making process if the WWC's contribution can be deemed "meaningful." As can be seen in the above findings, surveys of administrators and teachers—both in 2010 as well as in 2019—suggest that neither group's usage was ubiquitous enough to meet the adjusted benchmark. In addition, an implicit assumption embedded within this goal is that educators use the WWC repeatedly. In other words, even if 37.5% of educators did report using the WWC to guide their decision-making, could we really conclude that the WWC had contributed to educators' research usage if they had only used it once? Evidence from both the GAO and EdTech Evidence Exchange evaluations suggest that educators' frequency of usage is low. For example, although the GAO (2010) survey found that 34% of school districts reported accessing the WWC, only 11% said they had done so at least seven times that year. Likewise, although the EdTech Evidence Exchange survey did not inquire into respondents' usage frequency of IES research access points, it did find that those who reported accessing research via the formats used by the WWC—such as "reports" or "executive summaries"—tended to only do so once or twice a year (Barton & Tindle, 2019). Taken together, this evidence refutes the notion that educators are accessing the WWC with enough regularity to constitute a meaningful contribution.

### *The Contribution Claim*

After an initial review of the evidence has occurred, Mayne (2012b, 2019) recommends that a contribution claim be made. In essence, a contribution claim is a statement "about whether the intervention made a difference as expected" (Mayne, 2012b, p. 273). Instead of answering with a simple "yes" or "no", Mayne (2019) has more recently recommended that the contribution claim also highlight "[h]ow and why has the intervention (or component) made a difference, or not, and for whom" (p. 175). Given these recommendations, a three-part contribution claim can be made based on the available evidence. First, and most generally, the WWC does not seem to be making a meaningful contribution to the IES's goal of "increasing [the] use of data and research in education decision-making" (IES, n.d.-a), nor does it appear to be fulfilling its own mission of being "...a central and trusted source of scientific evidence for what works in education" (WWC, 2020d, p. 1). Not only is the WWC failing to adequately reach potential users, but even when it does, the percentage of educators who report using the WWC to inform their decision-making is smaller than anticipated; the extent of their usage is also wanting.

Second, and more specifically, the reasons why the WWC is failing to make a meaningful difference may be due to the nature of its assumptions. Among the WWC's causal assumptions, its assumptions of reach and capacity change appear most fragile. In terms of reach, there appears to be some lack of awareness among teachers about the WWC and its resources. For those who are aware, they may lack the willingness to seek it out because the WWC fails to utilize dissemination strategies that educators prefer. Furthermore, even those who are interested in accessing the WWC often face barriers to

access that the WWC cannot account for. In terms of capacity change, there are numerous reasons why the WWC's outputs—and the information featured therein—may not be perceived as trustworthy or relevant by educators. Among the WWC's rationale assumptions, the prescriptive assumptions informing its systematic review approach in the first place, are most refutable.

Third, there was some evidence suggesting that the WWC's contribution may be more robust among some educators than others. Administrators' awareness of the WWC, as well as their eventual usage, appears to be greater than teachers'. Even so, there appears to be some heterogeneity among administrators as well, which suggests that state- and district-level officials are more engaged with the WWC than principals and other school-level administrators. These observations were helpful when deciding on additional evidence sources to pursue in Step 5.

Before proceeding with Step 5, though, it may be helpful to direct the reader to Table 4.4. Table 4.4 summarizes the extent to which each component of the WWC's theory of change is supported by the evidence examined in Step 4 (and later in Step 6). Specifically, the table demonstrates what evidence sources were considered when evaluating each component, and moreover, if those evidence sources provided confirmatory, mixed, or refuting evidence for the given component. Differences across stakeholder groups (i.e., across teachers and administrators) are represented as well. The table is also intended to help the reader quickly identify where the WWC's theory of change is most limited. For example, there is an abundance of light gray squares indicating refuting evidence in rows corresponding with the WWC's reach assumptions

and capacity change assumptions. This suggests, as noted above, that the WWC's lack of impact may be at least partially due to some faulty programmatic assumptions.

### Step 5: Gathering Additional Evidence

Given the nature of the contribution claim, as well as the evidence supporting it, Mayne's (2008, 2012b, 2019) instructions to gather additional evidence in Step 5 seemed prudent. Indeed, the contribution claim suggests that school-level practitioners—particularly teachers—engaged with the WWC less than administrators at the district and state levels (e.g., GAO, 2010). Admittedly, however, some of the evidence available on educators' WWC usage failed to include teachers in the first place (e.g., Penuel et al., 2017). Likewise, although the contribution claim implicated some of the WWC's causal and rationale assumptions as empirically unsound, much of this evidence comes solely from the literature. Will other sources of evidence suggest similarly?

I approached Step 5 with two goals. The first goal was to better understand how teachers engage with the WWC. There are several reasons why gathering additional evidence related to teacher engagement could help strengthen my contribution story. One reason, as alluded to above, is the lack of available evidence about their current engagement. Several of the evidence sources cited extensively in Step 4—namely, the GAO's (2010) evaluation of the WWC and Penuel et al.'s (2017) large-scale survey of research use—reported findings from administrators but not from teachers. Although literature reviews of research use (Cain, 2016; Schaik et al., 2018) have generally included more studies of teachers than administrators, the reverse seems to be true when considering WWC usage specifically. Because we currently know more about how

administrators engage with the WWC than how teachers do, the contribution claim could be enhanced by studying teachers' engagement further.

A second rationale for further investigating teachers' WWC engagement—and particularly the engagement of younger, less experienced teachers—is that most existing evidence sources have sampled teachers with extensive backgrounds in education. For example, educators participating in Barton and Tindle's (2019) survey reported having worked in education for an average of 21 years. Similarly, the teachers in Lysenko et al's (2016) survey of Canadian teachers had an average of 12 years of experience. Knowing how less-experienced teachers engage with the WWC is important because it might better-position the WWC to make a meaningful contribution in the future. For example, these educators have had less time to establish routines in how they access educational research, so they may be more open to trying new research access points than their more seasoned colleagues. Also, because these teachers are likely to remain in the education sector for longer than those who have already taught for several years, engaging them might be more conducive to making a lasting contribution.

A third reason why gathering additional evidence on teachers' use is important is because the WWC itself is trying to reach them. One example is its increased focus on practice guides. Though not all practice guides are specific to teachers, many are. Mark Schneider, the current director of IES, has made retooling practice guides one of his primary goals, because "they are central to translating research into practice" (Schneider, 2019). He has even committed significant financial resources to practice guides, making them "two-year, million-dollar endeavors" (Schneider, 2019). Another example can be found in how IES has tried to solicit teacher input through its partnership with the

EdTech Evidence Exchange (i.e., Barton & Tindle, 2019). Part of that partnership involved a "listening tour" intending to hear from teachers about their uses and perceptions of educational research (Sparks, 2018). So, one more reason to gather additional evidence on teachers' engagement with the WWC is to maximize this project's relevance to the WWC and its current ambitions.

Along with gathering additional evidence about teachers' WWC engagement, my approach to Step 5 also involved further investigation into the WWC's assumptions. Focusing on these assumptions was necessary given how the current contribution story implicated them as a reason why the WWC has failed to maximize its impact. I was especially eager to find applicable qualitative data, as most of the WWC-specific evidence reviewed above is quantitative in nature. Open-ended responses and/or discussions involving the intended beneficiaries of the WWC could demonstrate the degree to which the WWC's assumptions align with the assumptions that WWC users—especially teachers—bring to educational research.

Accordingly, I analyzed two additional sources of evidence, including (a) a set of focus group transcripts provided by the U.S. Department of Education, and (b) a survey of preservice teachers conducted specifically for this project. Both sources are described in Chapter 3.

### Step 6: Updating the Contribution Claim

In light of this additional evidence, should the contribution claim concluding Step 4 be updated? Or does this new evidence simply bolster the claims already made? The objective of Step 6 was to determine if insights from the survey and focus group data justify modifying and/or solidifying the contribution story, and in turn, the resulting

contribution claim. As will be described below, these additional data sources both corroborate and complicate the conclusions reached in Step 4, especially those related to the WWC's rationale and causal assumptions.

***Normative Assumptions***

An observation emerging out of the Mathematica focus groups is that researchers and administrators were keenly aware of a research-practice gap, and moreover, forthright about it being a problem. For instance, during one of the in-person focus groups, a respondent vocalized their awareness of a gap as well as their belief that closing it was desirable:

> "But the issue I still have is there such a gap between people's understanding of research and appropriate use of it. I mean, it's something I know that REL and everybody, it's a continuous challenge and I have it when I'm out there. I mean, if there can be resources developed that help support the logic in some way."

In response, a fellow focus group member similarly acknowledged the presence of a gap. In fact, they hoped that the WWC could do something about it, urging that:

> "…it's still a gap. I don't know if you [the WWC] can fix it. But if you can fix it, please do."

During one of the virtual focus groups, a researcher posed many of the same questions that, in effect, motivated the WWC's creation:

> "More and more, I realize that an important link may be missing that brings best practice to the classroom. When we have so many solid resources, why isn't practice changing? Where is the disconnect? How can we bring the research into the hands of those that directly touch the children? How do those folks have time to learn about these resources? What structures are missing to enable connection between research on best practice and the practitioners?"

Interestingly, the preservice teachers seldom recognized the presence of a gap. When defining "research" (in the context of education) as part of my survey, however, many implied that their practice should be guided by research. Indeed, one of the themes

emerging in their responses was that research was meant to guide instruction. For example, one teacher offered the following definition:

> Research means finding information whether that is through observation or online research that informs your practice in some way.

For another, educational research involved:

> Finding out new ways and techniques to teach your students and to get them more engaged.

A third suggested similarly:

> Research is finding the best way to teach future students.

Still, a fourth conceptualized research as:

> Collecting data to improve instruction. Searching for different tools to use in the classroom.

One went as far as to say that:

> Research is how we improve education and shape the future. Without research, we would not change or improve.

In sum, these responses strengthen the WWC's normative assumption that "research should guide practice" by suggesting that many educators believe similarly.

***Diagnostic Assumptions***

As a reminder, the WWC was built upon a diagnosis of the research-practice gap as the result of deficiencies in the supply and demand for educational research. Not only was educational research condemned as being low quality, inaccessible, and overly theoretical, but users of this research—especially teachers—were thought to be oblivious and ill-equipped to consume it.

Just as evidence in Step 3 did, the additional evidence gathered in Step 5 complicates this conclusion. It does so by demonstrating that teachers—even those in the

beginning stages of their career—are actually quite sure about the relevance of educational research and their ability to consume it. For example, over 80% of survey respondents disagreed or strongly disagreed with the statement "I find most educational research to be irrelevant to practice." Likewise, close to 70% rejected the statement that "It is challenging for me to determine whether or not research is of high-quality." These response patterns suggest that these teachers do not find educational research to be irrelevant, nor do they feel ill-equipped to vet its quality themselves. Even so, it is worth noting that responses to the statement "I rarely come across research findings that are presented in an easy-to-understand way" were more ambiguous, with 48% of respondents disagreeing and 53% of respondents agreeing. This suggests that teachers do, however, find research to be presented in inaccessible ways—a finding that aligns with the WWC's own diagnostic assumptions.

Unfortunately, because the focus groups had more to do with specific WWC products and less about the state of educational research generally, little could be gleaned about teachers' and administrators' beliefs about the causes of a research-practice divide. Some participants in the "Academic/Researcher" focus group alluded to the diagnostic assumptions of the WWC. Several comments were made suggesting that they, too, find both the supply of and demand for educational research to be inadequate. On the demand side, one respondent echoed the concerns of G. Reid Lyon (see Step 2) about teachers being unable to vet research themselves. Not only did they question educators' awareness of the WWC, but they also questioned their ability to vet and/or find research:

> "Sadly, most teachers have never heard of WWC and don't know how to go about finding reliable information."

On the supply side, another implied that educational research is sometimes lacking in quality, and, in turn, that the WWC could play a role in improving it:

> "I thnk [sic] the WWC could do more to help improve the quality of research in particular areas by providing guidance to researchers, policymakers, and funders about where the deficits are greatest."

While researchers and academics might harbor concerns about the supply or demand of educational research, teachers who responded to the survey seemed not to.

***Prescriptive Assumptions***

The two-part prescriptive assumption on which the WWC is based contends that the causes of the research-practice gap are most appropriately addressed through (a) increased government control, and (b) a form of systematic research review. Little additional evidence was relevant to the former. Even so, the idea that systematic research review would be the best way of addressing the gap was indirectly mentioned in both the survey and the focus groups. Based on that evidence, the assumption enjoyed mixed support. In the survey, respondents enthusiastically endorsed the statement that "The best way to share research with educators is by publishing free, easy-to-understand reports." Not a single respondent strongly disagreed, and despite the disagreement of 14%, the remaining 86% of the sample either agreed or strongly disagreed. Given that the results of systematic research reviews, particularly those pertaining to educational programs and practices, are disseminated through these types of short, written reports, there is some support for the belief that this type of strategy would be a good way of bridging the research-practice gap. That support, however, can only be cautiously extended to the concept of systematic review itself.

Similarly, although focus group respondents never named the systematic review process explicitly, they did imply that the usual ways of disseminating information from a systematic review are better-suited for researchers and/or administrators compared to teachers. In response to a question about how the WWC "should get the word out", one participant believed that this type of dissemination was shortsighted:

> "The bridge between the research and practice probably will not happen easily by simply asking teachers to read the guides…I think this is one of my great frustrations. There are so many resources available, many for free, resources that would greatly improve the educational experience of our children. The task is connecting the resources with educators. Researchers are using the resources, I believe. We need to make sure that educators and preservice educators receive the resources in a usable way."

In a similar way, a participant from the in-person focus groups suggested that certain types of WWC outputs that more faithfully represent the systematic review process (i.e., intervention reports) are more appealing to administrators than teachers. Teachers, on the other hand, desire something like a practice guide, which, as explained earlier, is not a direct byproduct of the systematic review process:

> "I think going back to your question about audience, I think each tool has a different audience. I think the intervention reports are going to be looked at much more at the district of the building level. But I think the practice guides really are more of a practitioner. That's the teacher level. Looking at what is it that that I can do in my classroom. Certainly building level as well, but I think there's an audience of teachers for those practice guides."

Ultimately, more information is still needed to confidently conclude whether or not educators hold the same prescriptive assumptions as the WWC.

***Reach Assumptions***

Reach assumptions specify the conditions necessary for the WWC's outputs to reach beneficiaries. Three conditions are specified in Step 3's theory of change, including (a) beneficiaries' awareness of outputs, (b) beneficiaries' willingness to seek them out,

and (c) beneficiaries' ability to seek them out. In Step 4, evidence largely determined these assumptions to be ill-founded. Much of the additional evidence corroborates this conclusion.

**Awareness of outputs.** Both the survey responses and focus group discussions suggest that teachers are largely unaware of the WWC. In the survey, only one-fifth of all respondents (i.e., 21 out of 105; 20%) reporting having heard of the WWC; all 21 reported that they had heard about it from professors in their college coursework. This finding is somewhat congruent with other estimates of WWC awareness, such as those from the EdTech Evidence Exchange survey (2019), which found that 36% of their convening attendees had heard of the WWC. In the focus groups, 90% of discussants in the Teacher focus group and 100% of the School/District Administrators focus group had heard of the WWC, but this was unsurprising given recruitment procedures (e.g., contacting educators who had subscribed to the WWC listserv). Despite this self-reported awareness, lack of awareness in the general teacher population was a common conversation point across the focus groups. Those in the Researchers/Academics group noted how researchers were more aware of the WWC than teachers were:

> "Continue to communicate with educators by whatever electrontic [sic] methods you can. I think this is one of my great frustrations. There are so many resources available, many for free, resources that would greatly improve the educational experience of our children. The task is connecting the resources with educators. Researchers are using the resources, I believe. We need to make sure that educators and preservice educators receive the resources in a usable way."

> "Sadly, most teachers have never heard of WWC and don't know how to go about finding reliable information. Everyone says their stuff is research-based, and WWC (and all other excellent programs) are competing with publishers for teachers' attention…"

Similarly, teachers themselves alluded to this lack of awareness. In response to a question about improving the WWC's dissemination practice, one teacher suggested "…advertising that you have these products -- don't think teachers know." Later in the discussion, when shown a video overviewing the WWC's resources, another teacher admitted that "[they] did not know these services existed." An interesting moment occurred in one of the in-person focus groups when the facilitator themselves acknowledged that teachers in their life were not aware of the WWC:

> "I would say my mom had no — my mom's been teaching for 15 years now — and she had no idea until I started working for Mathematica. And she's like, "What are you doing?" And then now her school knows, her principal knows because — but without that personal connection there's no [inaudible 01:15:34]."

With this assumption, responses in both the surveys and focus groups again converged on the conclusion that teachers are largely unaware of the WWC. Thus, the additional evidence only strengthens what was concluded in the initial analysis.

**Ability to seek out outputs.** The new sources of evidence also aligned with previous evidence when suggesting that time constraints impede teachers' ability to seek out educational research. Constraints on teachers' time were mentioned in the focus groups, and some survey responses spoke of those constraints as well. Although none of the survey items explicitly asked about time constraints, one respondent explained in an open-ended response that they had not used the WWC because "I have very little time to add in any other resources." This was similarly acknowledged by state leaders, administrators, and teachers participating in the focus groups. When asked about for their opinions about a new WWC resource, participants stated the following:

> "As a PowerPoint I might share it with the elementary teachers I teach; however, for them to print it may be a waste of time, paper, and ink. They are swamped and have minimal time to refer to a guide, pamphlet, or notebook of information. They

> need easy access, straight to the point, easy to read and jot down. Please no color unless it prints with clarity in grayscale."
>
> "The shorter the better. There was probably some things that could be left out to keep it shorter. Teachers and education personnel need information quickly, with as little fluff and possible."
>
> "Again conciseness is important -- give the information in the most concise and practical way for teachers who have limits on time."

In general, focus group participants applauded the WWC for developing resources that were "quick and fast" or "short, sweet, and to the point." Yet, this does not change the fact that teachers may have little time to actively pursue these resources.

What's more, a few of the focus group conversations hinted at the issue of active versus passive dissemination strategies (e.g., Levin, 2011), which was previously implicated in Step 4. As a reminder, the WWC largely publishes its outputs in formats that beneficiaries must seek out themselves (i.e., passive dissemination). This passive dissemination strategy, along with the time constraints faced by teachers and administrators, means that the WWC's assumption about stakeholders' ability to reach their resources is largely unfounded. Focus group participants came to this conclusion as well. For example, one participant posed the following question:

> "…there's so many things bombarding us that are research based now, from different policy think tanks and so on that are active bombardiers or us — this is a question, not a statement — is What Works in like attack mode or defensive mode? Do we have to go find it or does it come at us? And [inaudible 00:19:00] all, we have to go find it. And I'm just— well I'm struck by this. Not the clarity, but it's kind of how do you get to it. How does it connect to people that might benefit from it?"

Answers to these questions were never offered, yet it became clear in subsequent conversations that participants desired a more active, "attack mode" approach. Furthermore, the inadequacy of passive dissemination was acknowledged by one of the

Mathematica facilitators as well. During one of the in-person discussions, the facilitator goes off-script and acknowledges the following:

> "That's another question. Is that another area of dissemination that the clearinghouse should seek? I have to say when the clearinghouse first started, they had a Field of Dreams attitude towards the clearinghouse. And so over the last ten or so years, we've been trying to change that, and let them know that just because it's there people aren't going to show up magically. So this dissemination piece is pretty critical. They stuck with email blasts for a while, and they reached out to Twitter and to Facebook and things like that. So when engaging with Ed Week or places like be a beneficial experience and reach audiences."

Indeed, the WWC has taken steps to be more active in their approach, such as the utilization of social media. Yet many of their primary outputs, as described elsewhere (e.g., GAO, 2010), continue to ask teachers to "show up magically."

**Willingness to seek out outputs.** The additional evidence paints a similar picture to the one mentioned earlier. Although many educators possess a willingness to seek out educational research generally, many may not have a willingness to seek out WWC outputs specifically. As mentioned above, in my survey of beginning teachers, most respondents rejected the idea that research was irrelevant to practice. In an opened-ended response, a respondent who had previously used the WWC said that they did so because of their eagerness to "use research base [sic] practice in the classroom." Likewise, teachers' focus group responses suggest an eagerness to engage with educational research. Of the six participants in the virtual teacher focus group, all of them answered affirmatively to the question "Do you seek out research-based information related to education?" Some of their responses included the following:

> "Yes, I seek out the information I need to use as research evidence for best practice. I typically begin by researching with search terms on the internet. Those serches [sic] usually lead to other reading of others research, or practice."

> "Yes, I am constantly seeking research concerning education. I often "google" topics for information. I have also joined several online learning communities that email frequently with available information through documents, webinars, on demand videos, and the like."

Together this data suggests that many teachers are, indeed, willing to consult educational research.

Even so, this willingness may not translate to WWC outputs specifically. One reoccurring explanation for why teachers may not have a similar willingness to engage with the WWC's resources is that they already have their own preferred ones. Of the 21 respondents who had heard of the WWC, five had never used it. When asked why this was the case, two mentioned consulting "other resources first." A similar sentiment was voiced by the 16 respondents who reported having both heard of and used the WWC. Four of those respondents stated that the WWC was their "primary resource for information about research-based educational practices," but the others reported using other sources as well.

It is thus important to ask why other information sources might be more central to teachers' research access than the WWC itself. One reason emerging in both data sources is that the WWC has failed to build relationships with teachers, and moreover, it does not adequately utilize formats that encourage relationship-building. This idea emerged in one of the in-person focus groups, when a respondent offered the following observation:

> "You know may be too simple thinking on my part, but you know we talk so much in school about how important it is to build relationships with kids. I think it's the same thing with this — and just like you're talking about Pinterest. They've built a relationship. They've proven how friendly and supportive and good they can be — and helpful. I think, as you make those connections, whether it's with an organization or faculty, somebody starts getting this stuff out to you."

In a similar way, some survey respondents discussed how their teaching practices were more often informed by "strategies that my professors have taught me" or "practices that have been taught to me in my classes" than those recommended in WWC resources. This may be because they were shared directly through an existing professor-student relationship. Interestingly, evidence emerged suggesting that virtual communities help facilitate this relationship-building. Focus group participants suggested that the WWC reach teachers through Pinterest:

> "…all my teacher friends are on Pinterest and they're Yeah. And they're constantly posting stuff to do in their classrooms. It's incredible if got on my Pinterest feed — whatever thing it is — it's all like things to do in the classroom. So, that one's, I think, a big one. [Cross talking 00:59:07]"
>
> "Do you have a Pinterest board? Teachers love pinterest. :)"
>
> "Featured articles in the NEA (the teachers' national union) magazine, pinterest - by subject matter, youtube by subject matter when possible."

Although the WWC does disseminate through other forms of social media (i.e., Facebook, Twitter), these platforms are often used to share information in a linear, one-directional way rather than as a relational exchange of ideas (e.g., Lavis et al., 2003). Given that both the survey and focus group data suggest that relational exchanges are a popular access point to educational research for teachers, these data only further the original conclusion that teachers may not be willing to seek out WWC resources because they favor other resources instead.

***Capacity Change Assumptions***

Once reached, the WWC's outputs are meant to motivate teachers to alter their practice in accordance with educational research. Just because WWC resources reach beneficiaries, however, does not mean they will be motivated to change their behavior.

Two general assumptions, which Mayne (2015, 2017) calls "capacity change assumptions," need to be met. In the WWC's theory of change, these assumptions specify that outputs must be perceived as relevant and trustworthy. These two conditions may even be contingent on one another (e.g., Cain, 2016), such that the relevance of a study is used by teachers as an indicator of credibility. The evidence summarized before suggested that teachers are unlikely to find the WWC's resources to be relevant or credible, and the additional evidence suggests similarly.

**Outputs are perceived as relevant.** One way this is evident is in how the research interests of teachers in both the surveys and focus groups diverge from those typically featured in WWC resources. The survey did not ask about respondents' specific topical interests, but it did pose several questions about the types of information respondents find most relevant. Although nearly half (42%) of respondents agreed or strongly agreed that research examining if an intervention "worked" was more relevant to them than research exploring "why" or "how" it worked, 58% disagreed or strongly disagreed. This pattern of responses suggests that the "what works" focus of the WWC is well-aligned to the interests of some, but not all, teachers.

Responses to other items more dramatically demonstrate the divergence between teachers' interests and the WWC's offerings. For example, nearly three-quarters (i.e., 72%) of respondents disagreed or strongly disagreed that researching the desired benefits of an intervention was more important than researching its unintended consequences. Finally, 46% of respondents disagreed and 31% strongly disagreed that interventions aimed at improving students' test scores mattered more than interventions focusing on other outcomes. This latter finding suggests that one of the most common outcome

variables in research featured by the WWC—that is, the outcome of student achievement—may be of less interest to teachers than other types of outcomes.

In turn, the focus group data offered additional insights into the research questions and topic areas that teachers are most interested in. Again, the focus groups demonstrated that the "what works" question is not necessarily off-base. When asked what kinds of questions classroom teachers are trying to find answers to, the in-person focus group participants responded as follows:

> "I was just going to say, how can I help my students improve — what's here that I can do to make a difference in the classroom to make kids improve, get better?"
>
> "They want the answer to what works [laughter]."
>
> "I'll just say that I know that my interaction with the districts, in particular, I work with district context. When they're going to adopt new interventions, particularly, it seems to be a big question. They want to know what works best…"
>
> "I think it's both for the teachers to know what works but also to know how to voice and respond to things that are coming down into their classroom. That they need a way to assess is this the right fit for my students."

This final excerpt is notable because it suggests that the "what works" question may only approximate teachers' true concerns; in other words, it does not encompass the intricacies inherent to the research questions teachers are often seeking answers to. Just as the above responses suggested that teachers want to know if a program or practice is "the right fit for my students," other respondents believed this was the case as well, especially given that teachers often work with students from diverse backgrounds and with a wide range of ability levels:

> "We've been trying to figure it out for a long time, right. So not just what works, but for whom? Under what conditions? Because what we've found is everything works to some degree. Well, most of it. Probably some of them don't. But I think that really is what and by what we see around the country, districts are struggling with. They're trying to find a program. Some are still there. And others are really

> saying, "This works for some students, but we don't know what to do for the others."
>
> "Well, one of the things that I know my teachers look for is that we have kids at so many different ability levels in a classroom. So, as a district, we've been trying to work on assisting with differentiated instruction, but really helping with groups and such — and such like that — and I know that the teachers look all the time for ways to make that run smoother and see the data that's showing that it's working."
>
> "Do you have stuff for special education? Because that population is really growing in our state and in our schools. And I think getting more resources to our staff— not only our special-ed teachers, but the classroom teachers as well."
>
> "We're getting a lot of Hispanic students in our state as well and at our school and so we're struggling to even get enough ELL teachers, and then how to teach students who are not English speaking when we have little support. Well, the support we have — I mean, the teachers we have — we should have more and we just can't find them."

As noted in Step 4, although the WWC does feature topic areas on "Children and Youth with Disabilities" and "English Learners," the number of studies reviewed in each of those areas is far smaller than those in others (e.g., "Literacy" or "Mathematics"). Thus, teachers who consult the WWC with a more nuanced question than "what works" may find little information of relevance.

Similarly, they may be disappointed to find the WWC offers little support for teachers interested in researching these questions themselves. Indeed, throughout the Teacher focus group, participants implied that one of the topic areas they wished the WWC featured was not a content area per se; rather, it was a desire for guidance in the research process. In response to focus group questions such as "What other topics would most interest you?" or "Do you have any suggestions for other video topics?," teachers offered responses such as the following:

> "Forms that help school staff begin to record the information they need toward progress monitoring academic success, or behavioral change."

"Reading and interpreting data."

Ultimately, these responses suggest that teachers may wish the WWC to behave less like a data repository and more like a professional development tool for building research capacity within their own schools.

The WWC's relevance to educators is not only dependent on their topic areas of interest. As suggested in Step 4, the relevance of information to teachers also depends on its format, and more specifically, its ability to be used in practice. If they have little application to practice, the WWC's outputs will be unlikely to motivate teachers to change their practice in ways that are evidence-based. There are two ways in which the focus group discussions further this claim. The first is that participants repeated the need for WWC resources to show—not just tell—teachers how research findings can be implemented in practice. As an example, though most practice guides make written recommendations for teachers, they rarely demonstrate what they look like when applied properly. Such demonstrations were requested across the focus groups:

> "As said previously, links to vidoes [sic] showing actual implementation in a classroom, exemplary lesson plans or stories of how teachers use the strategies in the classroom. Anything that helps teachers "see" how they can be put into action in a classroom. There is reading about it and seeing it. Since teachers are so isolated in their own classrooms, it is hard for them to view other teachers in action and anything we can do to show them ways to teach improves instruction."

> "…video demonstrations or video modeling is always helpful."

> "Teachers love to see practices in action. Recording live data meetings or interviewing practitioners as they go through this process might be helpful. But really, any videos actually showing the recommendations (from the practice guides) in action would be great!"

There is evidence that, following the completion of these focus groups, the WWC made efforts to supplement their practice guide recommendations with video demonstrations.

In fact, even for practice guides that were released several years ago, the WWC has since added videos to accompany written content. However, few videos of this nature currently exist (see Step 4). Accordingly, although the WWC has likely improved its relevance by adding videos of recommendations in practice, it is still failing to attract teachers compared to other access points that allow teachers to *see* what works (i.e., Pinterest).

A related observation made by focus group participants is that the WWC's resources fail to acknowledge that teachers need more than just knowledge of research-based practices in order to implement them. Additional resources are needed if teachers are to move from *knowing* what works to *doing* what works. Consider these responses by teachers, researchers, and administrators when asked what else the WWC could provide to improve its practice guides:

> "Teacher-friendly materials that go beyond the practice guides. Think about everything a teacher or school would need to carry-out the recommendations. They need planning templates, links to learn more about certain strategies, PD modules, checklists, observation forms or reflection tools to improve their practice, etc."
>
> "…listing specific maniupulatives [sic] that can be utilized [sic]…"
>
> "Links to examples of lessons, materials, videos of teaching practice."
>
> "Sample lesson plans and templates are quite useful."
>
> "…exemplary lesson plans or stories of how teachers use the strategies in the classroom."

Again, the WWC's practice guides are well-intentioned, but their relevance lags behind other information access points that provide teachers grab-and-go materials that can be integrated seamlessly into the classroom.

**Outputs are perceived as trustworthy.** As before, evidence in both the survey responses and focus groups demonstrated that some of the WWC's evidence standards

resonate with what teachers see as indicators of research quality. For instance, in the survey, a small majority (57%) of respondents agreed that controlled experiments were the best way to determine if an educational intervention works. They were also split in their preference for quantitative research, with 48% agreeing that researchers should collect quantitative data when testing the efficacy of an intervention. Not only do these responses indicate an alignment between some teachers' standards of evidence and the WWC's own, but it also indicated that the WWC is already perceived as trustworthy—at least according to the teachers in the focus groups. When asked why they visit the WWC, these teachers appeared to trust the WWC's ability to distinguish high- and low-quality research:

> "I visit the WWC to look at what educational products are found to be effective or not. I often look at WWC when I hear of something someone is using that they are saying they believe is working for their students."
>
> "To find programs that are research based for possible adoption/implementation for my district. To see if the programs being used have validity."
>
> "I visit the WWC website for educational supports that I know are quality and scientifically researched."

Admittedly, these responses were unsurprising given that focus group participants were already attracted to the WWC, but they still suggest that the WWC's vetting process is perceived as rigorous and trustworthy.

At the same time, however, focus group participants voiced concern about how the WWC's existing standards may actually hinder its credibility. Several of these concerns echo those noted in the initial analysis. For example, some educators expressed concern over the relatively small number of studies that earn the WWC's stamp of

approval. Indeed, the shortage of available studies limits the WWC's relevance to teachers and administrators:

> "I would like to see more programs covered. I frequently find that when I'm looking for a specific program it is not reviewed. I have also found programs/practices that I would be interested in that I couldn't find more information about outside of the clearing house. I spent quite a while searching for more information and finally gave up."

> "It is often disappointing to find that there is no acceptable research on specific products/programs and; therefore, no useful information."

> "There are so many educational products out there that aren't found on the WWC site."

This shortage of studies also affected how teachers perceived the WWC's credibility. Participants in nearly every focus group were concerned that practice guides recommendations based on "minimal evidence" would be a turn off for practitioners looking for practices to implement in the classroom:

> "To me, some of it depends, again, on who the audience is. If you're a researcher, then you might be more interested in some of these things where their evidence rating is minimal. My question is, if you're an educator and you're looking for What Works, then the first thing it says is "recommendation — do this, and here are all the steps," and at the end it says, "Minimal evidence," why would you care? Like to me, I would only show them the first one. If! was talking to educators, I would show them recommendation one and then have a little bit more information on what moderate means."

> "If there is minimal level of evidence, why are we advising teachers to perform these functions?"

> "I'd like just impose that thought that with having read this and playing the role of not multi managed teacher, but just a regular teacher into a newish teacher. And the things that were laid out really made sense to me and I can reinforce that. Then I get to the section that says the effects of these are minimal. Minimal, minimal, minimal."

> "I don't use these [the practice guides] anymore because I used to and I got that. 'Why are you giving this to us a minimal evidence? I can't site this. I can't pick this up. It's a waste of my time.'"

Accordingly, the focus group data bolster many of the same concerns raised in Step 4.

These data corroborated another claim from Step 4—namely, that the WWC's preoccupation with internal validity causes it to overlook the importance of external validity to teachers. In their survey responses, less than half of all preservice teachers (44%) agreed or strongly agreed that, when judging the quality of research, they care more about its methods or design than the context in which it was conducted. The importance of context when deciding whether to trust a research finding was also communicated in the focus groups, such as in the following:

> "I'm a teacher and I — from my experience with my current administrators or — because I've gone through a couple now. One retired and then the principal moved up into the superintendent's positions. But I think they're constantly looking at what worked in different schools and how can they implement what worked in that school within our school, especially if the school district has the same demographics as our district." (DM620529 REL West)

In this case, however, the participant is referencing how their administrators attend to findings in context, not that they do themselves. Even so, the role of external validity in teachers' standards of evidence again featured in this revision of the contribution story.

Finally, as noted initially, teachers have a propensity to trust information and/or recommendations shared by fellow educators. Focus group participants acknowledged this as well, and they spoke to how receptive teachers are when a colleague claims that something "works":

> "…there's two categories. What Works is kind of a condensed research on what works and there's a lot of research that we push on our students and that they're exposed to. There's another category of teachers putting on the web things that have worked for them. Tricks — they have all different kinds and they can be all subject matter [inaudible 00:03:28] so on and so forth. And that second category are not the first category of what works, but they are things that interest teachers and they — I've found many of our — my students anyway — find themselves spending more time with the anecdotal, 'Here's what I tried in my classroom and here's what I found.'"

When asked what could make the WWC practice guides more useful, another participant suggested:

> "…online communities where teachers and others can discuss experiences putting these practices into action."

Thus, though not overwhelming, there was additional evidence in the focus groups suggesting that teachers trust the experiences of each other as much—if not more—than the WWC's own resources.

***Behavior Change Assumptions***

In cases where teachers become motivated to alter their practice in accordance with WWC recommendations, certain conditions must still be met for behavior change to occur (Mayne, 2015, 2017). The presence of those conditions is assumed as part of the WWC's behavior change assumption, which states that instrumental research use is possible. In this context, instrumental use refers to occasions where teachers use WWC-provided information to guide their behavior, as opposed to using it to influence how they think about teaching practices or various educational issues (i.e., conceptual use).

The additional evidence gathered in Step 5 finds that, in certain cases, this assumption holds. Put simply, some teachers who are motivated to change their practice in accordance with WWC resources do, in fact, do so. There were also indications that teachers—at least those early in their career—believe that barriers to instrumental research use can be circumvented. For example, a large majority of survey respondents agreed (64%) or strongly agreed (22%) that any teacher who wished to can implement evidence-based practices in their classrooms. There was less agreement (48%) about whether or not it was reasonable to expect teachers to implement these practices without

altering them first. One survey respondent even claimed that they had already "used [the] WWC to create 3 different intervention programs in the classroom." Instances of instrumental use were mentioned in the focus group discussions as well. Consider teachers' responses to a question about whether they use the WWC website to inform their decision-making:

> "In the past, I have found research studies that I have found incredibly interesting and helpful through the WWC Website. I used the information found within the studies to change my classroom practice or try something new with my students. I have not purchased a curriculum or made administrative level decisions based on what I found through the WWC Website."
>
> "I have used the website to help inform decision making concerning classroom practices and purchase of reading series for our classrooms as well as to verify that what we have has some merit (or not)."

Again, though these teachers were sampled purposively, they still suggest that instrumental use may, indeed, be preferable and possible for some educators.

That said, the additional evidence also suggested that teachers' use of the WWC occurred in other forms. In fact, evidence in the survey and focus groups converged to suggest that the conceptual and symbolic use of WWC resources was just—if not more—common than instrumental use. In terms of conceptual use, both survey respondents and focus group participants alluded to the WWC's usefulness as a launchpad from which to locate new and interesting educational research. As explained by one respondent, the WWC is not their primary source of educational research, but they have "used it as a starting point in reviewing literature on specific practices." Another responded in the following way:

> "Yes and no. I may start with WWC; however, I also look into the literature…I then use WWC to help guide me in deciding whether or not the studies I refer to are of high quality and rigor."

This latter response again alludes to teachers' perception of the WWC as a research access point that, because of its rigorous standards, is trustworthy. Teachers in the focus groups also spoke to the WWC's value as a starting point when learning more about a specific program or research area:

> "I often visit [the WWC website] and find topics that intrigue me and as a result. I print reports and they become my nightly reading."

> "It's almost like it [a WWC video] can serve as a port of entry in to that as it's sustained from a teacher that might have rustled with this and wants to go beyond that somehow."

> "I often visit to check on products I am interested in as well as those my school may have expressed interest in. I also visit to seek out products/program that address specific intructional [sic] needs and in resonse [sic] to emails sent from WWC about research topics of interested to me."

Hence, teachers conceptual use of the WWC is apparent.

Even so, another commonly cited use of the WWC, at least in focus group discussions, was less redeeming. Teachers, as well as researchers and administrators, discussed using the WWC in order to strengthen (and, in some cases, "game") grant applications. Examples of this were abundant when focus group participants were asked why they visit the WWC website:

> "I mostly look when I'm looking for research results that can be used for preference priority points on U.S. Dept. of Ed grant applications."

> "To gain information about data for grants and search for qualified research data."

> "I have used it [the WWC's website] as a resource in writing grants for items that I need to find a sound rationale for."

> "I visit the site when I am looking for vetted research or programs considered to be effective by the US Department of Education or other funders who want to fund programs showing evidence of promise."

Using the WWC in this way isn't necessarily condemnable, especially if it eventually leads to the implementation of evidence-based programs. It does imply, however, that the WWC's work is being used in unintended ways—namely, to curry favor when seeking grant funding.

Finally, it is important to note that these additional data sources verified certain barriers to instrumental use identified in Step 4. Specifically, teachers' lack of agency to implement evidence-based practices—even when they wanted to—was mentioned in the survey and focus groups. When asked why they had heard of the WWC but not used it, one survey respondent explicitly mentioned this lack of agency, writing that "I am being held by my district to use the curriculum that they have provided…" This sentiment was reiterated in the focus group discussions, though somewhat more implicitly:

> "I am not sure I will share this video. At this point I am not involved in any curriculum/instructional product search. We just adopted a new math series and I am not sure what subject is our next purchase, nor who is conducting the search at this time."

> "I just think that if you target adminstrators [sic], and employees at the district offices the information would spread more quickly. In our district the folks at the district office are the ones that could really benefit from the information you provide because they are the ones that often make the decisions about what programs or curiculum [sic] that will be used throughout the district."

Because teachers are seldom in a position to make programmatic decisions, instrumental use of the WWC is far less feasible for them than it is for district- or school-level administrators. This may be one reason why previous work finds the WWC to be more popular among administrators (e.g., Penuel et al., 2017). Given this additional evidence, support for the behavior change assumption that instrument use is possible, particularly among teachers, is mixed. Though some report using WWC resources to retool their

classroom practices, others are not able to innovate because they lack decision-making power.

### *Intended Size of Contribution*

For any behavior change resulting from WWC engagement to constitute a contribution, it must occur to a "meaningful" extent. Originally, the IES established an "ambitious" (Baldwin et al., 2008, p. xvii) benchmark from which to gauge "meaningful" engagement—namely, that 25% of education decisionmakers would consult the WWC prior to making a decision. That goal was to be achieved by 2014; by 2020, it is reasonable to believe that the benchmark would be even higher (i.e., 37.5%). Regardless of the benchmark used, additional evidence from the preservice teacher survey finds usage rates to be meager. 16 out of 105 respondents indicated that they used the WWC—a usage rate of approximately 15%. This rate is comparable to the rate reported by convening attendees responding to the EdTech Evidence Exchange survey, who reported their usage to be 11% (Barton & Tindle, 2019). As was the case in the GAO survey (2010), respondents also reported using the WWC reported infrequently—namely, 87% of respondents said that they had used the WWC once per month or less; the remaining two teachers (i.e., 13%) reporting using the WWC 2-3 times a month. Thus, the evidence suggests that WWC can still not consider itself to be a "central" source of scientific evidence on educational programs and practices, especially among new teachers.

## Revising the Contribution Claim

Taken together, the additional evidence changes the existing contribution claim very little (see Table 4.4). If anything, this evidence strengthens all parts of the original three-part claim. First, it supports the claim that the WWC is likely failing to make a

meaningful contribution to the IES's goal of "…increas[ing] use of data and research in education decision-making" (IES, n.d.-a). More specifically, it is not being used as widely or as frequently as the adjusted benchmark intends. Second, as was also stated in the initial contribution claim, this wanting contribution may be the result of faulty causal assumptions about the WWC's reach and ability to facilitate capacity changes among educators. It may also be the result of faulty rationale assumptions about the perceived usefulness of systematic review efforts in closing the research-practice gap. The third part of the initial contribution claim suggested that the WWC's contribution is stronger (though still not sufficient) among administrators than it is among teachers. Based on this additional evidence, I can add that its contribution appears stronger among administrators than it does among a specific subset of teachers—namely, those who are still new to their profession. This is especially troubling given that these teachers represent the next generation of education decisionmakers, but it is still reconcilable. In the next chapter, I will outline recommendations that, if followed, might help the WWC better maximize its contribution as a part of the IES's research dissemination repertoire.

## Chapter 4 Tables and Figures

Figure 4.1 - WWC Theory of Change as Developed for this Contribution Analysis

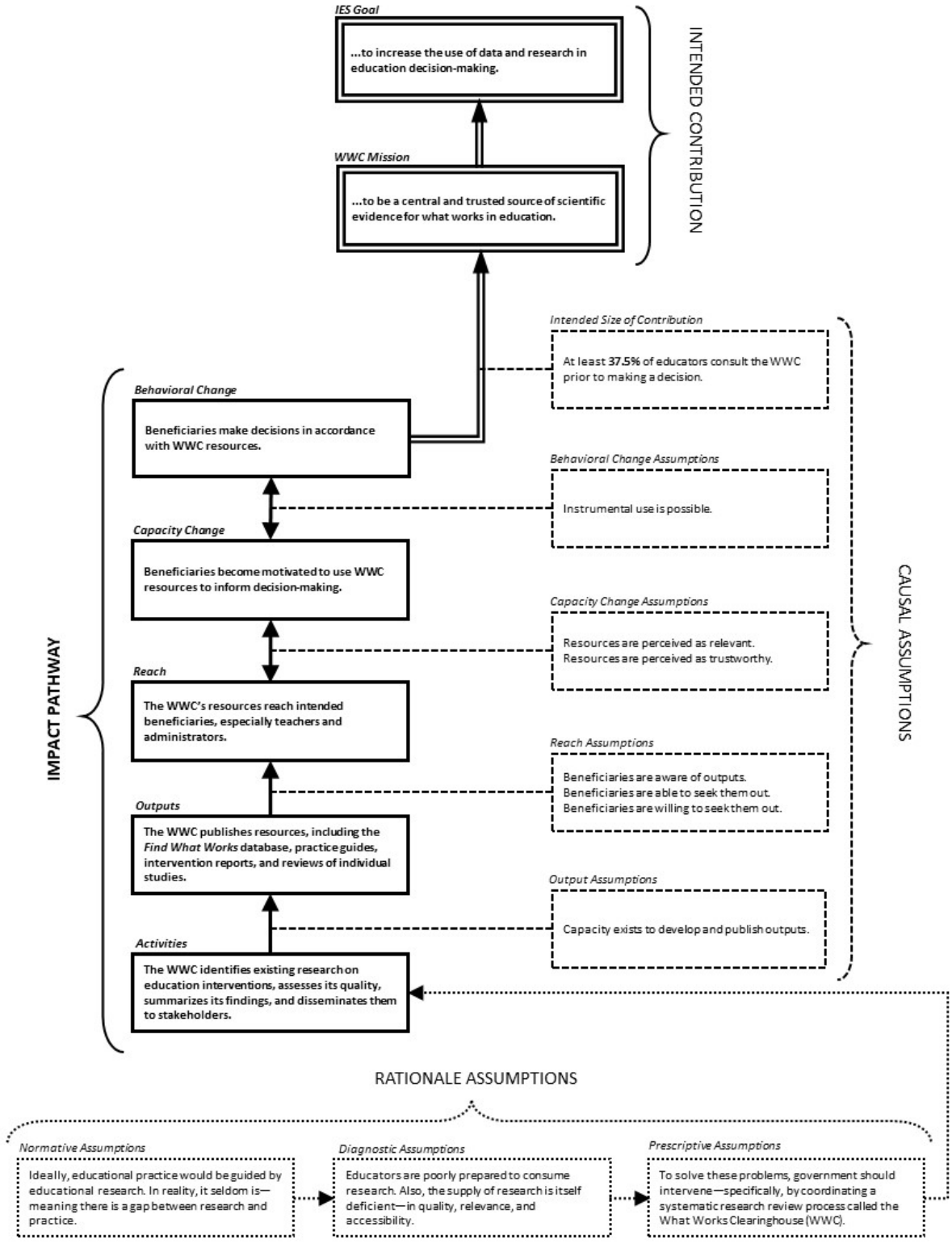
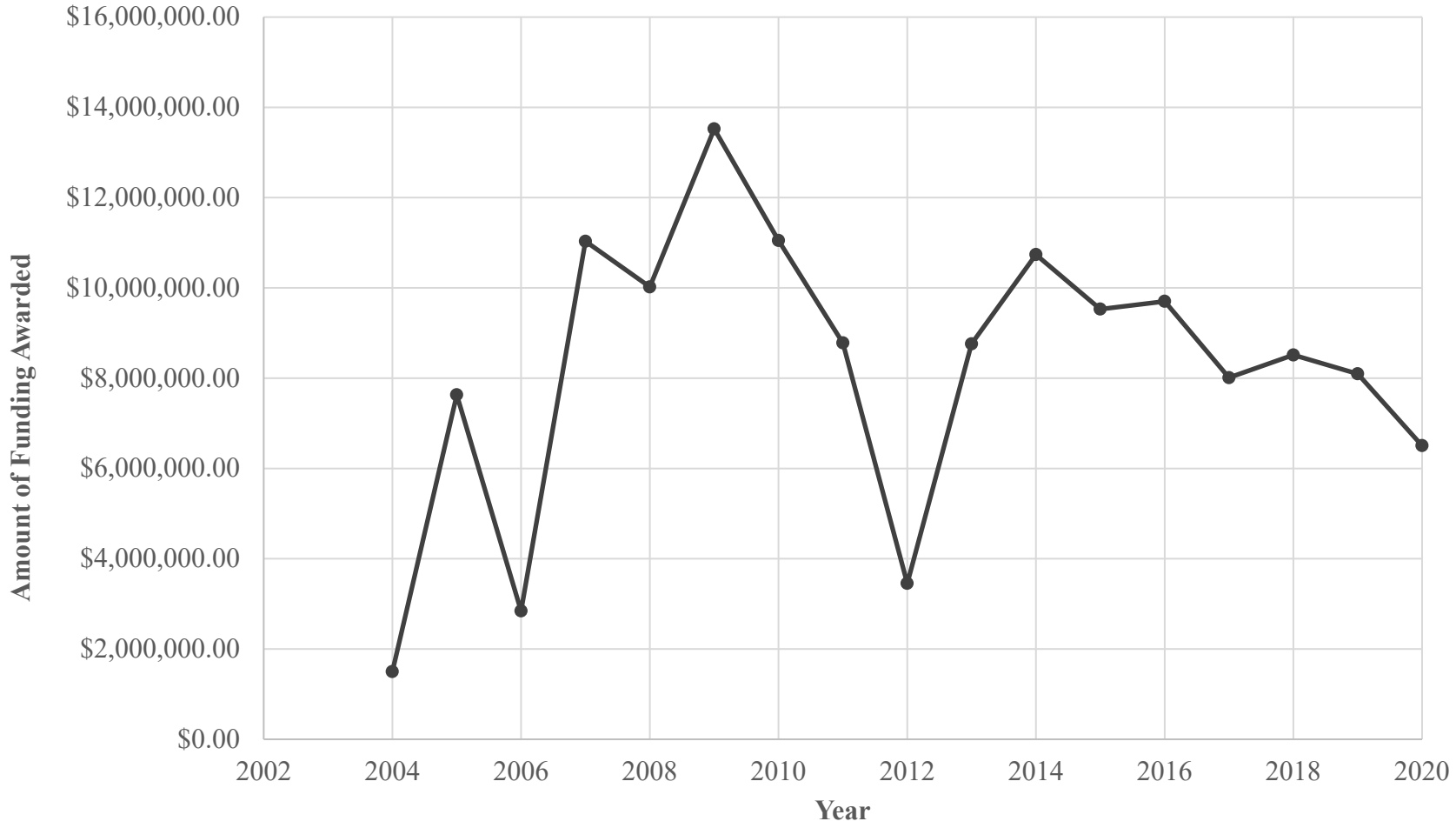


Figure 4.2 - Size of WWC-Related Contracts Afforded Over Time




*Note.* Data extracted from usaspending.gov did not include any WWC-related contracts awarded prior to 2004.

Table 4.1 - Grover Whitehurst's Statements on the What Works Clearinghouse, 2002-2004

| Source | Date | Text |
|---|---|---|
| Hearing before the House Subcommittee on Education Reform of the Committee on Education and the Workforce | February 2002 | "Again, that is exactly our intent with the What Works Clearinghouse effort. Currently, at least for some topics, and school size, the one you mentioned, is one of those topics with a fairly large amount of literature. If you go to our current dissemination effort, which is the Education Research and Information Clearinghouse, ERIC, and click on class size, the problem is that you will generate hundreds of hits. Some of those lead to articles or scientific papers or summaries that accurately and informatively describe the research in that area. Many others do not. The descriptions, when they are of high quality, in many cases, are not framed in a way that would be particularly useful to the school superintendent, and it is just very difficult. If you look at the local family physician or pediatrician, for example, we really don't expect those professionals to go do a thorough review of hundreds of papers themselves to decide which drug to dispense in the office. The Federal Government has a role in vetting that information and providing it to them. We think it is very important that the U.S. Department of Education provide that information to practitioners and school superintendents and educators in the form that you are mentioning, and so that it is user-friendly, pre-adjusted, understandable, and useful in decision-making." (Hearing before the House Committee, 2002c) |
| Hearing before the Senate Committee on Health, Education, Labor, and Pensions | June 2002 | "We also have an effort under way called the What Works Clearinghouse, which will be for the first time a place that people can turn for evidence with respect to educational products and programs and approaches that will not provide a list of accepted programs but simply will provide information on how much research exists, what its quality is and what its direction is. We are not without research in many important areas and we hope that the What Works Clearinghouse will be the major portal to the sort of evidence that is out there that can be usable by parents and educators." (Hearing before the Senate Committee, 2002b) |

Table 4.1 (continued) - Grover Whitehurst's Statements on the What Works Clearinghouse, 2002-2004

| | | |
|---|---|---|
| Interview with T.H.E Journal | January 2004 | "The work of the Clearinghouse is to provide an instrument that can be used by people, such as readers of T.H.E. Journal, which will provide them with such information, as is available, that's relevant to the decisions they have to make when they purchase technology, of a curriculum of a professional development model. It's to provide a well-respected source of information with regards to what the science says, and what evaluation says, about which programs work for whom." (Mageau, 2004) |
| IPR Distinguished Public Policy Lecture Series | April 2004 | The knowledge generated by research must be disseminated in a clear, user-friendly, and easily accessible format. To this end, IES created the What Works Clearinghouse. Its sole purpose is to deliver solid research into the hands of educators, policymakers, and the public. (Whitehurst, 2004) |

Table 4.2 - WWC Conceptualization of Its Mission and Goals Over Time, 2008-2019

| Source | Date | Text |
|---|---|---|
| WWC Handbook V2 | 2008 | "The mission of the Institute of Education Sciences' (IES) "What Works Clearinghouse" is to be a central and trusted source of scientific evidence for what works in education. By reviewing and synthesizing scientific evidence, the What Works Clearinghouse (WWC) is fulfilling part of IES's overall mission to bring "rigorous and relevant research, evaluation and statistics to our nation's education system." The IES is within the U.S. Department of Education and the WWC is within the institute's National Center for Education Evaluation and Regional Assistance." (WWC, 2008b) |
| WWC Handbook V2.1 | 2011 | "The What Works Clearinghouse (WWC) is part of the U.S. Department of Education's Institute of Education Sciences (IES), which was established under the Education Sciences Reform Act of 2002. With its critical assessments of scientific evidence on the effectiveness of education programs, policies, and practices (referred to as "interventions"), and a range of products summarizing this evidence, the WWC is an important part of IES's strategy to use rigorous and relevant research, evaluation and statistics to improve our nation's education system. The mission of the WWC is to be a central and trusted source of scientific evidence for what works in education." (WWC, 2011) |
| WWC Handbook V3 | 2014 | "The What Works Clearinghouse (WWC) is an initiative of the U.S. Department of Education's National Center for Education Evaluation and Regional Assistance (NCEE), within the Institute of Education Sciences (IES), which was established under the Education Sciences Reform Act of 2002. The WWC is an important part of IES's strategy to use rigorous and relevant research, evaluation, and statistics to improve our nation's education system. It provides critical assessments of scientific evidence on the effectiveness of education programs, policies, and practices (referred to as "interventions") and a range of products summarizing this evidence....The mission of the WWC is to be a central and trusted source of scientific evidence for what works in education." (WWC, 2014) |

Table 4.2 (continued) - WWC Conceptualization of Its Mission and Goals Over Time, 2008-2019

| | | |
|---|---|---|
| WWC Handbook V4 | 2017 | "The WWC is an initiative of the U.S. Department of Education's Institute of Education Sciences (IES), which was established under the Education Sciences Reform Act of 2002. It is an important part of IES's strategy to use rigorous and relevant research, evaluation, and statistics to improve our nation's education system. The mission of the WWC is to be a central and trusted source of scientific evidence for what works in education. The WWC examines research about interventions that focus on improving educationally relevant outcomes, including those for students and educators." (WWC, 2017c) |
| WWC What We Do Video | 2018 | "Not all education research is equal. Identifying well-designed studies, trustworthy research, and meaningful findings to inform decisions and improve student outcomes can be tricky. That's where What Works Clearinghouse, or the WWC, comes in. An investment of the Institute of Education Sciences in the U.S. Department of Education, the WWC is a resource that helps teachers, researchers, administrators, and policymakers make evidence-based decisions. We review the research, determine which studies meet rigorous standards, summarize the findings, and provide tools to help educators use research in practice." (IES, 2018c) |
| WWC Handbook V4.1 | 2020 | "The WWC is an initiative of the U.S. Department of Education's Institute of Education Sciences (IES), which was established under the Education Sciences Reform Act of 2002. It is an important part of IES's strategy to use rigorous and relevant research, evaluation, and statistics to improve our nation's education system. The mission of the WWC is to be a central and trusted source of scientific evidence for what works in education. The WWC examines research about interventions that focus on improving educationally relevant outcomes, including those for students and educators." (WWC, 2020d) |

Table 4.3 – IES Commentary on the What Works Clearinghouse, 2018-2020

| Source | Date | Text |
|---|---|---|
| Blog Post on "Changes are Coming to the WWC" | 2018 | "I would like to take this opportunity to tell you about some of the directions IES will be exploring to improve the What Works Clearinghouse (WWC), a flagship product of IES since its inception. The WWC recently passed its 15th birthday and has gone through many changes since then to ensure that it continues to help the nation identify what works for whom and under what conditions." (Schneider, 2018) |
| Blog Post on "First Year Accomplishments" | 2019 | "But more important is how we are rethinking the What Works Clearinghouse website and its Practice Guides. Both are among IES's marquee activities, and WWC staff have done excellent work in managing the growing volume of reports and studies that are now eligible for WWC review…[t]hat said, we are striving to make the WWC website as usable as possible for the wide range of users who come to it. All too often, our own rules and procedures have made the WWC and Practice Guides prone to burying the valuable information they contain in language that is difficult to understand." (Schneider, 2019) |
| IES Director's Biannual Report to Congress (2017-2018) | 2019 | "The Institute's What Works Clearinghouse (WWC) has primary responsibility for synthesizing existing evidence about educational practices, programs, and policies and disseminating their findings to federal, state, and local policymakers. The WWC released 30 Intervention Reports and Practice Guides in FY 17–18. Practice Guides, produced through the WWC, are among the most important products of the Institute, because they are central to translating research into practice." (IES, 2018a) |
| Blog Post on "A New Year's Update" | January 2020 | "We are investing more resources in practice guides, which are some of the most downloaded documents on our website…[t]hey are essential to our mission since they help to translate research for arguably our most important stakeholders --- educators." (Schneider, 2020) |

Table 4.3 (continued) - IES Commentary on the What Works Clearinghouse, 2018-2020

| | | |
|---|---|---|
| IES "About Us" Website | April 2020 | "The Institute of Education Sciences (IES) is the statistics, research, and evaluation arm of the U.S. Department of Education. We are independent and non-partisan. Our mission is to provide scientific evidence on which to ground education practice and policy and to share this information in formats that are useful and accessible to educators, parents, policymakers, researchers, and the public…[w]e provide resources to increase use of data and research in education decision making. Through the What Works Clearinghouse, we conduct independent reviews of research on what works in education." (IES, n.d.-a) |

Table 4.4 - Summary Table of Evidence and Conclusions from Contribution Analysis

| COMPONENT | WWC | Eval | Lit | Sur | FocG | TeSur | CONCLUSIONS |
|---|---|---|---|---|---|---|---|
| **Normative Assumptions** (beliefs about the existence of a problem—namely, that what is occurring in reality is not ideal) | | | | | | | |
| In reality, educational research rarely guides educational practice. | | | + / - | + / - | + | | Beliefs about a research-practice gap are justifiable, but more so for teachers than administrators. The belief that research "should" guide practice remains an open question. There is little proof that educators' research use ultimately benefits students. |
| Ideally, research should guide practice. | | | - | | + | + | |
| **Diagnostic Assumptions** (beliefs about the causes of a problem) | | | | | | | |
| Educational research is deficient—in quality, relevance, and accessibility. | | - | + / - | +/- | | - | There is also mixed support for the WWC's diagnostic assumptions. Though the qualities of educators and educational research factor into the research-practice gap, a more significant factor is the structural barriers inhibiting educators' research engagement. |
| Educators struggle to consume research. | | | +/- | - | + | - | |
| **Prescriptive Assumptions** (beliefs about how best to solve the problem) | | | | | | | |
| The solution must be government-controlled. | | - | | - | | | Neither systematic review nor federal involvement appear to be vital to closing the research-practice gap. Federal dissemination efforts are underutilized, and systematic review initiatives generally share findings in a passive, uncollaborative manner. |
| The solution should involve a systematic research review. | | - | +/- | - | - | + | |
| **Activities** (the actions taken by a program to bring about its intended results) | | | | | | | |
| Identification, assessment, and summarization of research occurs. | +/- | +/- | | | | | The WWC continues to systematically review research, but its review rate has slowed in recent years. |
| **Output Assumptions** (the conditions necessary for a program's activities to be translated into outputs) | | | | | | | |
| Capacity exists to develop/publish outputs. | +/- | +/- | | | | | The WWC may lack the capacity to translate its activities into outputs, but more evidence is needed. |
| **Outputs** (the tangible goods or services produced by a program to bring about its intended results) | | | | | | | |
| Resources are published. | +/- | +/- | | | | | The WWC's rate of publishing outputs has slowed, but its rate of updating existing outputs has improved. |

*Note.* WWC = WWC or IES Data; Eval = Previous Evaluations; Lit = Literature Reviews, Sur = Large-Scale Surveys; ForG = Focus Groups; TeSur = Teacher Survey

Teacher / Administrator — Confirmatory Evidence (+) — Mixed Evidence (+/-) — Refuting Evidence (-) — No Available Evidence

Table 4.4 - Summary Table of Evidence and Conclusions from Contribution Analysis (continued)

| COMPONENT | WWC | Eval | Lit | Sur | FocG | TeSur | CONCLUSIONS |
|---|---|---|---|---|---|---|---|
| **Reach Assumptions** (the conditions necessary for a program's outputs to reach its intended beneficiaries) | | | | | | | |
| Beneficiaries are aware of outputs. | | – / +/- | | | – | – | Overall, the evidence refutes these assumptions, especially for teachers. Awareness of the WWC depends on the type of educator examined, but awareness is wanting. For teachers who are aware of the WWC outputs and willing to engage with them, a host of barriers (e.g., a lack of time) may still keep them from seeking out these outputs, even if they want to. |
| Beneficiaries are able to seek out outputs. | | – | – | | – | – | |
| Beneficiaries are willing to seek out outputs. | | – | – / +/- | – / +/- | – | +/- | |
| **Reach** (the activities/outputs characterizing a program reach their intended beneficiaries) | | | | | | | |
| WWC outputs reach intended beneficiaries. | – | – | | + | | | Although some administrators may to be reached by WWC outputs, few teachers are. |
| **Capacity Change Assumptions** (the conditions necessary for a program's outputs, once reached, to elicit desirable cognitive changes) | | | | | | | |
| Outputs are perceived to be relevant. | – | – | – | – | – | +/- | Some educators perceive the WWC's outputs as credible. Others, however, might not. WWC outputs likely lack relevance given their untimeliness, topic areas, and lack of applicability to practice. |
| Outputs are perceived to be trustworthy. | – | – | +/- | | +/- | +/- | |
| **Capacity Change** (the cognitive changes that result from beneficiaries' engagement with activities/outputs) | | | | | | | |
| Beneficiaries becomes motivated to use WWC resources. | | | | | | | Unfortunately, little is known about how educators' capacities change after engaging with WWC outputs. |
| **Behavior Change Assumptions** (the conditions necessary for beneficiaries' cognitive changes to manifest in behavior changes) | | | | | | | |
| Instrumental use is possible. | | | – | | +/- | +/- | Teachers' lack of agency makes it hard for some to implement research in their classrooms. |
| **Behavior Change / Intended Contribution** (the behavioral changes that result from beneficiaries' engagement with activities/outputs) | | | | | | | |
| Users make decisions in accordance with WWC resources; 37.5% of educators report doing so. | | – | | | | – | Administrators use WWC outputs more than teachers; even so, both groups have usage rates below 37.5%. |

*Note.* WWC = WWC or IES Data; Eval = Previous Evaluations; Lit = Literature Reviews, Sur = Large-Scale Surveys; ForG = Focus Groups; TeSur = Teacher Survey

Teacher / Administrator — Confirmatory Evidence: + — Mixed Evidence: +/- — Refuting Evidence: - — No Available Evidence

# CHAPTER 5 – DISCUSSION

This project used a theory-based evaluative approach called contribution analysis (Mayne, 2008, 2012b, 2019) to articulate and scrutinize the What Works Clearinghouse's (WWC) theory of change. Its ultimate goal was to pursue answers to three guiding questions, each of which was motivated by tenants of evaluative thinking (Vo & Archibald, 2018). The first question inquired into the extent of the WWC's impact—namely, the extent to which it has fulfilled its mission of becoming a "central and trusted source of scientific evidence for what works in education" (WWC, 2020d, p. 1) while, in turn, contributing to the IES's larger goal of "…increas[ing] use of data and research in education decision-making" (IES, n.d.-a). The second question inquired into the reasons behind the WWC's wanting impact, with a special focus on the assumptions underlying its approach to identifying, assessing, summarizing, and disseminating educational research. The third question inquired into how the WWC might be improved.

In this chapter, I propose answers to this third question while reviewing my answers to the first two. Together the evidence suggests that the WWC's contribution is not as robust as was originally hoped, especially among classroom teachers. This may be because the WWC's approach appeals to certain rationale and causal assumptions that are shortsighted given the empirical evidence. Moreover, these assumptions are often incongruent with those held by educators. Thus, to better align its approach to practitioners' perspectives, needs, and realities, the WWC may consider the recommendations offered hereafter.

### Question 1: Is the WWC Working?

To ask if the WWC is working is, in essence, subjecting the WWC to the same question that it poses to educational programs and practices. As explained in Chapter 1, it is also an example of evaluative thinking (ET) in action (e.g., Vo & Archibald, 2018). Indeed, two of ET's central tenants have helped motivate this first guiding question, including (a) its appreciation for accountability as a functional part of healthy democracy, and (b) its belief that an evaluator's positionality must be both recognized and examined (Vo et al., 2018). With regard to the former, asking if the WWC is "working" is an example of metaevaluation—of holding evaluators (or evaluative bodies, like the WWC; Scriven, 1994) accountable to the same standards that they impose onto others. With regard to the latter, asking if the WWC is "working" is a challenge to my own biases and beliefs, which have questioned the WWC's ineffectiveness without full consideration of the evidence.

Upon consideration of the evidence, though, I am left with a conclusion that closely aligns with my initial beliefs. The WWC has failed to meaningfully contribute to the Institute of Education Science's larger goal of "increas[ing] use of data and research in education decision-making" (IES, n.d.-a), and in addition, it has failed to fulfill its own mission of becoming a "central and trusted source of scientific evidence for what works in education" (WWC, 2020d, p.1). Evidence of these failures is apparent after examining the WWC's impact pathway. More specifically, although there is evidence that the WWC's activities and outputs are occurring as planned, its reach and behavior change components are not. In the remainder of this section, takeaways from each step of the

impact pathway are summarized, along with limitations and future directions when applicable.

**The WWC continues to execute its process of systematic research review, but its current review rate is slower than it was a decade ago.** One indicator of the WWC's activities was the rate of its individual study reviews, as those reviews serve as building blocks for many of its outputs. Data cataloguing the pace at which WWC contractors have reviewed individual studies suggested that this pace has slowed in recent years—from an average review rate of 1000 studies per year to a current rate of just over 100. The rate at which WWC review protocols are developed was analyzed as well. Although the data showed a dip in the number of new protocols developed each year, it also revealed an increase in the number of older protocols being updated. Ultimately, these patterns may be more indicative of a shift in focus than a slacking in protocol development.

**Although the WWC's outputs are currently being published at a slower rate than they were a decade ago, its rate of updating existing outputs has quickened.** A noticeable peak in the publishing of intervention reports and practice guides occurred in the late 2000s (i.e., 2005-2009), during which virtually half of all available intervention reports and practice guides were produced. In more recent years, intervention report and practice guide development has slowed. This suggests that many of these outputs are outdated, though this is more so the case for intervention reports than practice guides. In fact, the WWC's rate of revising its practice guides has improved, such that six of its seven revisions/updates have occurred since 2017. This could again signal a shift in priorities.

An interesting observation linked to this pattern of findings—and one that might warrant further investigation—is that the WWC's peak output publication rate mirrors a peak in funding levels. Put another way, as funding levels essentially doubled between the WWC's first (from 2002-2007; $28.5 million) and second five-year contract cycles (from 2007-2012; $53.3 million), its practice guide publication rate quadrupled. This may be for reasons other than funding (e.g., focusing on other, more traditional research synthesis outputs—such as intervention reports—in its early years), but it is provocative nonetheless. Likewise, in 2007 alone, the WWC produced more intervention reports than it has in all other program years combined (i.e., 303 vs. 290), which is harder to attribute to funding increases given its occurrence during a transition year between the two contract cycles. Future examinations of the WWC should explore the factors responsible for this blip in output publishing, and moreover, if increasing current funding could lead to swifter output production.

**Only a minority of educators—especially teachers—have been reached by the WWC's outputs.** Early estimates from the GAO (2010) suggested that administrators are more likely to have been reached by a WWC output than teachers, but both groups had not been reached with much success. Whereas 34% of district-level respondents had accessed the WWC in 2010, only 15% of principals and 5% of teachers had reported doing so. More recent rates reported by Barton and Tindle (2019) are nearly identical, with 14% of the sample saying they accessed research through the WWC. Although there are some indications that access rates among administrators have increased (e.g., Penuel et al., 2017), all of these estimates are lower than my adjusted benchmark of how many educators would *use* information from WWC outputs to inform

their decision-making (i.e., 37.5%). Thus, the WWC's lack of reach is a clear threat to its impact pathway, especially for teachers.

If anything, the WWC seems to be reaching the wrong beneficiary groups. Specifically, although the WWC has inconsistently monitored the types of people that do end up visiting its website, existing findings suggest that the website is more popular among researchers than it is among teachers/administrators. In 2008, the most common visitors (at least according to pop-up survey responses) were teachers and administrators; in 2014, however, site visitors who completed a pop-up survey most commonly identified themselves as "researchers or college/university staff" (WWC, n.d.-f). It is unclear what the current breakdown looks like, but it seems plausible—based on these surveys as well as responses in the Mathematica focus groups—that the WWC is doing a better job at reaching researchers than educators.

**Little is known about how educators' capacities change when engaging with WWC outputs.** This is perhaps the biggest limitation of my evaluation. As noted in both Step 4 and Step 6 of the contribution analysis, the "Capacity Change" component of the WWC's impact pathway could not be refuted or confirmed. Because there is sufficient evidence refuting the WWC's capacity change assumptions, it is tempting to conclude that few educators who engage with WWC outputs end up feeling motivated enough to try and put them to use.

Assuming this would be overstepping, though, especially given the responses of some focus groups participants. In those groups, participants were shown prototypes of WWC outputs and probed for feedback. One of those outputs included a beta version of the *Find What Works* tool—a tool that has since become a central part of the WWC

website. Results of the focus group, which were summarized in a 2016 memo, concluded that "respondents believed the tool will be useful for their work and are likely to share it with others, especially with their colleagues" (Mathematica, 2016, p. 1). This finding is indicative of a capacity change—that is, when users engaged with the *Find What Works* tool, they became *motivated* to share it with colleagues. Even so, this finding is limited for several reasons. First, it comes from a self-selected sample of stakeholders who volunteered themselves to be part of the focus groups. They were already aware of the WWC and likely had a positive view of it. A second limitation is that the *Find What Works* tool is best thought of as a search engine through which educators can reach individual study reviews, intervention reports, and practice guides. Educators' motivations after engaging with this search engine reveal little about their motivations once these more traditional outputs are reached. Future projects or feedback groups would yield more compelling information about users' capacity changes if (a) participants were randomly selected, and if (b) feedback was solicited about the WWC's existing outputs, such as intervention reports.

**WWC outputs appear to be underutilized, though administrators' use may be more common than teachers' use.** Even so, both groups still report usage rates that fail to meet the adjusted benchmark. For example, in the GAO's 2010 survey, 34% of district-level administrators reported using the WWC to guide their curriculum decisions, help design professional development opportunities, and even to help develop school improvement plans. Such a finding is promising, but also tempered by the fact that only 5% of teachers had ever even accessed the WWC. The more recent EdTech Evidence Exchange survey estimated that 32% of association members reported using information

gathered from the WWC, whereas only 11% of convening attendees did so (Barton & Tindle, 2019). Even the survey conducted as part of this evaluation, which assessed beginning teachers' knowledge and use of the WWC, found that only 15% had used it.

If we adjust the IES's original usage goal—namely, that 25% of educators would consult the WWC to inform their decision-making by the 2013-2014 fiscal year—for the present day, then we would assume that 37.5% of teachers and administrators would need to be using the WWC in order for its intended contribution to be realized. Both groups are not, which suggests that it is not working to further the IES's goal of "increas[ing] use of data and research in education decision-making" (IES, n.d.-a). It is worth recognizing, however, that my extension of IES's benchmark is based on arithmetic and not on consultation with WWC representatives. Subsequent evaluations should work with the IES to determine what their current expectations for use are. They should also ask the IES to formalize a benchmark regarding educators' frequency of use, as existing data indicates that use is relatively infrequent. For example, most of the administrators who reported using the WWC in the GAO (2010) survey were estimated as using it to a "small" or "moderate" extent. Similarly, only 2 of the 16 teachers who reported usage in my survey said they did so more than once a month. This suggests that the WWC may entertain a great deal more casual users than diehard ones—a finding that further speaks to its wanting contribution. Finally, additional work should inquire into *how* the WWC is used. Although evidence from both the focus groups and preservice teacher survey suggest that educators use the WWC instrumentally (i.e., to inform decision-making) and conceptually (i.e., to access/read research that they find compelling), it is unclear how much usage is "imposed" (i.e., for compliance reasons; Weiss, 2008). Nevertheless, it is

reasonable to conclude that not all usage of the WWC is intended to directly inform decision-making, so the usage rates reported here are likely inflated when considering that the WWC's intended contribution is one of instrumental use.

### Question 2: Why Isn't the WWC Working?

Arriving at an answer to my first research question—namely, that the What Works Clearinghouse is likely failing to realize its intended contribution—neither satisfies the evaluative principles nor the theoretical motivations of this project. Indeed, as both an exercise in evaluative thinking (e.g., Vo & Archibald, 2018) and a test of Farley-Ripple et al.'s (2018; see also CRUE, n.d.) model of educational research use, this project investigated "why" the WWC's contribution looks the way that it does. In doing so, it focused on the assumptions underlying the WWC's creation in the first place (i.e., its "rationale assumptions"; Nkwake & Morrow, 2016) as well as the assumptions about how its activities as a systematic review initiative would bring about increased research use among educators (i.e., its "causal assumptions"; Mayne, 2015, 2017). Focusing on assumptions is necessitated by this project's allegiance to evaluative thinking. As explained in Chapter 1, Vo et al. (2018) considered the critical examination of biases and assumptions to be a hallmark of evaluative thinking. It is also necessitated by this project's adherence to Farley-Ripple's (2018) theoretical framework, which extends Caplan's (1975) Two-Communities Theory. Their framework proposes that the gap between educational researchers and educational practitioners is a byproduct of gaps in their assumptions about how research should be conceptualized, conducted, and shared. This is yet another reason why this project studied the assumptions undergirding the WWC's theory of change with as much care as its activities and outcomes.

Together the findings suggested that the WWC may not be working because its assumptions are not realistic. Not only were many of the WWC's causal assumptions challenged by the evidence, but some of its rationale assumptions were discovered to be unfounded as well. A recap of this evidence is provided below. In a functional sense, the evidence indicated that the WWC's theory of change was riddled with holes, whereby one component in the theory of change was unable to lead to the next because the assumptions in-between were left unrealized. In a more theoretical sense, the evidence reinforces Farley-Ripple et al.'s thinking about how incompatible assumptions held by the research and practice communities are at least partly responsible for the research-practice gap.

**Rationale Assumptions**

When identifying and evaluating the underlying assumptions of a program's theory of change, Mayne (2015, 2017) recommended attending to "the underlying hypotheses or premise(s) on which the intervention is founded" (2017, p. 157). These hypotheses or premises are referred to as *rationale assumptions*, and other work (e.g., Nkwake, 2013; Nkwake & Morrow, 2016) has similarly advocated for examining these assumptions in an evaluative context. Nkwake and colleagues have worked to typologize rationale assumptions as belonging to three categories, including: normative assumptions, which are value-based assumptions about the existence of a problem and the need to respond to it; diagnostic assumptions, which are assumptions about the core causes or contributors of a problem; and prescriptive assumptions, which are assumptions about how to best address the problem in question. Each of those categories was examined in the context of the WWC, and while doing so, the following observations emerged.

**The educational research-practice gap is not a ubiquitous reality, and the desirability of closing it remains an open question.** Even before the formalization of the IES, the undesirability of a research-practice gap was a chief motivator of U.S. educational policy developments. This "gulf between the bench and the trench" (Whitehurst, 2003, p. 5) has remained in the policy crosshairs because of a certain assumption—namely, the normative assumption that educational practice should be guided by research. Why would such a reality be desirable? Because, as spotlighted by the infamous *A Nation at Risk* report (U.S. National Commission on Excellence in Education, 1983) and echoed repeatedly thereafter (Hargreaves, 1996; Slavin, 2002, 2004), there was (and still is) a prevailing belief that an increasingly research-based U.S. education system could help ameliorate, among other things, the country's loosening grasp on "commerce, industry, science, and technological innovation…throughout the world" (U.S. National Commission on Excellence in Education, 1983, p. 10).

One emergent finding from this project is that the research-practice gap is not a blanket educational phenomenon. Hemsley-Brown and Sharp's (2004) early empirical review, which featured literature available prior to the development of the WWC, concluded that different patterns of research use existed among administrators and teachers. Whereas teachers' research use was found to be rare, "[p]rincipals considered themselves to be regular, thoughtful users of research knowledge" (p. 9). Work by Biddle and Saha (2002, 2006) suggested similarly, as does contemporary work conducted by Penuel and colleagues (2017, 2018), in which "school and district leaders alike reported frequent use of research use" (Penuel, 2017, p. 1). Admittedly, more work is needed when considering the looseness with which extant work has characterized the role of

"administrator." As one example, Penuel's (2017) survey included a diverse mix of principals, program directors, curriculum supervisors, superintendents, and assessment coordinators; educators in each of these roles likely have different data and research needs (e.g., Coburn & Talbert, 2006). Even so, the WWC's assumption that educators seldom engage with educational research is not nearly as defensible for administrators as it is for teachers.

Likewise, although research-informed educational practice may seem like an "unexceptionable" ideal (Buchmann, 1984, p. 421; see also Cook, Smith, & Tankersley, 2012), it—like any other target of evaluative thinking—warrants deeper consideration. In doing so, my contribution analysis found scant evidence that bringing research evidence to bear in practice actually manifests in its intended downstream effects. If anything, the existing evidence suggests that doing so may be neither desirable nor beneficial. For example, Gorard and colleagues (2020) acknowledged that, when evidence-based interventions validated by the WWC are subsequently implemented, they are often found to be "ineffective" (p. 11); Pogrow (2017) has similarly documented how implementation of evidence-based practices can "mislead rather than inform practice" (p. 2); Zhao (2017) even discussed how evidence-based practices can "harm" students and teachers. Together these observations suggest that, if the goal is to elicit lasting benefits in student outcomes, then additional evidence is necessary before we can assume that educational research "should" be brought to bear in practice. We also must ensure that the benefits of doing so outweigh any potential risks (e.g., Zhao, 2017, 2019), and that we think critically about whether or not research-based practice is desirable in the first place (for commentaries, see Biesta, 2007, 2010; Cain, 2016; Hammersley, 2005).

**The most frequently-cited barrier to educators' research use is a structural one—not a barrier that is intrinsic to them.** Statements made during the pre-IES congressional hearings (Hearing before the House Committee, 2002a; Hearing before the House Subcommittee, 2002a), as well as current examples of how the WWC describes itself (WWC, n.d.-g), implicated certain features of educational research *and* educational research users as causes of the research-practice problem. Whitehurst and others blamed educational research for lacking in rigor, relevance, and accessibility (Hearing Before the House Subcommittee, 2002c; Hearing Before the Senate Committee, 2002b). At the same time, they acknowledged how most educators were unprepared to access and/or assess research themselves. The WWC's own Procedures Handbook (WWC, 2020d) recognizes this as well, stating that "…it can be difficult, time consuming, and costly for decisionmakers to access and draw conclusions from relevant studies about the effectiveness of these interventions" (p. 1). These concerns, about both the supply and demand of educational research, characterize the WWC's diagnostic assumptions.

Findings from the contribution analysis offered mixed support for the WWC's assumption that educators find educational research to be irrelevant and lacking in credibility; some do, yet other do not. The evidence was far more convergent when it came to WWC's assumption about educators' own abilities as research consumers. Not only are many administrators and teachers thoughtful consumers of research (e.g., Williams & Coles, 2007; Lysenko et al., 2014; Cain, 2017), but external factors are greater contributors to the research-practice gap than educators' own abilities as research consumers. Indeed, a robust body of evidence suggests that educators' research use is constrained by "organizational structure, culture, and leadership" (Farley-Ripple et al.,

2018, p. 240). In fact, Schaik and colleagues (2018) went as far as to conclude that teachers use of research-based information is "…largely a matter of how organizations operate, which deeply affects the way how individual teachers work" (p. 58). Reviews of the literature (e.g., Dagenais et al., 2012; Cain, 2016) have consistently recognized that organizational structures, especially the lack of time allocated for educators' research engagement, are the predominant barriers to research use. Recent, large-scale surveys have likewise found that organizational factors predict the research use of both teachers (e.g., Lysenko et al., 2014) and administrators (Penuel et al., 2017). The WWC was not designed to address these larger structural issues, meaning that it was built upon questionable assumptions about the core causes of a research-practice gap.

**The most successful research access points are interpersonal and exchange-based rather than passive and unidirectional.** Resting on these assumptions about the existence of a research-practice gap and its causes are two prescriptive assumptions about how best to address them. Those assumptions are that (a) government intervention is needed, and that (b) the government should coordinate a systematic research review process to identify, assess, summarize, and disseminate research on behalf of educators. Both assumptions were neatly laid out by Whitehurst, as well as his contemporaries, in the 2002 congressional hearings (Hearing Before the House Subcommittee, 2002a; Hearing Before the Senate Committee, 2002a). But why, then, have federal research dissemination initiatives generally failed to captivate educators?

One of the most robust findings in the knowledge utilization literature is that, in order to effectively connect research and practice, efforts must cultivate and/or capitalize on connections between people. This has been noted in commentaries (Levin, 2011; Neal

et al., 2018; Tseng, 2012) as well as in reviews of the empirical literature (Dagenais et al., 2012; Gorard et al., 2020; Hemsley-Brown & Sharp, 2004; Schaik et al., 2018). While this might necessitate the "co-creation of knowledge" between members of the research and practice communities (Schaik et al., 2018, p. 59; see examples below), it might also entail capitalizing on educators' propensity to consult one another. Accessing research through a well-informed colleague is not only one of the most common access points cited by educators, but it appears to be one of the most frequently utilized as well (Barton & Tindle, 2019; Penuel et al., 2017). For example, whereas 25% of educators reported consulting "reports or executive summaries" 1-2 times a year, 30% consulted "colleagues who read research" seven or more times a year (Barton & Tindle, 2019).

This discrepancy brings up another point. Literature celebrating the potential of interpersonal relationships as research access points has, at the same time, commented on the futility of efforts that simply summarize research for educators. Systematic research reviews, especially when their results are written up in reports and left for educators to seek out (e.g., passive dissemination; Cooper, 2012) are one such example. Indeed, after their review of the literature, Gorard and colleagues (2020) went as far as to conclude that "[e]ven the best systematic syntheses of evidence often have little impact in practice" (p. 17). These same observations have been made about the WWC's approach in particular (e.g., GAO, 2010), including by the Mathematica focus group participants, who noted that systematic review is a more useful way to engage researchers and upper-level administrators than teachers. Therefore, the assumptions underlying an initiative like the WWC are largely unfounded.

## Causal Assumptions

In Mayne's (2008, 2012b, 2019) descriptions of contribution analysis, special attention is paid to the underlying events or conditions likely necessary for a program to function as planned. These events or conditions are called *causal assumptions*. Causal assumptions underlie each component in a program's theory of change, and more specifically, they underlie the logic of how each component is assumed to lead to the next. For example, for a program's activities to manifest in desired outputs, there must be capacity to develop and publish those outputs; for those outputs to reach their intended audience, they must be known, accessible, and desirable. Not only did this contribution analysis explicate a host of assumptions underlying the WWC's program components, but it also scrutinized those assumptions based on the available evidence. The core takeaways from that analysis are reviewed below, along with their limitations and implications for future research.

**The WWC may lack the capacity to translate its activities into outputs, but more evidence is needed.** The first causal assumption specified in the WWC's theory of change relates to how the findings resulting from its systematic review process are translated into sharable outputs for educators. Plainly stated, does the WWC have enough capacity to do so? Evidence reviewed over the course of the contribution analysis was mixed. Not only did the GAO's (2010) original evaluation find a substantial backlog in the WWC's output production, but my own analysis of current data detected something similar. Indeed, a gap was discovered between the time that an individual study review was conducted and the date that it was eventually posted on the WWC website; the median number of days between these two dates was 54 days.

Though indicative of a backlog similar to that which existed a decade ago, this finding should be treated cautiously. First, it examined a different type of output than the one examined by the GAO. The GAO looked at the backlog in the publication of intervention reports and quick reviews, whereas I examined individual study reviews. A second difference is in the possible explanation of the backlog. The GAO (2010) evaluation implicated the IES peer review process as the reason for the backlog, but the reason for the individual study review backlog is less clear. The WWC Procedures Handbook (2020d) confirms that:

> "…each WWC publication is submitted to IES, which reviews the document internally and sends it for peer review by researchers who are knowledgeable about WWC standards and are not staff with the WWC contractor that prepared the draft publication" (p. C-3).

It is unclear, however, which WWC outputs are considered to be "publications" and which ones are not, as this passage goes on to name "intervention reports" and "practice guides" (WWC, 2020d, p. C-4)—but never individual study reviews. Instead, it appears that the WWC only depends on "certified reviewers" (WWC, 2020d, p. C-3) —not IES staff—to vet these individual study reviews. Accordingly, the amount of time between when a review is conducted and posted may reflect on the promptness of certified reviewers external to the IES.

The WWC does bear some responsibility, however, in ensuring that it has enough certified reviewers to handle the study review process. It also bears responsibility for ensuring that these reviewers are well-trained. As shown in Chapter 4, only a minority of individual reviewers are certified in the latest WWC Group Design Standards, and almost none are certified in the latest standards set for Single Case Designs or Regression Discontinuity Designs. More information is needed to say with confidence that the WWC

is lacking capacity to conduct its reviewers and publish its outputs, especially given that only data from individual reviewers (and not review organizations) could be examined.

**Many educators—especially classroom teachers—are still unaware of the WWC, and even those who are aware of it may be unable or unwilling to access it.** This lack of awareness was true a decade ago, and it remains true today. For instance, the GAO (2010) evaluation discovered that, while awareness was common among state-level administrators (i.e., 87% had heard of the WWC), less than half of all district- and school-level administrators in the sample had heard of the WWC. Even more alarming was the finding that only 13% of teachers reported having heard of it. A more recent estimate provided by the EdTech Evidence Exchange suggests that awareness has likely grown, but that it is still dependent on the type of educator in question. Whereas 62% of association members had heard of the WWC, only 36% of convening attendees had heard of it (Barton & Tindle, 2019). My survey of preservice teachers showed that they may be especially unaware, as only about 20% of respondents had heard of the WWC. Taken together, these findings imply that classroom teachers are less aware of the WWC than administrators are. Furthermore, given that the WWC's adjusted benchmark for *usage* (not simply awareness) specifies a usage rate of 37.5%, these rates of awareness are likely lower than they need to be for the WWC's theory of change to remain intact.

Educators who are aware of the WWC's outputs, however, may still lack the willingness or the ability to access them. Evidence unearthed during the contribution analysis calls into question educators' willingness to engage with the WWC, even when they are receptive to the idea of research-informed practice in the first place. Not only do many educators already have favored research access points that the WWC must compete

with, but the WWC's outputs do not share the characteristics of other popular access points that make them appealing. Both administrators (e.g., GAO, 2010; Penuel et al., 2017) and teachers (e.g., Barton & Tindle, 2019; Drill et al., 2013) tend to favor research access points that involve relational exchanges with colleagues—something the WWC's written reports do not allow for. The WWC is also unequipped to address external constraints to educators' research-seeking (e.g., Levin, 2011; Tseng, 2012), such as lack of time. To be sure, the WWC is well aware of these constraints. Specifically, it explicitly recognizes that it can be "time-consuming" for practitioners to engage with educational research, which is why it tries to identify, vet, and summarize research for them (see WWC, 2020d). It also does a commendable job developing resources that are concise and quickly digestible, which was noted in the focus groups. Even so, the WWC cannot afford potential users the time to engage with it; doing so is beyond its control. As such, educators' inability to seek out the WWC's outputs is another likely fracture in its theory of change.

**Educators' perceptions of the trustworthiness and/or relevance of research are far more nuanced than the WWC's own.** To become motivated to use the WWC's outputs to guide practice, practitioners must find them trustworthy and relevant. The extant evidence suggests that the WWC's characterization of high-quality research—which emphasizes indicators of internal validity—resonates with the characterizations held by some teachers (Borg, 2009, 2012; Cain, 2016). At first blush, this evidence bodes well for the WWC and its perceived rigor among these educators. Upon closer examination, however, this contribution analysis uncovered several patterns in the WWC's extractable data that may—and do—leave these educators skeptical of the

WWC's quality. How so? Educators who view large sample sizes as an indicator of research quality might be put off by the fact that the median sample size among intervention reports synthesizing at least one available study is 242. Likewise, they may be alarmed that over three-quarters of those intervention reports only synthesize two studies or less. Also noteworthy is the finding that almost half of all practice guide recommendations are rated by the WWC as having "minimal" supporting evidence, and furthermore, that this was concerning to those participating in the Mathematica focus groups. Hence, the WWC's own preoccupation with issues of internal validity may, in effect, be setting a standard that is hindering rather than helping its cause, and others have suggested similarly (e.g., GAO, 2010).

Despite some educators' concerns about internal validity, most are more concerned about the external validity of the research they encounter (e.g., Neal et al., 2018; Joram et al., 2020). This was recognized by the WWC, who, shortly after conducting their user feedback campaign, redesigned their website so that educators could more easily locate research conducted with "students like yours" (IES, 2018c). Unfortunately, analyses conducted as part of the contribution analysis found that over half of all intervention report outcomes (based on at least one eligible study) lacked critical demographic information—such as a student's ELL or free/reduced lunch status—that would allow users to make these determinations. Many lacked information about school-level factors as well, such as a school's urbanicity. Given that teachers are more "receptive to generalization/transfer as long they [can] establish a high degree of similar between their own students/context and that of the original study" (Joram et al.,

2020, p. 6), this lack of contextualizing information featured in the WWC's outputs may limit both their trustworthiness and their relevance.

The WWC's relevance may be lacking for other reasons. Consider, for example, the issue of timeliness. Outdated information is of little relevance to educators, especially when considering that they are most likely to consult research when confronting an immediate problem in their practice (e.g., Drill et al., 2012). Not only did my analysis find evidence of a sizable backlog between when an individual study review was completed and when it was posted, but it also found that over half of all intervention reports and practice guides available to beneficiaries were published prior to 2010. This datedness has been noted elsewhere (e.g., Slavin, 2017). Furthermore, not all of the information available is of interest to educators. The WWC's topic areas and outcomes of interest sometimes align with educators' own, but other times they do not (e.g., Barton & Tindle, 2019; Penuel et al. 2018). This may be why respondents in the GAO (2010) evaluation said they would be more likely to use the WWC if it featured additional topic areas. In their current state, however, the topic areas do not account for the diversity of research interests held by administrators and educators; this, too, diminishes the WWC's relevance.

**Teachers face an array of barriers to using educational research—even if they wanted to.** Perhaps the most devastating blow to the WWC's theory of change is this one. Even when everything goes according to plan—namely, when WWC outputs reach educators, who, in turn, become motivated to use them in practice—research-informed practice is still at the mercy of school-, district-, and state-level institutional barriers. Besides the structural factors noted above—such as the lack of time educators

are given for research engagement (e.g., Cain, 2016)—there are also impediments posed by the organizational cultures in which educators function. For example, investigations of educators' research use have repeatedly found a research-friendly organizational culture to be a predictor of use for both teachers (e.g., Lysenko et al., 2014) and administrators (e.g., Penuel et al., 2017). A component of this culture is the amount of agency afforded to teachers to try new things, including the implementation of research-based practices. Just as previous work has found a lack of agency to interfere with teachers' research use (e.g., Joram et al., 2020), this contribution analysis found evidence that a lack of agency may interfere with teachers' use of WWC outputs in their decision-making. Participants in both the Mathematica focus groups and the teacher survey spoke to how they were at the mercy of curricular decisions made by administrators, and as such, that the information produced by the WWC may be better-suited for those administrators than for teachers with little decision-making power. Such a finding does not rule out the WWC's impact in, at the very least, informing the actions of administrators, but it does suggest that its goal of facilitating teachers' research use may be more aspirational than achievable.

### Question 3: How Can the WWC Be Improved?

Until now, this chapter has focused on the summative elements of my contribution analysis—meaning, the elements lending themselves to a judgement about whether or not the WWC is making a meaningful contribution to closing the research-practice gap. In short, this project finds that it is not, and moreover, that its failure is (at least partially) due to refutable programmatic assumptions. But, as explicated in Chapter 1, this project began with a formative intention—namely, an intention to "support the

process of improvement" (Scriven, 1991, as cited in Patton, 1996, p. 131). Specifically, it sought to elicit improvements by offering recommendations to (a) help the WWC better fulfill its own mission and (b) help it more thoroughly contribute to the IES's larger research dissemination goals. Not only are these formative intentions implied in the very nature of my third evaluation question, but they are demanded by both the evaluative and methodological inspirations of this project. For example, evaluative thinking theorists have spoken of evaluation as an educative (e.g., Vo, Schreiber, & Martin, 2018) or pedagogical (e.g., Patton, 2017) activity. Likewise, proponents of theory-based evaluation—a school of thought to which contribution analysis belongs (Mayne, 2012b)—explain how understanding a program's innerworkings is a critical step in "providing important information for program improvement" (Chen & Rossi, 1989, p. 302). Even users of the contribution analysis methodology have promoted its usefulness as a quality improvement tool (Biggs, Farrell, Lawrence, & Johnson, 2014). It would therefore be a mistake to consider this project's summative insights without also remembering its formative intentions.

Given the answers to my first two evaluation question, I foresee two ways to answer my third. The first is to offer suggestions for reconciling the WWC's questionable causal assumptions. Indeed, based on the evidence uncovered in my contribution analysis, countless changes could be made to the WWC—many of which have been recommended elsewhere. For example, to build it reviewing capacity, the WWC could take after other systematic review initiatives (e.g., the United Kingdom's EPPI-Centre; UCL, n.d.) or research enterprises (e.g., the Collective Replications and Education Project; Grahe et al., 2020) and invite undergraduate and/or graduate students to serve as

individual study reviewers. To better reach educators, the WWC could heed the suggestions of the Mathematica focus group participants and partner with teacher education programs to connect with novice teachers. To improve its relevance, it could consider additional topic areas of research. All of these are viable improvements.

Though viable, they may not be enough. Why? Because recommendations targeting the WWC's causal assumptions would do little, if anything, to remedy the fallibility of its rationale assumptions. As demonstrated in Chapter 4 and repeated earlier in this chapter, the WWC is built upon several refutable beliefs about the nature of the research-practice gap, the causes of the gap, and the best ways of closing it. Accordingly, in the remainder of this paper, I offer four recommendations for retooling the WWC given what the evidence says about these rationale assumptions. This is the other way of answering my third guiding question, and it is the path forward that I support.

**Gather data on how the implementation of WWC-provided information has affected practice, if at all.** This project would not have been possible without the generous amount of data the WWC collects on itself, and moreover, makes available to the public. Although there is evidence that the WWC uses some of this data to assess its own progress (e.g., WWC, n.d.-f, 2018, 2019), additional self-study could and should occur. Indeed, implicit to the WWC's normative assumption that educational research should guide educational practice is the belief that "when research tells practitioners 'what works', their teaching becomes research-informed and thereby, of better quality" (Cain, 2017, p. 3). Unfortunately, as acknowledged by Cain (2016), there is surprisingly little evidence that bridging the educational research-practice gap actually has "trickle down" benefits for students, and this holds true to the WWC as well (e.g., Gorard et al.,

2020). For example, what is known about the effects of a teacher deciding to implement a practice guide recommendation in their classroom? Do student outcomes really change as a result? Or, along similar lines, what is known about the effects of a principal implementing an educational intervention that is deemed to have "positive effects" according to a WWC intervention report?

Answering these questions requires several things—among them a commitment for the WWC to study itself, and moreover, an availability of resources for it to do so. I can envision at least three ways in which this kind of self-study might occur. The first is to encourage educators to offer testimonials about their experiences executing WWC practice guide recommendations. Just as blogs and social media sites allow for users to comment on content, the WWC website might consider doing the same. Specifically, I envision each webpage with links to a specific practice guide to also allow users to comment on their experiences using it. So, too, do the WWC's users. One participant in the Mathematica focus groups suggested that the WWC add "…links to curricula, connections to online communities where teachers and others can discuss experiences putting these practices into action."

There are several potential benefits to doing so. One benefit is that, by offering these testimonials, users will be providing the WWC with additional metadata about which guides are most popular, which are most/least effective at eliciting their desired outcomes, and which need updating. It would also provide users an opportunity to offer recommendations for how the guides might be improved moving forward. A second benefit would be the cultivation of a more active research access point, and perhaps more importantly, one that enables educators to build relationships and exchange information

with one another. Teachers favor these types of access points (e.g., Tseng, 2012), so offering this type of functionality could make the WWC a favored research access point in turn.

A second way this self-study might occur is by asking schools and/or districts implementing WWC-vetted interventions to report back on the effects of that implementation. This could be feasible given that some users likely engage with the WWC for "imposed" reasons (e.g., Weiss, 2008), such as when securing funding for a new curriculum or program requires it to have been vetted by the WWC (i.e., the Investing in Innovation Fund; Boulay et al., 2018). In these cases, the grantee might collect data as part of grant requirements, or they may choose to conduct an internal evaluation in response to their own interests. Results, or the raw data themselves, would then be reported back to IES, who could then use them to supplement the WWC's current evidence base. Although the WWC may want to keep these findings separate from its traditional research synthesis work, these school and/or district-generated findings would lend a unique crowdsourced element to the WWC, and doing so poses several potential benefits. For example, when considering the large number of intervention reports only featuring one or two studies (WWC, 2020b), the WWC may not be in a position to forsake additional data—even if kept supplemental from the actual reports. These user-generated findings would also offer insights into which types of interventions are more easily implemented with fidelity, and additionally, which interventions still prove effective despite variations in implementation. Finally, this crowd-sourced component may have a special allure to educators, some of which distrust research done by those

who have not experienced the realities of the classroom (e.g., Behrstock-Sherratt et al., 2011).

A third way the WWC could engage in self-study would be to conduct research itself. This option is likely the most resource-intensive, but it may also be the most informative. It could also proceed in several ways. As an example, rather than ask WWC users for feedback about potential WWC outputs (as was done during its user feedback study; WWC, n.d.-f), it could be more useful to hear from WWC users about how they engage with current outputs, such as intervention reports. A different kind of study might involve experimental research, in which a random group of teachers is given a WWC practice guide and a comparison group is not. Teachers could be surveyed about their use of the practice guide, and the outcomes of their students (particularly in domains addressed in the practice guide) could be compared to those from teachers who were not given the guide. Regardless of how this kind of self-study occurs, the simple act of thinking evaluatively about itself (e.g., Vo & Archibald, 2018) could help the WWC think beyond its mission and instead consider the assumption that attaining it will, indeed, benefit students.

**Commit to a renewed focus on empowering educators to be thoughtful consumers of research.** One of the diagnostic assumptions harbored by Whitehurst (2003) and others (Hearing Before the House Subcommittee, 2002b) implicated educators' inability to critically consume educational research. They believed that most educators lacked the "information literacy" (e.g., Williams & Coles, 2007b) or "data literacy" (e.g., Henderson & Corry, 2020) needed to be thoughtful consumers of research. In its current form, the WWC does little to foster educators' data literacy; instead, it

attempts to circumvent the issue by finding, reviewing, and summarizing educational research on their behalf. This approach overlooks the fact—which is documented in the literature as well as comments in the Mathematica focus groups—that many educators are eager to refine their data literacy skills. What's more, the literature generally supports the notion that educators' participation in research can help close the research-practice gap to various degrees, either by (a) improving their attitudes about the value of education research (e.g., Cantalini-Williams et al., 2015) and/or by (b) increasing the likelihood of research use (e.g., Cousins & Simon, 1996); however, more research is needed with regard to this second point.

I envision at least two contributions the WWC could make with regard to educators' data literacy. The first involves updating a research review area that, though increasingly relevant, has been neglected since 2009. In September 2009, the WWC published a practice guide entitled "Using Student Achievement Data to Support Instructional Decision Making" (WWC, 2009a; see also WWC, 2009b). The guide, which includes recommendations such as "Provide supports that foster a data-driven culture within the school" and "Develop and maintain a districtwide data system," may have been ahead of its time, especially when considering that all of its recommendations were based on "minimal evidence." This suggests that there was a dearth of research evidence when the guide was developed. Updating this guide necessitates an update to its associating review protocol. One reason to update the protocol is the surge in scholarly attention being paid to educators' data literacy (e.g., Mandinach & Gummer, 2013; Henderson & Corry, 2020), which means that there is additional research requiring review. Another is that much of that literature examines interventions designed to

improve data literacy (e.g., van Geel et al., 2017; Kippers, Poortman, Schildkamp, & Visscher, 2018). If the WWC wishes for school- and district-level administrators to choose data literacy trainings that do, in fact, "work," then research evaluating these interventions should be identified, vetted, summarized, and disseminated. This requires virtually no changes to the WWC's current approach, and it would also ensure that the WWC reviews a topic area that is clearly of interest to educators.

A second, more drastic change to the WWC involves a shift in priorities—from its disseminative focus to a more educative one. Instead of conducting the research review process on behalf of its users, what if it coached them in how to review research themselves? Although the IES's Regional Education Laboratories (RELs) are already confronting this challenge (IES, n.d.-b), the WWC is well-positioned to do similarly. For example, along with its public list of WWC-certified reviewers (WWC, 2020a), the WWC also maintains its own training and certification process (WWC, n.d.-h) that aspiring reviewers must complete in order to become certified. Its nine-module *Group Design Standards Online Training* involves modules dedicated to "confounding factors" and "cluster-level assignment", as well as a multiple-choice certification test. Likewise, the WWC has produced videos and webinars showing WWC users on how best to navigate its outputs (e.g., WWC, 2016a). Given that users, such as those participating in the Mathematica focus groups, have research interests that are yet to be featured by the WWC, the WWC could enhance its usefulness by offering additional modules or videos on how its users might find, review, and make sense of research evidence *outside* of the WWC. Any concern that doing so would, in effect, be self-defeating for the WWC should be considered in light of the potential benefits that might result.

**Consider other approaches to research synthesis, and in turn, how elements from those approaches could be incorporated into the WWC's own.** As explained in Chapter 1, the WWC is just one example of a larger trend involving the use of systematic research reviews to bridge research-practice gaps—both in education (Oakley, 2002) as well as in other policy areas (Boaz, Ashby, & Young, 2002). It embodies the assumption that systematic research review is an effective way of closing research-practice gaps. The hallmark of systematic review is:

> "produc[ing] *new* knowledge by making explicit connections and tensions between individual study reports that were not visible before. It involves purposeful selection, review, analysis, and synthesis of primary research reports on a similar topic." (Suri, 2013, p. 889, emphasis in original)

This can be accomplished through a variety of approaches, each of which has certain defining characteristics. The WWC's approach to systematic review has been called a "threshold, rule-based approach" (Stockard & Wood, 2017), or, in a similar way, an iteration of Slavin's (1986) "best-evidence" approach. In both cases, the WWC is exercising what has been called a "positivist orientation" towards research synthesis, which is characterized by (a) the use of a priori review protocols that are meant to minimize research bias, (b) the goal of estimating the overall effectiveness of an educational strategy, and (c) assuming a top-down approach to educational change (e.g., Suri, 2013).

Although some scholars have lauded the WWC's review approach (Slavin, 2004, 2008), others have noted its limitations. One limitation is that it sets an unreasonable standard given the realities of educational research (e.g., Berliner, 2002; Lykins, 2012); in other words, its standards of evidence are too strict. As a result, studies have found that it overlooks potentially useful studies when making conclusions, and consequently, that

the conclusions it does make are often made according to a small number of studies (GAO, 2010; Stockard & Wood, 2017). This may be why some more inclusive reviews of evidence-based educational practices have yielded contradictory conclusions to those drawn by the WWC (Stockard & Wood, 2017), as have systematic reviews that prioritize their own set of standards (e.g., Slavin & Madden, 2008). Another limitation is that positivist syntheses, despite appearing more systematic than other approaches (and thus less prone to bias and errors), are no less fallible than their alternatives. Indeed, there is evidence of the WWC's rules-based approach being poorly implemented despite quality-control procedures (e.g., McArthur, 2008). Others have shared inconsistencies and errors in their personal communications with WWC staff (Stockard, 2010), or even discussed how their opinions were actively suppressed while serving in a content advisory role for the WWC (Schoenfeld, 2006). Any effort to facilitate educational reform is sure to face criticism, but the abundance of criticism about the WWC's systematic review approach is concerning, especially when it comes from those who championed the initiative in the first place (e.g., Slavin, 2017)

Criticism aside, however, it is important to recognize that a positivist, rules-based approach is not the only way to synthesize research, nor is it the only way to do so in a manner that makes conclusions about "what works." Other types of reviews offer alternatives that may be worth integrating as the WWC moves towards its third decade of existence. Consider participatory approaches to research synthesis. By shifting priorities from the "objective distancing of an unbiased expert" to "practical experience, local knowledge, and serendipitous leaps of intuitive understanding" (Suri, 2013, p. 899), a participatory orientation towards systematic review might also strengthen a review's

usefulness to practitioners (Boaz, Ashby, & Young, 2002). Likewise, Wolgemuth and colleagues (2017) advocate for the importance of interpretative/critical synthesis approaches, which attend just as much—if not more—to how studies operationalize a certain concept or content area (i.e., its construct validity) than to its methods and design.

At first blush, integrating elements from these other approaches may sound unrealistic. Nevertheless, there is an example of a systematic review initiative doing exactly that. Preceding the WWC by several years, a systematic review initiative called the *Evidence for Policy and Practice Information and Coordinating Centre* (EPPI Center) was funded by the United Kingdom's *English Department for Education and Skills* in 2000. Oakley (2003) details the Center's aims and processes in a way that echoes the WWC's own—so much so that she likens it to the What Works Clearinghouse, which had recently been announced at the time of her paper's publishing. There are several notable differences, though. These can be seen in, among other things, the key principles informing the EPPI Center's review process. Like the WWC, it believed that systematic reviews must proceed according to a predetermined protocol "specifying a particular, answerable research question, and criteria about what kinds of studies will be included in, and excluded from, the domain of literature to be surveyed" (Oakley, 2003, p. 24). Unlike the WWC, however, it also believed that a systematic review "is credible only if it has involved input from research users at all stages of the review process" (Oakley, 2003, p. 24). This "user-driven" approach to research review attempts to include students, parents, and educators when "deciding which topics need most urgently to be reviewed, and analyzing and disseminating the results of reviews" (EPPI Center, n.d.). Surely, as admitted by Oakley (2003), facilitating this kind of inclusion is challenging. Even so, the

WWC could start by (a) more actively eliciting topic area suggestions by educators through the WWC Help Desk, and (b) including educators in the front end of its review process (i.e., in review protocol development) rather than just the back end (i.e., when developing practice guides).

**Consider alternatives to systematic review.** While gathering evidence for this contribution analysis, it became clear that alternatives to the WWC's approach abound. Closing the research-practice gap is challenging no matter the approach taken, however there are alternatives to systematic research review that have shown promise. Here I examine three alternatives that may lend themselves better to the IES's goal of "…increas[ing] use of data and research in education decision-making" (IES, n.d.-a), including (a) research champions, (b) research-practice partnerships, and (c) teacher action research.

***Research champions.*** In response to the question "what is the best way to get evidence into use in education?," Gorard and colleagues' (2020) review of the literature found the use of "research champions" to be a promising approach. Research champions are school-level educators who are trained to engage with educational research, collaborate with colleagues to encourage research use, and work to build research capacity at their institutions (e.g., Burn, Conway, Edwards, & Harries, 2020). Although there is little evidence that the employment of a research champion is associated with improved student outcomes, Gorard et al. (2020) did find three studies documenting the positive impact of research champions on the research-related attitudes of their colleagues. Hence, there is some evidence that the research champion model is effective.

The research champion model is also compelling because it capitalizes on what is known to facilitate research use among educators in the first place. For instance, it capitalizes on the fact that educators are most likely to access research through their colleagues (e.g., GAO, 2010; Penuel et al., 2017). It also offers educators an interactive research access point, as research champions can—unlike a written report—converse with information users, clarify points of confusion, and acknowledge points of concern. Given their experience as educators, research champions might also be well-positioned to demonstrate research-based recommendations in the classroom, as opposed to just telling their fellow educators what to do.

At this point, the research champion model seems to have received greater consideration in the United Kingdom than the United States. Perhaps the closest thing to a research champion model managed by IES is its Regional Education Laboratories (RELs). Although the RELs are devoted to generating "genuine partnerships" with the educators they serve (IES, n.d.-b), they are best understood as a derivative of the research champion model—not as an embodiment of it. There are several key differences between the two. Indeed, though the RELs core function of providing "training, coaching, and technical support" (IES, n.d.-b) to educators is similar to that of a research champion, these services are provided by research contractors rather than fellow educators. For example, the REL serving the U.S. Appalachia region (which includes Kentucky, Tennessee, Virginia, West Virginia as defined by IES) is managed by a research institute called SRI International (IES, n.d.-b), which is based out of California. A second difference, which is perhaps a result of the first, is that administrators and teachers use the RELs even less frequently than they use the WWC (Barton & Tindle, 2019; Penuel et

al., 2017). Given the findings reviewed by Gorard et al. (2020), a true iteration of the research champion model should be more (rather than less) compelling to educators.

There are two ways the IES might contribute to a research champion model moving forward. If the IES proves reluctant to alter the WWC is some of the ways I suggest, it could train WWC-specific liaisons at the school- or district-level. These liaisons could help distribute WWC resources to colleagues, respond to colleagues' research questions by reviewing the WWC database themselves, offer professional development seminars, and even relay feedback from their colleagues (e.g., ideas about new WWC topic areas) to the WWC so that subsequent changes could be made. A second option would be for the IES to help train research champions that are not WWC-specific. They would be permitted to offer guidance to fellow teachers without being limited by the WWC's standards of evidence or topic areas.

Both options would help address oft-cited barriers to research use as well as barriers to educators' use of the WWC more specifically. Because most teachers do not have time to engage with research themselves (e.g., Cain, 2016; Schaik et al., 2018), having one teacher do so on behalf of their colleagues helps manage these time constraints, especially if that liaison is recognized at the district- and/or school-level and thus given a reduced teaching load. A second barrier is teachers' data literacy (e.g., Williams & Coles, 2007), which, though improving, is still not ubiquitous. While efforts to improve all teachers' data literacy could and should continue, building the data literacy of a single teacher—specifically, that of the WWC liaison—would be more feasible. The liaison position could also try to recruit teachers who possess that knowledge already. A third barrier involves some educators' skepticism of educational research, and more

specifically, their concerns that "research is done by people who don't spend time in classrooms and who don't know students" (Behrstock-Sherratt et al., 2011, p. 5). When that research is shared through a fellow educator, however, skeptical teachers may be more receptive.

***Research-practice partnerships (RPPs).*** Social scientists have continued to advance methodologies premised on the power of fostering collaboration between members of the research and practice communities—a true "bridging" of the research-practice gap. Rather than a researcher on one side and a practitioner on the other, these approaches advocate for the practitioner becoming part of the research process. One example is research-practice partnerships (RPP; e.g., Coburn & Penuel, 2016), which are "long-term collaborations between practitioners and researchers that are organized to investigate problems of practice and solutions for improving schools and school districts" (p. 48).

A growing body of research speaks to both the enthusiasm for and the efficacy of this approach. Despite the challenges inherent to these types of collaborations (e.g., Agans et al., 2020), both practitioners and researchers have voiced excited about these opportunities (e.g., Beveridge, Mockler, & Gore, 2018). Existing reviews of the literature (Dagenais et al., 2012; Schaik et al., 2018) suggest research-practice partnerships to be a useful way of encouraging research use. Some evidence even suggests that these partnerships hold promise in improving "cross-cultural" relations between the research and practice communities—such as mitigating negative attitudes about educational research held by some teachers (e.g., Ratcliffe et al., 2005) or improving researchers'

own appreciation for the practical knowledge held by teachers (e.g., Beveridge et al., 2018).

Indeed, since 2013, the IES has recognized the fruitfulness of such an approach. They have operated a funding mechanism dedicated to research-practice partnerships, which "supports partnerships composed of research institutions and state or local education agencies that have identified an education issue or problem of high priority" (IES, n.d.-c). A high-level analysis of data on the IES's "Research-Practice Partnerships in Education Research" funding program indicated that, since its conception, 62 RPP grants have been awarded for a total of $24.3 million (IES, 2020). Evidence on the success of these partnerships is still emerging, but based on preliminary findings released by the IES's own National Center for Research in Policy and Practice (e.g., Farrell et al., 2018), investment in these partnership should continue. If anything, I recommend that more funding be made available—even if that means diverting funds from the WWC. Not only is this justifiable because of the WWC's wanting impact, but because of the vast funding differences between the two initiatives. Indeed, compared to the $24 million dollars awarded to RPPs since 2013, the IES has spent roughly three times as much (i.e., $69.8 million) on WWC-related contracts (see Figure 4.2).

***Teacher action research.*** Finally, a third alternative to systematic review would be to encourage practitioners to conduct educational research themselves. Teacher action research is founded on the idea that "educational research should be an integral part of the work of teachers in schools rather than an activity carried out on schools by outsiders" (Hammersley, 1993, p. 425; see also Schön, 1995 and Stenhouse, 1980). This alternative is most directly aligned with the recommendations of Dewey (1929) and

Freire (1968) addressed in Chapter 1, and it remains a popular one in studies of the educational research-practice divide (e.g., Ostinelli, 2016). Yet, of the three alternatives to systematic review discussed in this chapter, it is also least congruent with the WWC's current approach.

Despite this incongruence, there are a host of reasons why teacher action research would be one of the most effective ways—if not the most effective—of bringing educational research to bear in practice. The first is that research dissemination becomes a non-issue. As demonstrated in this contribution analysis, one of the main failures in the WWC's theory of change is that its outputs are failing to reach teachers, and to a lesser extent, administrators (Barton & Tindle, 2019; GAO, 2010). When research is conducted by teachers within their own educational settings, it is already within their reach. A second is that the WWC's capacity change assumptions, which require that the WWC is perceived as relevant and trustworthy, also become non-issues. In teacher action research, teachers are free to investigate issues that are of the upmost relevance to them. They are also free to subject their work to broader definitions of "quality" and "rigor" than are currently allowed for by the WWC (e.g., Groothuijsen et al., 2019). As a result, their research is likely perceived to have great external validity than research conducted in settings that are not their own.

However, just as an educator's use of WWC outputs is contingent upon institutional factors—such as a school or district's openness to educational research—so too is the reality of teacher action research. In settings where teachers may not be afforded much agency to deviate from the curriculum or explore other research-based teaching strategies, they are likely to also be discouraged from conducting their own

research projects (e.g., Borg, 2008, 2009). Accordingly, facilitating teacher action research requires vast institutional change—one where this type of research is cultivated, supported, and valued. The IES might help facilitate this change by demonstrating support for teacher research themselves. As recommended above, this support might include transitioning the WWC from a research dissemination platform to one that helps educators develop their own data literacy and research skills. Another way they could show support is by shifting the WWC into a clearinghouse for teacher action research—that is, a place where teacher-researchers could share their findings with their colleagues. A third idea is for the IES provide funding to sponsor teacher buy-outs, whereby individual teachers could be granted semester or year-long sabbaticals to conduct their own research projects while being relieved from their teaching responsibilities. Regardless of the path forward, the IES can no longer afford to overlook the powerful institutional factors that interfere with educators' research engagement.

## Conclusion

To conclude, it seems appropriate to return to this project's three guiding questions. First, is "what works" working? Just as Finn (2008) admitted with regard to his *What Works* booklet, I have similarly concluded that the *What Works Clearinghouse* (WWC) has "had little impact on its primary audiences" (Finn, 2008, p. 135)—especially among teachers. Second, why might its impact be wanting? The primary explanation for its lack of impact is that its theory of change rests on several ill-founded assumptions. Not only were many of the WWC's causal assumptions refuted by the evidence, but some of its rationale assumptions—such as the belief that a systematic research review initiative would be an effective way of bringing educational research to practice—were

refuted as well. Third, how might the WWC be improved? Given the fragility of its rationale assumptions, the WWC may only be able to strengthen its impact if it fundamentally retools its approach to systematic research review or to research dissemination more generally. When taken together, these three points compose my contribution claim. They also ring true to the words of Coburn, Honig, and Stein (2009), who, in a single sentence, convey an analogous argument to the one presented in these five chapters:

> "Yet, underneath the calls for school districts to use evidence sit a set of assumptions about the nature of evidence and evidence use that may not accurately reflect the realities of decision making in public bureaucracies such as school districts." (p. 3)

Only when there is congruence between the IES's assumptions about educational research and the assumptions held by educators will genuine progress be made in closing the research-practice gap. And, as is so often suggested in the literature (e.g., Hemsley-Brown & Sharp, 2003; Dagenais et al., 2012; Schaik et al., 2018), this congruence likely requires more collaboration between the research and practice communities.

APPENDICES

## Appendix A - R Syntax for WWC Extractable Data

```
######################## Load datasets ########################

library(readxl)
Studies <- read_excel("~/Downloads/Studies.xlsx")

library(readxl)
Studies <- read_excel("~/Downloads/InterventionReports.xlsx")

######################## Load packages ########################

library(Rcmdr)
library(dplyr)
library(summarytools)

################# Number of studies reviewed over time #################

#Create new variable for ReviewedDate
Studies$ReviewedDate_new <- Studies$ReviewedDate

#Trim timestamp off of ReviewedDate_new
Studies$ReviewedDate_new <-
format(as.POSIXct(Studies$ReviewedDate_new,format='%m/%d/%Y
%H:%M:%S'),format='%m/%d/%Y')

#Extract year from ReviewedDate_new
Studies$ReviewDate_year <- as.numeric(substring(Studies$ReviewedDate_new,7,10))

#Frequenecy Table
summarytools::freq(Studies$ReviewDate_year, order = "freq")

################ Number of studies posted over time ###############

#Create new variable for PostingDate
Studies$PostingDate_new <- Studies$Posting_Date

#Trim timestamp off of PostingDate_new
Studies$PostingDate_new <-
format(as.POSIXct(Studies$PostingDate_new,format='%m/%d/%Y
%H:%M:%S'),format='%m/%d/%Y')

#Extract year from PostingDate_new
Studies$PostingDate_year <- as.numeric(substring(Studies$PostingDate_new,7,10))
```

```
#Frequenecy Table
summarytools::freq(Studies$PostingDate_year, order = "freq")

############## Backlog between study review and study posting#################

#Backlog

Studies$Backlog <-
difftime(Studies$PostingDate_new,Studies$ReviewedDate_new,units = 'day')

Studies$Backlog_num <- as.numeric(Studies$Backlog)


numSummary(Studies[,"Backlog_num", drop=FALSE], statistics=c("mean", "sd",
                                              "IQR", "quantiles"), quantiles=c(0,.25,.5,.75,1))

## Number of Studies Meeting Standards Used in Each Intervention Report Outcome ##

#Make "NumStudiesMeetingStandards" numeric
InterventionReports$NumStudiesMeetingStandards <-
as.numeric(InterventionReports$NumStudiesMeetingStandards)

#Histogram
with(InterventionReports, Hist(NumStudiesMeetingStandards, scale="frequency",
                        breaks="Sturges", col="darkgray"))

#Descriptive Statistics
numSummary(InterventionReports[,"NumStudiesMeetingStandards", drop=FALSE],
         statistics=c("mean", "sd", "IQR", "quantiles"), quantiles=c(0,.25,.5,.75,1))

#Frequency Table
local({
  .Table <- with(InterventionReports, table(NumStudiesMeetingStandards))
  cat("\ncounts:\n")
  print(.Table)
  cat("\npercentages:\n")
  print(round(100*.Table/sum(.Table), 2))
})

############ Average sample size in intervention report outcomes ############


#Filter for cases with a least one eligible study
InterventionReports_atLeast1 <- subset(InterventionReports,
                                subset= NumStudiesMeetingStandards >= 1)

#Descriptive statistics for "Sample_Size_Intervention"
```

```
numSummary(InterventionReports_atLeast1[,"Sample_Size_Intervention",
                                drop=FALSE], statistics=c("mean", "sd", "IQR", "quantiles"),
quantiles=c(0,.25,.5,.75,1))


############ Demographic data from intervention reports ############


#Load "summarytools" package
library(summarytools)

#Frequencies for Ethnicity
freq(InterventionReports_atLeast1$Ethnicity_Hispanic)
freq(InterventionReports_atLeast1$Ethnicity_Not_Hispanic)

#Frequencies for Race
freq(InterventionReports_atLeast1$Race_White)
freq(InterventionReports_atLeast1$Race_Black)

#Frequencies for Disabilities
freq(InterventionReports_atLeast1$Demographics_of_Study_Sample_Students_with_dis
abilities)

#Frequencies for Free/Reduced Lunch
freq(InterventionReports_atLeast1$Demographics_of_Study_Sample_Free_or_reduced_
price_lunch)

#Frequencies for ELL status
freq(InterventionReports_atLeast1$Demographics_of_Study_Sample_English_language_
learners)

#Frequencies for Gender
freq(InterventionReports_atLeast1$Gender_Female)
freq(InterventionReports_atLeast1$Gender_Male)

#Frequencies for School Type
freq(InterventionReports_atLeast1$School_type_Charter)
freq(InterventionReports_atLeast1$School_type_Parochial)
freq(InterventionReports_atLeast1$School_type_Public)
freq(InterventionReports_atLeast1$School_type_Private)

#Frequencies for School Setting
freq(InterventionReports_atLeast1$Urbanicity_Rural)
freq(InterventionReports_atLeast1$Urbanicity_Suburban)
freq(InterventionReports_atLeast1$Urbanicity_Urban)

######## Demographic data from individual studies ########
```

```
#Load "summarytools" package
library(summarytools)

#Load "dplyr" package
library(dplyr)

#Only select studies that met WWC standards with or without reservations
Studies_onlyMetStandards <- filter(Studies, Study_Rating == "Meets WWC standards
without reservations" |
                        Study_Rating == "Meets WWC standards with reservations")

#Frequencies for Ethnicity
freq(Studies_onlyMetStandards$Ethnicity_Hispanic)
freq(Studies_onlyMetStandards$Ethnicity_Not_Hispanic)

#Frequencies for Race
freq(Studies_onlyMetStandards$Race_White)
freq(Studies_onlyMetStandards$Race_Black)

#Frequencies for Disabilities
freq(Studies_onlyMetStandards$Demographics_of_Study_Sample_Students_with_disabi
lities)

#Frequencies for Free/Reduced Lunch
freq(Studies_onlyMetStandards$Demographics_of_Study_Sample_Free_or_reduced_pri
ce_lunch)

#Frequencies for ELL status
freq(Studies_onlyMetStandards$Demographics_of_Study_Sample_English_language_le
arners)

#Frequencies for Gender
freq(Studies_onlyMetStandards$Gender_Female)
freq(Studies_onlyMetStandards$Gender_Male)

#Frequencies for School Type
freq(Studies_onlyMetStandards$School_type_Charter)
freq(Studies_onlyMetStandards$School_type_Parochial)
freq(Studies_onlyMetStandards$School_type_Public)
freq(Studies_onlyMetStandards$School_type_Private)

#Frequencies for School Setting
freq(Studies_onlyMetStandards$Urbanicity_Rural)
freq(Studies_onlyMetStandards$Urbanicity_Suburban)
freq(Studies_onlyMetStandards$Urbanicity_Urban)
```

```
#############Rating Breakdown of Reviewed Studies ###############

local({
  .Table <- with(Studies_onlyMetStandards, table(Study_Rating))
  cat("\ncounts:\n")
  print(.Table)
  cat("\npercentages:\n")
  print(round(100*.Table/sum(.Table), 2))
})

##################Topic areas of individual studies##################

#Load "summarytools" package
library(summarytools)

#Load "dplyr" package
library(dplyr)

#Only select studies that met WWC standards with or without reservations
Studies_onlyMetStandards <- filter(Studies, Study_Rating == "Meets WWC standards
without reservations" | Study_Rating == "Meets WWC standards with reservations")

#Frequencies of Topic Areas
freq(Studies_onlyMetStandards$Topic_Behavior)
freq(Studies_onlyMetStandards$Topic_Mathematics)
freq(Studies_onlyMetStandards$Topic_Science)
freq(Studies_onlyMetStandards$Topic_Postsecondary)
freq(Studies_onlyMetStandards$Topic_Teacher_Excellence)
freq(Studies_onlyMetStandards$Topic_Literacy)
freq(Studies_onlyMetStandards$Topic_Early_Childhood)
freq(Studies_onlyMetStandards$Topic_K_to_12th_Grade)
freq(Studies_onlyMetStandards$Topic_Charter_Schools)
freq(Studies_onlyMetStandards$Topic_Path_to_Graduation)
freq(Studies_onlyMetStandards$Topic_ELL)
freq(Studies_onlyMetStandards$Topic_SWD)

############Protocols Featured in Studies / Intervention Reports ##############

#Frequencies of Protocols Featured in Intervention Reports

local({
  .Table <- with(InterventionReports_atLeast1, table(Protocol))
  cat("\ncounts:\n")
  print(.Table)
  cat("\npercentages:\n")
  print(round(100*.Table/sum(.Table), 2))
```

```
})

#Frequencies of Protocols Featured in Studies

local({
  .Table <- with(Studies_onlyMetStandards, table(Protocol))
  cat("\ncounts:\n")
  print(.Table)
  cat("\npercentages:\n")
  print(round(100*.Table/sum(.Table), 2))
})

############Outcome Domains Featured in Intervention Reports #############

local({
  .Table <- with(InterventionReports_atLeast1, table(Outcome_Domain))
  cat("\ncounts:\n")
  print(.Table)
  cat("\npercentages:\n")
  print(round(100*.Table/sum(.Table), 2))
})

##################Effectiveness Ratings of Intervention Reports###################

local({
  .Table <- with(InterventionReports_atLeast1, table(Effectiveness_Rating))
  cat("\ncounts:\n")
  print(.Table)
  cat("\npercentages:\n")
  print(round(100*.Table/sum(.Table), 2))
})
```

## Appendix B – Focus Group Protocols

### In-Person Focus Group Protocol

### 1) Familiarity with the WWC

a) Thinking about classroom teachers as a target audience:

- What types of questions do teachers seek to answer when they visit the WWC?
- Are the report ratings easy to understand to teachers?
- (If difficult to understand) What changes can you suggest to report ratings to make them more understandable to teachers?

b) Reflecting now on administrators as a target audience:

- What types of questions do teachers seek to answer when they visit the WWC?
- Are the report ratings easy to understand to teachers?
- (If difficult to understand) What changes can you suggest to report ratings to make them more understandable to teachers?

c) Reflecting now on researchers a target audience:

- What information and/or materials are most useful?
- What additional materials could be useful?

d) Are there audiences that you feel the WWC (IES) may not be targeting? Which ones? What type of information would they want to access?

### 2) Resource 1: Video

a) Let's talk a bit about the video:

- Is the presentation of the information in the video clear? If not, in what ways could it be clearer or improved?
- Would you recommend a video like this to others? If so, whom? (That is, would you show it to your school leaders? District staff?) If not, why not?
- What changes, if any, would you make to how the information is presented?
- Is the length of the video appropriate?
- Do you have any suggestions for other video topics?

### 3) Resource 2: Practice Guide Summary

a) Let's spend a few minutes talking about the practice guide summary:

- Is the presentation of the information clear? If not, in what ways could it be clearer or improved?
- Is the length appropriate?
- Would you recommend a summary like this to others? If so, whom? (That is, would you show it to your school leaders? District staff?) If not, why not?
- Does it pique your interest in the full guide? Why or why not?
- What additional materials, if any, would make the practice guides more useful?

**4) Resource 3: Topical Blast**

a) Let's spend a few minutes discussing the topical information email campaign we just explored.

- Is the presentation of the information clear? If not, in what ways could it be clearer or improved?
- Is the length appropriate?
- Does it pique your interest to explore additional resources on the website? Why or why not?
- What changes, if any, would you make to how the information is presented?
- Do you have any suggestions for other email campaign topics?

**5) Dissemination**

a) The WWC currently engages in quite a few dissemination strategies. In addition to the email campaigns we mentioned, the WWC hosts webinars and uses Facebook and Twitter to alert users to new content.

- Do you have any suggestions for how the WWC can get the word out abouts its resources?
- Are there audiences that you believe the WWC might not be targeting? Which ones? What type of information would they want to access, and how would they want to access it?

b) Thinking about all of the resources the WWC produces, including the resources we just viewed and the other resources we shared with you via email in advance of this discussion, do you have any suggestions for other resources you think the WWC should produce?

- If so, what do you suggest?
- What formats (for example, print, video) would be useful for these resources? Why?

**6) Wrap-Up**

a) As we close, I just wanted to remind you that WWC feedback effort is working to evaluate ways in which the WWC and its associated products could be improved. With this in mind, is there anything that I haven't asked you about regarding the What Works Clearinghouse and your work with the REL that you'd like to comment on? Is there anything else we should be aware of?

## Virtual Focus Group Protocol

### 1) Participant Background Knowledge of the WWC

a) Have you ever visited the WWC website?
b) Why do you visit the WWC website?
c) When you visit, is it easy to find what you're looking for?
d) How could the website be improved?
e) Do you use what you find on the website to inform decision making? If yes, what decisions? (for example, purchase a curriculum, change classroom practice)

### 2) Product 1: Video

a) Did you watch the video?
b) What did you like about the video?
c) What do you wish there had been more detail about?
d) What would you change to improve the video?
e) Who will you share this practice guide with? If you don't think you will share it with anyone, why not?
f) What do you think about the length of the practice guide summary? Was it too long, too short?
g) Does it pique your interest in the full guide? Why or why not?
h) What additional materials, if any, would make the practice guides more useful?

### 3) Product 2: Practice Guide Summary

a) What do you like about the practice guide summary?
b) What do you wish there had been more detail about?
c) What would you change to improve this approach to summarizing the information in the practice guide?
d) Who will you share this video with? If you don't think you will share it with anyone, why not?
e) What do you think about the length of the video? Was it too long, too short?
f) Do you have any suggestions for other video topics?

### 4) Product 3: Special Features

a) What did you like about the webpage?
b) What do you wish there had been more detail about?
c) What would you change to improve this webpage?
d) Who will you share this webpage with? If you don't think you will share it with anyone, why not?

e) Does it pique your interest in reviewing some of the other products on the website? Why or why not?
f) What other topics would most interest you?

**5) Dissemination**

a) Do you seek out research-based information to education? If so, what are your primary methods for getting this information? If not, why not?
b) How should the WWC get the word out about its products, all of which are free?
c) Are there other audiences you believe the WWC should target? If so, who? What type of information would they want?
d) What else could the WWC provide to help you with your work? What formats (for example, print, video) would be useful for these products and resources?

**6) Conclusion**

a) Do you have any additional comments or suggestions for the WWC?
b) Thank you for taking the time to participate in our virtual focus group. Please enter the email address where you'd like to receive your Amazon gift card.

# REFERENCES


Agans, J. P., Burrow, A. L., Kim, E. S., Garbo, C., Schroeder, M., Graf, S., & Davis, T. (2020). "You're going to burn some bridges if you come at it the wrong way": Reflecting on the realities of research-practice partnerships. *Community Development*, *51*(1), 36-52.

American Evaluation Association. (2018). *Guiding principles for evaluators.* Retrieved from https://www.eval.org/p/cm/ld/fid=51

Anwaruddin, S. M. (2015). Teachers' engagement with educational research: Toward a conceptual framework for locally-based interpretive communities. *Education Policy Analysis Archives*, *23*(40), 1-22.

Arbour, G. (2020). Teaching programme evaluation: A problem of knowledge. *Evaluation and Program Planning*. Advance online publication.

Archibald, T., & Buckley, J. (2018, July). *Evaluative thinking: Principals and practices.* Presented at the Summer Institute of the American Evaluation Association (AEA), Atlanta, GA.

Archibald, T., Sharrock, G., Buckley, J., & Cook, N. (2016). Assumptions, conjectures, and other miracles: The application of evaluative thinking to theory of change models in community development. *Evaluation and Program Planning*, *59*, 119-127.

Baldwin, S.E., Muller, P.A., Akey, T.M., McManus, J., Phillips, M., Plucker, J., and Sharp, S. (2008). *An evaluation of the effectiveness of the Institute of Education Sciences in carrying out its priorities and mission: Final report*. Institute of Education Sciences. U.S. Department of Education. Washington, DC.

Ball, A. F. (2012). To know is not enough: Knowledge, power, and the zone of generativity. *Educational Researcher*, *41*(8), 283-293.

Barton, E. A., & Tindle, K. (2019). *Educator voices on educational research*. Retrieved from EdTech Evidence Exchange website: https://drive.google.com/file/d/1Tsq88yyunLvtpktzlmM7x3pwamGHOvOw/view

Behrstock-Sherratt, E., Drill, K., & Miller, S. (2011). Is the supply in demand? Exploring how, when, and why teachers use research. Retrieved from the American Institutes of Research website: https://www.air.org/resource/supply-demand-exploring-how-when-and-why-teachers-use-research

Berliner, D. C. (2002). Educational research: The hardest science of all. *Educational Researcher*, *31*(8), 18–20.

Better Evaluation. (2020). *Contribution analysis.* Retrieved from https://www.betterevaluation.org/en/plan/approach/contribution_analysis

Beveridge, L., Mockler, N., & Gore, J. (2018). An Australian view of the academic partner role in schools. *Educational Action Research*, *26*(1), 25-41.

Bhatti, Y., Dahlgaard, J. O., Hansen, J. H., & Hansen, K. M. (2015). Getting out the vote with evaluative thinking. *American Journal of Evaluation*, *36*(3), 389-400.

Biddle, B. J., & Saha, L. J. (2002). *The untested accusation: Principals, research knowledge, and policy making in schools*. Lanham, MD: R&L Education.

Biddle, B. J., & Saha, L. J. (2006). How principals use research. *Educational Leadership*, *63*(6), 72-77.

Biesta, G. J. (2007). Why "what works" won't work: Evidence-based practice and the democratic deficit in educational research. *Educational Theory*, *57*(1), 1–22.

Biesta, G. J. (2010). Why ‘what works’ still won’t work: From evidence-based education to value-based education. *Studies in Philosophy and Education*, *29*(5), 491-503.

Boardman, A. G., Argüelles, M. E., Vaughn, S., Hughes, M. T., & Klingner, J. (2005). Special education teachers' views of research-based practices. *The Journal of Special Education*, *39*(3), 168-180.

Boaz, A., Ashby, D., & Young, K. (2002). *Systematic reviews: What have they got to offer evidence based policy and practice?* [White paper]. ESRC UK Centre for Evidence Based Policy and Practice. Retrieved from https://www.voced.edu.au/content/ngv%3A26838

Bogenschneider, K., & Corbett, T. J. (2010). *Evidence-based policymaking: Insights from policy-minded researchers and research-minded policymakers*. London, United Kingdom: Routledge.

Bogenschneider, K., Corbett, T. J., & Parrott, E. (2019). Realizing the promise of research in policymaking: Theoretical guidance grounded in policymaker perspectives. *Journal of Family Theory & Review*, *11*(1), 127-147.

Boulay, B., Goodson, B., Olsen, R., McCormick, R., Darrow, C., Frye, M.,…Sarna, M. (2018). *The Investing in Innovation Fund: Summary of 67 evaluations: Final Report* (NCEE 2018-4013). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Borg, S. (2008). English language teachers’ beliefs about research: Perspectives from the Netherlands. *Levende Talen Tijdschrift*, *9*(3), 3-13.

Borg, S. (2009). English language teachers' conceptions of research. *Applied Linguistics*, *30*(3), 358–388.

Borg, S., & Alshumaimeri, Y. (2012). University teacher educators' research engagement: Perspectives from Saudi Arabia. *Teaching and Teacher Education*, *28*(3), 347-356.

Bracht, G. H., & Glass, G. V. (1968). The external validity of experiments. *American Educational Research Journal*, *5*(4), 437-474.

Bråten, I., & Ferguson, L. E. (2015). Beliefs about sources of knowledge predict motivation for learning in teacher education. *Teaching and Teacher Education*, *50*, 13-23.

Breuer, E., Lee, L., De Silva, M., & Lund, C. (2016). Using theory of change to design and evaluate public health interventions: A systematic review. *Implementation Science*, *11*(1), 63-80.

Buchmann, M. (1984). The use of research knowledge in teacher education and teaching. *American Journal of Education*, *92*(4), 421-439.

Buckley, J., Archibald, T., Hargraves, M., & Trochim, W. M. (2015). Defining and teaching evaluative thinking: Insights from research on critical thinking. *American Journal of Evaluation*, *36*(3), 375-388.

Budhwani, S., & McDavid, J. C. (2017). Contribution analysis: Theoretical and practical challenges and prospects for evaluators. *Canadian Journal of Program Evaluation*, *32*(1), 1-24.

Burkhardt, H., & Schoenfeld, A. H. (2003). Improving educational research: Toward a more useful, more influential, and better-funded enterprise. *Educational Researcher*, *32*(9), 3-14.

Burn, K., Conway, R., Edwards, A., & Harries, E. (2020). The role of school-based research champions in a school–university partnership. *British Educational Research Journal*. Advance online publication.

Cain, T. (2015). Teachers' engagement with research texts: Beyond instrumental, conceptual or strategic use. *Journal of Education for Teaching*, *41*(5), 478–492.

Cain, T. (2016). Research utilisation and the struggle for the teacher's soul: A narrative review. *European Journal of Teacher Education*, *39*(5), 616-629.

Cain, T. (2017). Denial, opposition, rejection or dissent: Why do teachers contest research evidence? *Research Papers in Education*, *32*(5), 611–625.

Cantalini-Williams, M. T., Curtis, D., Eden-DeGasperis, K., Esposto, L., Guibert, J., Papp, H., & Roque, C. (2015). Exploring the benefits of a collaborative inquiry team in education (CITE) initiative to develop a research community and enhance student engagement. *Brock Education: A Journal of Educational Research and Practice*, *25*(1), 57-72.

Caplan, N. (1979). The two-communities theory and knowledge utilization. *American Behavioral Scientist*, *22*(3), 459-470.

Caplan, N., Morrison, A. & Stambaugh, R. J. (1975). *The use of social science knowledge in policy decisions at the national level: A report to respondents.* Ann Arbor, MI: Center for Research on Utilization of Scientific Knowledge, University of Michigan.

Center for Data-Driven Reform in Education. (n.d.). *Best Evidence Encyclopedia*. Retrieved from http://www.bestevidence.org/

Center for Research Use in Education (CRUE) (n.d.). Our guiding framework. Retrieved from https://www.research4schools.org/research-framework/

Chen, H. T., & Rossi, P. H. (1983). Evaluating with sense: The theory-driven approach. *Evaluation Review*, *7*(3), 283-302.

Chen, H. T., & Rossi, P. H. (1989). Issues in the theory-driven perspective. *Evaluation and Program Planning*, *12*(4), 299-306.

Clark, D. L., & Carriker, W. R. (1961). Educational research and the Cooperative Research Program. *The Phi Delta Kappan*, *42*(6), 226-230.

Coburn, C. E., Honig, M. I., & Stein, M. K. (2009). What's the evidence on district's use of evidence? In J. Bransford,D. J. Stipek, N. J. Vye, L. Gomez, & D. Lam (Eds.) *Educational improvement: What makes it happen and why?* (pp. 67–86). Cambridge, MA: Harvard Educational Press.

Coburn, C. E., & Penuel, W. R. (2016). Research-practice partnerships in education: Outcomes, dynamics, and open questions. *Educational Research,* 45(1), 48-54.

Coburn, C. E., & Talbert, J. E. (2006). Conceptions of evidence use in school districts: Mapping the terrain. *American Journal of Education*, *112*(4), 469-495.

Cook, B. G., Smith, G. J., & Tankersley, M. (2012). Evidence-based practices in education. In K. R. Harris, S. Graham, and T. Urdan (Eds.) *APA Educational Psychology Handbook: Vol. 1. Theories, Constructs, and Critical Issues* (pp. 493-526), Washington, DC: American Psychological Association.

Cooper, A. M. (2012). *Knowledge mobilization intermediaries in education: A cross-case analysis of 44 Canadian organizations* (Doctoral dissertation). Retrieved from ProQuest Dissertations Publishing. (Order No. NR97561)

Cooper, A. M. (2013). Research mediation in education: A typology of research brokering organizations that exist across Canada. *Alberta Journal of Educational Research*, *59*(2), 181-207.

Cooper, A., Edelstein, H., Levin, B., & Leung, J. (2010). *Use of web-based research materials in education: Is uptake occurring?.* Paper presented at the Canadian Society for the Study of Education (CSSE) Conference, Montreal, Canada.

Cooper, A., Levin, B., & Campbell, C. (2009). The growing (but still limited) importance of evidence in education policy and practice. *Journal of Educational Change*, *10*(2-3), 159-171.

The Cooperative Research Act, 68 Stat. § 533 (1954).

Council for the Accreditation of Educator Preparedness (CAEP). (n.d.). *CAEP standards*. Retrieved from http://caepnet.org/standards/introduction

Cousins, J. B., & Earl, L. M. (1992). The case for participatory evaluation. *Educational Evaluation and Policy Analysis*, *14*(4), 397-418.

Cousins, J. B., & Simon, M. (1996). The nature and impact of policy-induced partnerships between research and practice communities. *Educational Evaluation and Policy Analysis*, *18*(3), 199-218.

Cousins, J. B., & Walker, C. A. (2000). Predictors of educators' valuing of systematic inquiry in schools. *Canadian Journal of Program Evaluation*, *25*, 25-52.

Dagenais, C., Lysenko, L., Abrami, P. C., Bernard, R. M., Ramde, J., & Janosz, M. (2012). Use of research-based information by school practitioners and determinants of use: A review of empirical research. *Evidence & Policy: A Journal of Research, Debate and Practice*, *8*(3), 285–309.

Delahais, T., & Toulemonde, J. (2012). Applying contribution analysis: Lessons from five years of practice. *Evaluation*, *18*(3), 281-293.

Dewey, J. (1929). Sources of a Science of Education. In Boydston, J. A. and Hickman, L. A. (Eds.) *The Later Works of John Dewey, 1925-1953. Volume 5: 1929-1930, Essays, The Sources of a Science Education, Individualism, Old and New, and Construction and Criticism* (pp. 1-40). Charlottesville, VA: InteLex Corp., 1996.

Dirkx, J. M. (2006). Studying the complicated matter of what works: Evidence-based research and the problem of practice. *Adult Education Quarterly*, *56*(4), 273–290.

Dresner, M., & Worley, E. (2006). Teacher research experiences, partnerships with scientists, and teacher networks sustaining factors from professional development. *Journal of Science Teacher Education*, *17*(1), 1-14.

Drill, K., Miller, S., & Behrstock-Sherratt, E. (2013). Teachers' perspectives on educational research. *Brock Education: A Journal of Educational Research and Practice*, *23*(1), 3-17.

Dunn, W. N. (1980). The two-communities metaphor and models of knowledge use: An exploratory case survey. *Knowledge*, *1*(4), 515-536.

Dunn, W. N., & Swierczek, F. W. (1977). Planned organizational change: Toward grounded theory. *The Journal of Applied Behavioral Science*, *13*(2), 135-157.

Dybdal, L., Nielsen, S. B., & Lemire, S. (2011). Contribution analysis applied: Reflections on scope and methodology. *The Canadian Journal of Program Evaluation*, *25*(2), 29-57.

Ebbeler, J., Poortman, C. L., Schildkamp, K., & Pieters, J. M. (2017). The effects of a data use intervention on educators' satisfaction and data literacy. *Educational Assessment, Evaluation and Accountability*, *29*(1), 83-105.

*Education research by the Federal Government: Hearing before the Senate Health, Education, Labor, and Pensions and House Education and Workforce Committees.* 106th Cong. (1999). Federal News Service, Inc. Lexis-Nexis. [June 21, 1999] (William Goodling comment)

Ehren, M. C., Leeuw, F. L., & Scheerens, J. (2005). On the impact of the Dutch Educational Supervision Act: analyzing assumptions concerning the inspection of primary education. *American Journal of Evaluation*, *26*(1), 60-76.

The Elementary and Secondary Education Act, Statute 79 Stat. § 27 (1965).

Everton, T., Galton, M., & Pell, T. (2002). Educational research and the teacher. *Research Papers in Education Policy and Practice*, *17*(4), 373-401.

Evidence for Policy and Practice Information and Coordinating Centre (EPPI Center). (n.d.) *User-driven evidence-informed policy and practice*. Retrieved from http://eppi.ioe.ac.uk/cms/Default.aspx?tabid=65

*Expanding opportunity by funding what works: Using evidence to help low-income individuals and families get ahead. Hearing before the Committee on Ways and Means, U.S. House of Representatives*, 114th Congress (2015) (testimony of Grover J. Whitehurst).

Farley-Ripple, E. N. (2012). Research use in school district central office decision making: A case study. *Educational Management Administration & Leadership*, *40*(6), 786-806.

Farley-Ripple, E., May, H., Karpyn, A., Tilley, K., & McDonough, K. (2018). Rethinking connections between research and practice in education: A conceptual framework. *Educational Researcher*, *47*(4), 235-245.

Farrell, C. C., Davidson, K. L., Repko-Erwin, M. E., Penuel, W. R., Quantz, M., Wong, H., Riedy, R., & Brink, Z. (2018). *A descriptive study of the IES Researcher–Practitioner Partnerships in Education Research Program: Final report* (Technical Report No. 3). Boulder, CO: National Center for Research in Policy and Practice.

Finn, C. E. (2008). *Troublemaker: A personal history of school reform since Sputnik*. Princeton, NJ: Princeton University Press.

Finn, C. E., & Tomlinson, T. M. (1987). Response to Glass on what works. *Educational Researcher*, *16*(7), 24-26.

Finnigan, K. S., Daly, A. J., & Che, J. (2013). Systemwide reform in districts under pressure: The role of social networks in defining, acquiring, using, and diffusing research evidence. *Journal of Educational Administration*, *51*(4), 476-497.

Fox, J., & Bouchet-Valat, M. (2020). Rcmdr: R Commander. (Version 2.7-1) [Computer software]. Retrieved from https://socialsciences.mcmaster.ca/jfox/Misc/Rcmdr/.

Freire, P. (1968). *Pedagogy of the oppressed*. New York, NY: Seabury Press.

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

Glass, G. (1987). What works: Politics and research. *Educational Researcher*, *26*(3), 5-10.

Gorard, S., See, B. H., & Siddiqui, N. (2020). What is the evidence on the best way to get evidence into use in education?. *Review of Education*, *8*(2), p. 570-610.

Gould, R. V., & Fernandez, R. M. (1989). Structures of mediation: A formal approach to brokerage in transaction networks. *Sociological Methodology*, *19*, 89-126.

Grahe, J. E., Brandt, M. J., Wagge, J. R., Legate, N., Wiggins, B. J., Christopherson, C. D., … Andreychik, M. (2020). Collaborative Replications and Education Project (CREP). https://doi.org/10.17605/OSF.IO/WFC6U

Greene, J. C. (1998). Balancing philosophy and practicality in qualitative evaluation. In R. Davis (Ed.) *Proceedings of the Stake Symposium on Educational Evaluation* (pp. 35-49). Champaign/Urbana, IL: University of Illinois.

Groothuijsen, S. E. A., Bronkhorst, L. H., Prins, G. T., & Kuiper, W. (2019). Teacher-researchers' quality concerns for practice-oriented educational research. *Research Papers in Education*, 1–22.

Guba, E. G., & Lincoln, Y. S. (1989). *Fourth generation evaluation*. Thousand Oaks, CA: Sage Publications.

Guthrie, J. (1989). *Regional Educational Laboratories: History and prospect.* Washington, DC: Office of Educational Research and Improvement, U.S. Department of Education.

Hammersley, M. (1993). On the teacher as researcher. *Educational Action Research*, *1*(3), 425–445.

Hammersley, M. (2001a). On 'systematic' reviews of research literatures: A 'narrative' response to Evans & Benefield. *British Educational Research Journal*, *27*(5), 543-554.

Hammersley, M. (2001b). *Some questions about evidence-based practice in education*. Paper presented at the Annual Conference of the British Educational Research Association (BERA), Leeds, United Kingdom.

Hammersley, M. (2005). The myth of research-based practice: The critical case of educational inquiry. *International Journal of Social Research Methodology*, *8*(4), 317–330.

Hargreaves, D. (1996). *Teaching as a research-based profession: Possibilities and prospects.* Paper presented at The Teacher Training Agency Annual Lecture.

Harkin, J. (2005). Fragments stored against my ruin: The place of educational theory in the professional development of teachers in further education. *Journal of Vocational Education & Training*, *57*(2), 165–179.

Hayes, N. (1997). *Doing qualitative analysis in psychology.* East Sussex, England: Psychology Press.

Hemsley-Brown, J., & Sharp, C. (2003). The use of research to improve professional practice: A systematic review of the literature. *Oxford Review of Education*, *29*(4), 449–471.

Henderson, J., & Corry, M. (2020). Data literacy training and use for educational professionals. *Journal of Research in Innovative Teaching & Learning*. Advance online publication.

House, E., & Howe, K. R. (1999). *Values in evaluation and social research*. Thousand Oaks, CA: Sage Publications.

Huberman, M. (1990). Linkage between researchers and practitioners: A qualitative study. *American Educational Research Journal*, *27*(2), 363-391.

Hsieh, H. F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, *15*(9), 1277-1288.

Institute of Education Sciences (IES). (n.d.-a). *About IES: Connecting research, policy, and practice*. Retrieved from https://ies.ed.gov/aboutus/

Institute of Education Sciences (IES). (n.d.-b). *The Regional Educational Laboratory program: About Us*. Retrieved from https://ies.ed.gov/ncee/edlabs/about/

Institute of Education Sciences. (n.d.-c). *Research-practitioner partnerships in education research.* Retrieved from https://ies.ed.gov/ncer/projects/program.asp?ProgID=81

Institute of Education Sciences. (2013). *Institute of Education Sciences director's biennial report to congress: Fiscal years 2011 and 2012.* Retrieved from https://ies.ed.gov/pdf/2014001.pdf

Institute of Education Sciences. (2017). *Institute of Education Sciences biennial report to congress 2013-2014 and 2015-2016.* Retrieved from https://ies.ed.gov/pdf/IESBR2013_2016.pdf

Institute of Education Sciences. (2018a). *Institute of Education Sciences director's biennial report to congress: Fiscal years 2017 and 2018.* Retrieved from https://ies.ed.gov/pdf/IESBR2017_2018.pdf

Institute of Education Sciences. (2018b). *Understanding the Department of Education's evidence definitions*. Retrieved from https://ies.ed.gov/ncee/wwc/multimedia/42

Institute of Education Sciences (Producer). (2018c). *What we do: Learn about the What Works Clearinghouse.* Retrieved from https://www.youtube.com/watch?v=k7MHz6swwi4

Institute of Education Sciences. (2020). *Funding opportunities* [Data file]. Retrieved from https://ies.ed.gov/funding/grantsearch/index.asp?mode=2&sort=1&order=1&all=1&search=ProgramName&slctProgram=81

Ion, G., & Iucu, R. (2014). Professionals' perceptions about the use of research in educational practice. *European Journal of Higher Education*, *4*(4), 334–347.

Jacob, S., & Boisvert, Y. (2010). To be or not to be a profession: Pros, cons and challenges for evaluation. *Evaluation*, *16*(4), 349-369.

The Joint Committee on Standards for Educational Evaluation. (1994). *Program evaluation standards.* Thousand Oaks, CA: Sage Publishing.

The Joint Committee on Standards for Education Evaluation. (2010). *Program evaluation standards*. Retrieved from https://evaluationstandards.org/program/

Joram, E. (2007). Clashing epistemologies: Aspiring teachers', practicing teachers', and professors' beliefs about knowledge and research in education. *Teaching and Teacher Education*, *23*(2), 123–135.

Joram, E., Gabriele, A. J., & Walton, K. (2020). What influences teachers' "buy-in" of research? Teachers' beliefs about the applicability of educational research to their practice. *Teaching and Teacher Education*, *88*, 1-11.

Joyce, K. E., & Cartwright, N. (2019). Bridging the gap between research and practice: Predicting what will work locally. *American Educational Research Journal*, *57*(3), 1045-1082.

Judkins, M., Stacey, O., McCrone, T., & Inniss, M. (2014). *Teachers' use of research evidence: A case study of United Learning Schools*. Retrieved from the National Foundation for Educational Research (NFER) website: https://www.nfer.ac.uk/publications/IMUL01.

Kaestle, C. F. (1993). Research news and comment: The awful reputation of education research. *Educational Researcher*, *22*(1), 23-31.

Kearney, C. P. (1967). *The 1964 Presidential Task Force on Education and the Elementary and Secondary Education Act of 1965* [Doctoral dissertation]. ProQuest Dissertations Publishing. (T-13704)

Kennedy, M. M. (1999). A test of some common contentions about educational research. *American Educational Research Journal, 36*(3), 511-541.

Kippers, W. B., Poortman, C. L., Schildkamp, K., & Visscher, A. J. (2018). Data literacy: What do educators learn and struggle with during a data use intervention?. *Studies in Educational Evaluation*, *56*, 21-31.

Lather, P. (2004). This is your father's paradigm: Government intrusion and the case of qualitative research in education. *Qualitative Inquiry*, *10*(1), 15–34.

Lavis, J. N., Robertson, D., Woodside, J. M., McLeod, C. B., & Abelson, J. (2003). How can research organizations more effectively transfer research knowledge to decision makers?. *The Milbank Quarterly*, *81*(2), 221-248.

Le Fevre, D. M. (2014). Barriers to implementing pedagogical change: The role of teachers' perceptions of risk. *Teaching and Teacher Education*, *38*, 56–64.

Leat, D., Reid, A., & Lofthouse, R. (2015). Teachers' experiences of engagement with and in educational research: what can be learned from teachers' views?. *Oxford Review of Education*, *41*(2), 270-286.

Leeuw, F. L. (1991). Policy theories, knowledge utilization, and evaluation. *Knowledge and Policy*, *4*(3), 73-91.

Leeuw, F. L. (2003). Reconstructing program theories: Methods available and problems to be solved. *American Journal of Evaluation*, *24*(1), 5-20.

Leeuw, F. L. (2012). Linking theory-based evaluation and contribution analysis: Three problems and a few solutions. *Evaluation*, *18*(3), 348-363.

Leko, M. M., Roberts, C., Peyton, D., & Pua, D. (2019). Selecting evidence-based practices: What works for me. *Intervention in School and Clinic*, *54*(5), 286–294.

Levin, B. (2011). Mobilising research knowledge in education. *London Review of Education*, *9*(1), 15-26.

Lindblom, C. E., & Cohen, D. K. (1979). *Usable knowledge: Social science and social problem solving*. New Haven, CT: Yale University Press.

Lipsey, M. W., & Cordray, D. S. (2000). Evaluation methods for social intervention. *Annual Review of Psychology*, *51*(1), 345-375.

Love, A. (2004). Implementation evaluation. In J. S. Wholey, H. P. Hatry, & K. E. Newcomer's (Eds.), *Handbook of practical program evaluation*, (pp. 63-97). Hoboken, NJ: Wiley.

Lykins, C. (2012). Why “what works” still doesn’t work: How to improve research syntheses at the What Works Clearinghouse. *Peabody Journal of Education*, *87*(4), 500–509.

Lysenko, L. V., Abrami, P. C., Bernard, R. M., & Dagenais, C. (2015). Research use in education: An online survey of school practitioners. *Brock Education: A Journal of Educational Research and Practice*, *25*(1), 35-54.

Lysenko, L. V., Abrami, P. C., Bernard, R. M., Dagenais, C., & Janosz, M. (2014). Educational research in educational practice: Predictors of use. *Canadian Journal of Education*, *37*(2). 1-26.

Mageau, T. (2004). Determining 'what works': An interview with Dr. Grover 'Russ' Whitehurst. *THE Journal (Technological Horizons in Education)*, *31*(6), 32-47.

Mandinach, E. B., & Gummer, E. S. (2013). A systemic view of implementing data literacy in educator preparation. *Educational Researcher*, *42*(1), 30-37.

Martinovic, D., Wiebe, N., Ratkovic, S., Willard-Holt, C., Spencer, T., & Cantalini-Williams, M. (2012). ‘Doing research was inspiring’: Building a research community with teachers. *Educational Action Research*, *20*(3), 385-406.

Mathematica Policy Research. (2016, March 16). *Feedback from virtual focus group participants on revised Find What Works tool* [Memorandum].

Mayne, J. (2008). *Contribution analysis: An approach to exploring cause and effect* (Institutional Learning and Change [ILAC] Briefs 16). Retrieved from https://ideas.repec.org/p/ags/ilacbr/52525.html

Mayne, J. (2011). Contribution analysis: Addressing cause and effect. In K. Forss, M. Marra & R. Schwartz (Eds.) *Evaluating the complex*, (pp. 53-96). Piscataway, NJ: Transaction Publishers.

Mayne, J. (2012a). Contribution analysis [Special issue]. *Evaluation*, *18*(3).

Mayne, J. (2012b). Contribution analysis: Coming of age?. *Evaluation*, *18*(3), 270-280.

Mayne, J. (2015). Useful theory of change models. *Canadian Journal of Program Evaluation*, *30*(2), 119-142.

Mayne, J. (2017). Theory of change analysis: Building robust theories of change. *Canadian Journal of Program Evaluation*, *32*(2), 155-173.

Mayne, J. (2019). Revisiting contribution analysis. *Canadian Journal of Program Evaluation*, *34*(2) 171-191.

McArthur, G. (2008). Does What Works Clearinghouse work? A brief review of Fast ForWord®. *Australasian Journal of Special Education*, *32*(1), 101-107.

Mertens, D. M. (2015). Philosophical assumptions and program evaluation. *Enrico Guglielminetti Luciana Regina*, *75*, 75-85.

Mertens, D. M. (2016). Assumptions at the philosophical and programmatic levels in evaluation. *Evaluation and Program Planning*, *59*, 102-108.

Michie, S., Atkins, L., & West, R. (2014). The behaviour change wheel. A guide to designing interventions. Gloucestershire, Great Britain: Silverback Publishing.

Midwest Regional Education Laboratory. (2019). *ESSA tiers of evidence: What you need to know*. Retrieved from https://ies.ed.gov/ncee/edlabs/regions/midwest/pdf/blogs/RELMW-ESSA-Tiers-Video-Handout-508.pdf

Morris, M. (2015). Research on evaluation ethics: Reflections and an agenda. *New Directions for Evaluation*, *148*, 31-42.

Morton, S. (2015). Progressing research impact assessment: A 'contributions' approach. *Research Evaluation*, *24*(4), 405-419.

National Center for Education Evaluation and Regional Assistance (NCEE) (n.d.). *About us.* Retrieved from https://ies.ed.gov/ncee/aboutus/

Neal, J. W., Neal, Z. P., Kornbluh, M., Mills, K. J., & Lawlor, J. A. (2015). Brokering the Research-Practice Gap: A typology. *American Journal of Community Psychology*, *56*(3–4), 422–435.

Neal, J. W., Neal, Z. P., Lawlor, J. A., Mills, K. J., & McAlindon, K. (2018). What makes research useful for public school educators? *Administration and Policy in Mental Health and Mental Health Services Research*, *45*(3), 432–446.

Neal, J. W., Neal, Z. P., Lawlor, J. A., Mills, K. J., & McAlindon, K. (2019). What types of brokerage bridge the research-practice gap? The case of public school educators. *Social Networks*, *59*, 41-49.

Nelson, A. A. (2019). *Do no harm? Critiquing the beneficence of the What Works Clearinghouse.* Presented at the Midwestern Educational Research Association's annual conference, Cincinnati, OH.

Newman, J., Cherney, A., & Head, B. W. (2016). Do policy makers use academic research? Reexamining the "two communities" theory of research utilization. *Public Administration Review*, *76*(1), 24-32.

Niaz, M. (2009). Qualitative methodology and its pitfalls in educational research. *Quality & Quantity*, *43*(4), 535–551.

Nkwake, A. M. (2013). *Working with assumptions in international development program evaluation*. New York, NY: Springer.

Nkwake, A. M., & Morrow, N. (2016). Working with assumptions [Special issue]. *Evaluation and Program Planning*, *59.*

No Child Left Behind, 115 Stat § 1425 (2001).

Nutley, S., Jung, T., & Walter, I. (2008). The many forms of research-informed practice: A framework for mapping diversity. *Cambridge Journal of Education*, *38*(1), 53-71.

Nutley, S., Walter, I., & Davies, H. T. (2009). Promoting evidence-based practice: Models and mechanisms from cross-sector review. *Research on Social Work Practice*, *19*(5), 552-559.

Oakley, A. (2002). Social science and evidence-based everything: The case of education. *Educational Review*, *54*(3), 277-286.

Oakley, A. (2003). Research evidence, knowledge management and educational practice: Early lessons from a systematic approach. *London Review of Education*, *1*(1), 21-33.

*Options for the Future of the Office of Education Research and Improvement: Hearing before the House Committee on Education and the Workforce, Subcommittee on Early Childhood Youth and Families,* (Serial 106-107), 106th Cong. (2000a)

*Options for the Future of the Office of Education Research and Improvement: Hearing before the House Committee on Education and the Workforce, Subcommittee on Early Childhood Youth and Families,* (Serial 106-107), 106th Cong. (2000b) (Testimony of G. Reid Lyon).

*Options for the Future of the Office of Education Research and Improvement: Hearing before the House Committee on Education and the Workforce, Subcommittee on Early Childhood Youth and Families,* (Serial 106-107), 106th Cong. (2000c) (Testimony of Robert Slavin).

Ostinelli, G. (2016). The many forms of research-informed practice: A framework for mapping diversity. *European Journal of Teacher Education*, *39*(5), 534-549.

Pareja Roblin, N., Tondeur, J., Voogt, J., Bruggeman, B., Mathieu, G., & van Braak, J. (2018). Practical considerations informing teachers' technology integration decisions: The case of tablet PCs. *Technology, Pedagogy and Education*, *27*(2), 165–181.

Patton, M. Q. (1996). A world larger than formative and summative. *Evaluation Practice*, *17*(2), 131-144.

Patton, M. Q. (1998). Discovering process use. *Evaluation*, *4*(2), 225-233.

Patton, M. Q. (2007). Process use as a usefulism. *New Directions for Evaluation*, *2007*(116), 99-112.

Patton, M. Q. (2012). A utilization-focused approach to contribution analysis. *Evaluation*, *18*(3), 364-377.

Patton, M. Q. (Ed.). (2017). Pedagogy of evaluation [Special issue]. *New Directions for Evaluation*, *155.*

Paulsell, D., Thomas, J., Monahan, S., & Seftor, N. S. (2017). A trusted source of information: How systematic reviews can support user decisions about adopting evidence-based programs. *Evaluation Review*, *41*(1), 50-77.

Penuel, W. R., Briggs, D. C., Davidson, K. L., Herlihy, C., Sherer, D., Hill, H. C., ... & Allen, A. R. (2017). How school and district leaders access, perceive, and use research. *AERA Open*, *3*(2), 1-17.

Penuel, W. R., Farrell, C. C., Allen, A. R., Toyama, Y., & Coburn, C. E. (2018). What research district leaders find useful. *Educational Policy*, *32*(4), 540-568.

Pogrow, S. (2017). The failure of the US education research establishment to identify effective practices: Beware effective practices policies. *Education Policy Analysis Archives*, *25*(5), 1-19.

Pondiscio, R. (2019). *If education advocacy were more like pharmaceutical ads*. Retrieved from https://www.educationnext.org/if-education-advocacy-more-like-pharmaceutical-ads/

R Core Team, (2013). R: A language and environment for statistical computing [Computer software].Vienna, Austria: R Foundation for Statistical Computing.

Ratecliffe, M., Bartholomew, H., Hames, V., Hind, A., Leach, J., Millar, R., & Osborne, J. (2005). Evidence-based practice in science education: The researcher–user interface. *Research Papers in Education*, *20*(2), 169–186.

Reagan, R. (1986). *Remarks on receiving the Department of Education report on improving education.* Retrieved from https://www.reaganlibrary.gov/archives/speech/remarks-receiving-department-education-report-improving-education

*The Reauthorization of the Office of Educational Research and Improvement: Hearing before the House Committee on Education and the Workforce, Subcommittee on Education Reform* (Serial 107-47), 107th Cong. (2002a)

*The Reauthorization of Educational Research and Improvement: Hearing before the House Committee on Education and the Workforce, Subcommittee on Education Reform* (Serial 107-47), 107th Cong. 31 (2002b) (Testimony of Douglas Christensen).

*The Reauthorization of the Office of Educational Research and Improvement: Hearing before the House Committee on Education and the Workforce, Subcommittee on Education Reform* (Serial 107-47), 107th Cong. (2002c) (Prepared Testimony of Grover Whitehurst)

*The Reauthorization of the Office of Educational Research and Improvement: Hearing before the House Committee on Education and the Workforce, Subcommittee on Education Reform* (Serial 107-47), 107th Cong. (2002d) (Prepared Testimony of Michael Castle)

*The Reauthorization of the Office of Education Research and Improvement: Hearing before the Senate Committee on Health, Education, Labor, and Pensions* (Serial 80-479), 107th Cong. 27 (2002a).

*The Reauthorization of the Office of Education Research and Improvement: Hearing before the Senate Committee on Health, Education, Labor, and Pensions* (Serial 80-479), 107th Cong. 18 (2002b) (Prepared Testimony of Grover Whitehurst).

Rich, A. (2005). *Think tanks, public policy, and the politics of expertise*. Cambridge, England: Cambridge University Press.

Riley, B. L., Kernoghan, A., Stockton, L., Montague, S., Yessis, J., & Willis, C. D. (2018). Using contribution analysis to evaluate the impacts of research on policy: Getting to 'good enough'. *Research Evaluation*, *27*(1), 16-27.

Rogers, E. M. (1988). The intellectual foundation and history of the agricultural extension model. *Knowledge*, *9*(4), 492-510.

Rogers, P. J. (2008). Using programme theory to evaluate complicated and complex aspects of interventions. *Evaluation*, *14*(1), 29-48.

Rogers, P. (2014). *Theory of change* (UNICEF Methodological Briefs: Impact Evaluation 2). Retrieved from UNICEF Office of Research website: https://www.unicef-irc.org/publications/747-theory-of-change-methodological-briefs-impact-evaluation-no-2.html

Rossi, P. H., & Wright, J. D. (1984). Evaluation research: An assessment. *Annual Review of Sociology*, *10*, 331-352.

Sahni, S. D., Polanin, J. R., Zhang, Q., Michaelson, L. E., Caverly, S., Polese, M. L., & Yang, J. (2021a). *A What Works Clearinghouse rapid evidence review of distance learning programs* (WWC 2021-005REV). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, What Works Clearinghouse. Retrieved from https://ies.ed.gov/ncee/wwc/Docs/ReferenceResources/Distance_Learning_RER_508c.pdf

Sampson, S., Bradley, K., & Nelson, A. (in press). Creating a structurally sound instrument by utilizing principles of measurement theory. In U. Luhanda & A. Gregg's (Eds.) *Basic Elements of Survey Research in Education*.

Schaik, P. van, Volman, M., Admiraal, W., & Schenke, W. (2018). Barriers and conditions for teachers' utilisation of academic knowledge. *International Journal of Educational Research*, *90*, 50–63.

Schlosser, R. W. (2006). The role of systematic reviews in evidence-based practice, research, and development. *Focus*, *15*, 1-4.

Schneider, M. (2018). Changes are coming to the WWC [Blog post]. Retrieved from https://ies.ed.gov/director/wwc2018.asp

Schneider, M. (2019). First year accomplishments [Blog post]. https://ies.ed.gov/director/remarks/4-23-2019.asp

Schneider, M. (2020). A New Year's update [Blog post]. https://ies.ed.gov/director/remarks/1-8-2020.asp

Schön, D. A. (1992). The theory of inquiry: Dewey's legacy to education. *Curriculum Inquiry*, *22*(2), 119-139.

Schön, D. A. (1995). Knowing-in-action: The new scholarship requires a new epistemology. *Change: The Magazine of Higher Learning*, *27*(6), 27-34.

Schoenfeld, A. H. (2006). What doesn't work: The challenge and failure of the What Works Clearinghouse to conduct meaningful reviews of studies of mathematics curricula. *Educational Researcher*, *35*(2), 13-21.

Scriven, M. (1969). An introduction to metaevaluation. *Educational Products Report*, *2*, 36-38.

Scriven, M. (1972). The exact role of value judgments in science. In K. F Schaffner & R. S. Cohen's (Eds.) *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association* (pp. 219-247). Dordrecht, Holland: D. Reidel Publishing.

Scriven, M. (1981). *The logic of evaluation*. Reyes, CA: Edgepress.

Scriven, M. (2009). Meta-evaluation revisited. *Journal of Multidisciplinary Evaluation*, *6*(11), 3-8.

Sheldon, J. (2016). *"What works" doesn't work: The problem with the call for evidence based practices in the classroom* [White paper]. Bad Ass Teacher Association**.** Retrieved from https://files.eric.ed.gov/fulltext/ED574817.pdf

Sherratt, E., & Miller, S. (2011). *Is the supply in demand? Exploring how, when, and why teachers use research*. Retrieved from the American Institutes for Research website: https://www.air.org/resource/supply-demand-exploring-how-when-and-why-teachers-use-research

Slavin, R. E. (1986). Best-evidence synthesis: An alternative to meta-analytic and traditional reviews. *Educational Researcher*, *15*(9), 5-11.

Slavin, R. E. (2002). Evidence-based education policies: Transforming educational practice and research. *Educational Researcher*, *31*(7), 15-21.

Slavin, R. E. (2004). Education research can and must address "what works" questions. *Educational Researcher*, *33*(1), 27-28.

Slavin, R. E. (2005). *Evidence-based reform: Advancing the education of students at risk* (Report for Renewing Our Schools, Securing Our Future: A National Task Force on Public Education). Retrieved from the Center for America Progress's website: https://www.americanprogress.org/wp-content/uploads/kf/Slavin%203%2017%20FINAL.pdf

Slavin, R. E. (2017). *The mystery of the Chinese dragon: Why isn't the WWC up to date.* Retrieved from https://www.huffpost.com/entry/the-mystery-of-the-chinese-dragon-why-isnt-the-wwc_b_5a1f11a3e4b039242f8c815

Slavin, R. E., & Madden, N. A. (2008). Understanding bias due to measures inherent to treatments in systematic reviews in education. Paper presented at the Annual Meeting of the Society for Research on Effective Education, Crystal City, VA.

Smith, L. S., & Wilkins, N. (2018). Mind the gap: Approaches to addressing the research-to-practice, practice-to-research chasm. *Journal of Public Health Management and Practice*, *24*, S6-S11.

Sparks, S. D. (2018, November 27). Teachers waat education research. The feds spend millions on it. So why can't it get to the classroom? *Education Week*. Retrieved from https://blogs.edweek.org/edweek/inside-school-research

Stake, R. E. (2005). Qualitative case studies. In N. K. Denzin & Y. S. Lincoln (Eds.), *The Sage Handbook of Qualitative Research* (p. 443–466). Thousand Oaks, CA: Sage Publications Ltd.

Stein, D., & Valters, C. (2012). *Understanding theory of change in international development* (The Justice and Security Research Programme Paper 1). Retrieved from http://www.theoryofchange.org/wp-content/uploads/toco_library/pdf/UNDERSTANDINGTHEORYOFChangeSteinValtersPN.pdf

Stenhouse, L. (1980). Curriculum research and the art of the teacher. *Curriculum, 1*(1), 40-44.

Stiles, L. J. (1962). The Cooperative Research Program: Contributions and next steps. *The Phi Delta Kappan*, *43*(6), 231-236.

Stockard, J. (2010). An analysis of the fidelity implementation policies of the What Works Clearinghouse. *Current Issues in Education*, *13*(4). 3-24.

Stockard, J., & Wood, T. W. (2017). The threshold and inclusive approaches to determining “best available evidence” An empirical analysis. *American Journal of Evaluation*, *38*(4), 471-492.

St. Pierre, E. A. (2006). Scientifically based research in education: Epistemology and ethics. *Adult Education Quarterly*, *56*(4), 239–266.

Stufflebeam, D. L. (1974). *Meta evaluation*. Occasional Paper Series (No. 4). Western Michigan University, Kalamazoo, MI.

Stufflebeam, D. L. (1978). Meta evaluation: an overview. *Evaluation & the Health Professions*, *1*(1), 17-43.

Stufflebeam, D. L. (2001). The metaevaluation imperative. *American Journal of Evaluation*, *22*(2), 183-209.

Suri, H. (2013). Epistemological pluralism in research synthesis methods. *International Journal of Qualitative Studies in Education*, *26*(7), 889–911.

Tkachenko, O., Hahn, H. J., & Peterson, S. L. (2017). Research–practice gap in applied fields: An integrative literature review. *Human Resource Development Review*, *16*(3), 235-262.

Tseng, V. (2012). The uses of research in policy and practice and commentaries. *Society for Research in Child Development’s Social Policy Report*, *26*(2), 1-24.

University College London (UCL). (n.d.). *Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI-Centre)***.** Retrieved from https://www.ucl.ac.uk/ioe/departments-and-centres/centres/evidence-policy-and-practice-information-and-co-ordinating-centre-eppi-centre

University of Kentucky Institutional Research, Analytics, and Decision Support. (n.d.). Degrees awarded time series [Interactive data visualization]. Retrieved from https://www.uky.edu/irads/deRgrees-credentials-awarded

U.S. Department of Education (1986). *What works: Research about teaching and learning*. Washington, D.C.: Office of Educational Research and Improvement, U.S. Department of Education.

U.S. National Commission on Excellence in Education. (1983). *A nation at risk: The imperative for education reform*. Washington, D.C.: The National Commission on Excellence in Education, U.S. Department of Education.

U.S. Government Accountability Agency (2010). *Improved dissemination and timely product release would enhance the usefulness of the What Works Clearinghouse*. Retrieved from https://www.gao.gov/products/GAO-10-644

U.S. Office of Education Act, 14 Stat § 434 (1867).

van Geel, M., Keuning, T., Visscher, A., & Fox, J. P. (2017). Changes in educators' data literacy during a data-based decision making intervention. *Teaching and Teacher Education*, *64*, 187-198.

Vanderlinde, R., & van Braak, J. (2010). The gap between educational research and practice: Views of teachers, school leaders, intermediaries and researchers. *British Educational Research Journal*, *36*(2), 299-316.

Viadero, D. (2001). Whitehurst aims to retool education research. *Education Week*, *21*(1), 38-40.

Vinovskis, M. A. (1998). *Changing federal strategies for supporting educational research, development and statistics*. Paper prepared for the National Educational

Research Policy and Priorities Board, U.S. Department of Education, Washington, DC.

Vo, A. D. T. (2013). *Toward a definition of evaluative thinking.* Available from ProQuest Dissertations & Theses Global. (Order No. 3562824).

Vo, A. D. T. & Archibald, T. (Eds.). (2018). Evaluative Thinking [Special issue]. *New Directions for Evaluation*, *158.*

Vo, A. T., & Archibald, T. (2018). New directions for evaluative thinking. *New Directions for Evaluation*, *158*, 139-147.

Vo, A. D. T., Schreiber, J. S., & Martin, A. (2018). Toward a conceptual understanding of evaluative thinking. *New Directions for Evaluation*, *158*, 29-47.

Ward, V., House, A., & Hamer, S. (2009). Knowledge brokering: The missing link in the evidence to action chain?. *Evidence & Policy: A Journal of Research, Debate and Practice*, *5*(3), 267-279.

Weiss, C. H. (1979). The many meanings of research utilization. *Public Administration Review*, *39*(5), 426-431.

Weiss, C. H. (1988). Evaluation for decisions. Is anybody there? Does anybody care?. *Evaluation Practice, 9*(1)*,* 5-19.

Weiss, C. H. (1998). Have we learned anything new about the use of evaluation?. *American journal of evaluation*, *19*(1), 21-33.

Weiss, C. H., Murphy-Graham, E., & Birkeland, S. (2005). An alternate route to policy influence: How evaluations affect DARE. *American Journal of Evaluation*, *26*(1), 12-30.

Weiss, C. H., Murphy-Graham, E., Petrosino, A., & Gandhi, A. G. (2008). The fairy godmother—and her warts: Making the dream of evidence-based policy come true. *American Journal of Evaluation*, *29*(1), 29-47.

What Works Clearinghouse (WWC). (n.d.-a). *Foundational skills to support reading for understanding in kindergarten through 3rd grade.* Retrieved from https://ies.ed.gov/ncee/wwc/PracticeGuide/21

What Works Clearinghouse (WWC). (n.d.-b). *Frequently asked questions.* Retrieved from https://ies.ed.gov/ncee/wwc/FAQ

What Works Clearinghouse (WWC). (n.d.-c). *Glossary.* Retrieved from https://ies.ed.gov/ncee/wwc/Glossary

What Works Clearinghouse (WWC). (n.d.-d) *Procedures and standards handbooks.* Retrieved from https://ies.ed.gov/ncee/wwc/Protocols#procedures

What Works Clearinghouse (WWC). (n.d.-e) *Review protocols.* Retrieved from https://ies.ed.gov/ncee/wwc/Protocols

What Works Clearinghouse (WWC). (n.d.-f). *Thanks for the feedback!*. Retrieved from https://ies.ed.gov/ncee/wwc/Feedback

What Works Clearinghouse (WWC). (n.d.-g). *What we do*. Retrieved from https://ies.ed.gov/ncee/wwc/WhatWeDo

What Works Clearinghouse (WWC). (n.d.-h). *WWC Group Design Standards Online Training.* Retrieved from https://ies.ed.gov/ncee/wwc/OnlineTraining

What Works Clearinghouse (WWC). (n.d.-i). *What Works Clearinghouse: Search Publications.* Retrieved from https://ies.ed.gov/ncee/wwc/Publication#/ContentTypeId:1

What Works Clearinghouse (WWC). (2008a) *Procedures and standards handbook: Version 1.* Retrieved from https://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc_version1_standards.pdf

What Works Clearinghouse (WWC). (2008b) *Procedures and standards handbook: Version 2.0.* Retrieved from https://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc_procedures_v2_standards_handbook.pdf

What Works Clearinghouse (WWC). (2009a). *IES practice guide: Using student achievement data to support instructional decision making.* Retrieved from https://ies.ed.gov/ncee/wwc/Docs/PracticeGuide/dddm_pg_092909.pdf

What Works Clearinghouse (WWC). (2009b). *Review protocol for data-driven decision making practice guide* Retrieved from https://ies.ed.gov/ncee/wwc/Docs/ReferenceResources/wwc_data_pg_protocol_v2.0.pdf

What Works Clearinghouse (WWC). (2011) *Procedures and standards handbook: Version 2.1.* Retrieved from https://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc_procedures_v2_1_standards_handbook.pdf

What Works Clearinghouse (WWC). (2014) *Procedures and standards handbook: Version 3.0.* Retrieved from https://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc_procedures_v3_0_standards_handbook.pdf

What Works Clearinghouse (WWC) (Producer). (2016a). *Finding evidence with the WWC: A webinar for educators.* Retrieved from https://ies.ed.gov/ncee/wwc/Multimedia/26

What Works Clearinghouse (WWC). (2016b). *The What Works Clearinghouse: 2016 in review.* Retrieved from https://ies.ed.gov/ncee/wwc/2016yearinreview

What Works Clearinghouse (WWC). (2017a). *Level of evidence in What Works Clearinghouse practice guides.* Retrieved from https://ies.ed.gov/ncee/wwc/Docs/Multimedia/wwc_pg_loe_022718.pdf

What Works Clearinghouse (WWC) (Producer). (2017b). *Practice guide level of evidence*. Retrieved from https://www.youtube.com/watch?v=4DxtKyj9XKs

What Works Clearinghouse (WWC). (2017c) *Procedures handbook: Version 4.0.* Retrieved from https://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc_procedures_handbook_v4.pdf

What Works Clearinghouse (WWC). (2017d). *The What Works Clearinghouse: Looking back on 2017.* Retrieved from https://ies.ed.gov/ncee/wwc/2017YearInReview

What Works Clearinghouse (WWC). (2018). *Year in review: 2018.* Retrieved from https://ies.ed.gov/ncee/wwc/2018yearinreview

What Works Clearinghouse (WWC). (2019). *Year in review 2019*. Retrieved from https://ies.ed.gov/ncee/wwc/Docs/multimedia/YearInReview2019.pdf

What Works Clearinghouse (WWC). (2020a) *Certified reviewers* [Data file]. Retrieved from https://ies.ed.gov/ncee/wwc/Reviewer

What Works Clearinghouse (WWC). (2020b) *Data from individual studies* [Data file and code book]. Retrieved from https://ies.ed.gov/ncee/wwc/StudyFindings

What Works Clearinghouse (WWC) (Producer). (2020c). *How states can use the WWC to inform evidence use under ESSA*. Retrieved from https://www.youtube.com/watch?v=hGCjAz5HqAI

What Works Clearinghouse (WWC). (2020d). *Procedures handbook: Version 4.1.* Retrieved from https://ies.ed.gov/ncee/wwc/Docs/referenceresources/

WWC-Procedures-Handbook-v4-1-508.pdf

What Works Clearinghouse (WWC). (2020e). *WWC evidence review protocol for interventions for English learners, Version 4.0.* Retrieved from https://ies.ed.gov/ncee/wwc/Docs/ReferenceResources/EL_Protocol_v_4_0_Feb2020.pdf

Whitehurst, G. J. (2003). The Institute of Education Sciences: New wine, New bottles. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, IL.

Whitehurst, G. J. (2004). Making education evidence-based: Premises, principles, pragmatics, and politics. Paper presented at the IPR Distinguished Public Policy Lecture Series at Northwestern University, Evanston, IL.

Williams, D., & Coles, L. (2003). *The use of research by teachers: Information literacy, access and attitudes* (Department of Information Management Report 14). Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.488.3873&rep=rep1&type=pdf

Williams, D., & Coles, L. (2007a). Evidence-based practice in teaching: An information perspective. *Journal of Documentation*, *63*(6), 812–835.

Williams, D., & Coles, L. (2007b). Teachers' approaches to finding and using research evidence: An information literacy perspective. *Educational Research*, *49*(2), 185–206.

Woodworth, K.R., David, J.L., Guha, R., Wang, H., & Lopez-Torkos, A. (2008). *San Francisco Bay Area KIPP schools: A study of early implementation and achievement. Final report.* Retrieved from the SRI International website:

https://www.sri.com/publication/san-francisco-bay-area-kipp-schools-a-study-of-early-implementation-and-achievement-final-report/

Yeager, D. S., & Walton, G. M. (2011). Social-psychological interventions in education: They're not magic. *Review of Educational Research*, *81*(2), 267-301.

Zhao, Y. (2017). What works may hurt: Side effects in education. *Journal of Educational Change*, *18*(1), 1–19.

Zhao, Y. (2018). *What works may hurt: Side effects in education.* New York, NY: Teachers College Press

Zoellner, B. P. (2010). *Prescription for the classroom? The cultural perspectives behind the scientifically based research reforms influencing the practices of the What Works Clearinghouse*. Available from ProQuest Dissertations & Theses Global database. (UMI No. 3437079)

VITA

**1) Education Institutions Attended / Degrees Awarded:**

**University of Kentucky –** Lexington, KY (2015 to 2021)
*Master of Arts:* Experimental Psychology (with an emphasis in Social/Developmental areas)
*Graduate Certificate:* Research Methods in Education (RMinE)
*PhD (in progress):* Education Sciences, with a specialization in Educational Policy Studies and Program Evaluation

**Pacific Lutheran University –** Tacoma, WA (2010 to 2014)
*Major:* Bachelor of Arts in Psychology
*Minors:* Religion and Publishing / Printing Arts

**University of St. Andrews –** St. Andrews, Scotland (studied abroad for the Spring 2013 term)

**2) Professional Positions Held**

| | |
|---|---|
| **2017-2020** | *Graduate Assistant, University of Kentucky Evaluation Center* |
| **2016-2019** | *Graduate Research Fellow, National Science Foundation* |
| **Spring 2016** | *Statistical Assistant for University of Kentucky Psychology Assistant* |

**3) Scholastic and Professional Honors**

| | |
|---|---|
| **Fall 2019** | *Department Travel Award* – Received a travel stipend to offset conference travel expenses ($1,000). |
| **Fall 2018** | *Department Travel Award* – Received a travel stipend to offset conference travel expenses ($1,000). |
| **Spring 2018** | *Department Travel Award* – Received a travel stipend to offset conference travel expenses ($1,000). |
| **Spring 2017** | *Department Travel Award* – Received a travel stipend to offset conference travel expenses ($1,000). |
| **Spring 2016** | *Department Travel Award* – Received a travel stipend to offset conference travel expenses ($1,000). |
| **Fall 2015** | *NSF Graduate Research Fellow* – Received a fellowship to |

fund my next three years of graduate training (~$102,000).

**Fall 2015** *Kentucky Opportunity Fellowship* **–** Qualified for a tuition waiver and year-long stipend (~$15,000) for past demonstrations of academic excellence.

**Spring 2014** *Pacific Lutheran University's "Psychology Research Award"* **–** Acknowledged for excellence in psychological research as demonstrated by my manuscript "Men as Students of a Feminist Pedagogy: A Social and Political Imperative."

**Spring 2014** *Pacific Lutheran University's "Service to Department Award"* **–** Acknowledged for creating a short, animated video entitled "Three Years in Three Minutes," which briefly detailed my experience conducting undergraduate research at Pacific Lutheran University.

**Summer 2013** *APA Summer Science Fellow* **–** Selected along with 11 other psychology students to undergo intensive research training, graduate school preparation, and study of careers in psychology with faculty at George Mason University.

**Fall 2010-2014** *International Honors Program* **–** Admitted (based on college entrance exam scores and high school GPA) and continued study in Pacific Lutheran University's Honors Program.

**Summer 2010** *President's Scholar* **–** Received competitive award based on scholastic achievement (guaranteed $72,000).

**4) Professional Publications**

Bradley, K. D., Sampson, S. O., & Nelson, A. A. (in press). Creating a structurally sound instrument by utilizing principles of measurement theory. In *Basic Elements of Survey Research in Education.*

Lyon, E., Freeman, R., Bathon, J., Fryar, A., McGlue, M., Erhardt, A., Rosen, A., Sampson, S., Nelson, A., & Parsons, J. (2020). Attitudinal impediments to geoscience major recruitment in among ninth graders at a STEAM high school. *Journal of Geoscience Education.*

Patterson, M. M., Bigler, R. S., Pahlke, E., Brown, C. S., Hayes, A. R., Ramirez, M. C., & Nelson, A. A. (2019). Toward a developmental science of politics: Core foundations, new perspectives, and illustrative empirical approaches. *Monographs of the Society for Research in Child Development*.

Nelson, A. A. & Brown, C. S. (2018). Too pretty for homework: How sexualized gender stereotypes impact girls' academic outcomes. *The Journal of Early*

*Adolescence.*

Nelson, A. A., Grahe, J. E., Ramseyer, F. (2016). Interacting in flow: An analysis of rapport-based coordination as optimal experience. *SAGE Open.*

Mandell, A., Becker, A., Nelson A. A., & Shaw, T. H. (2015). Neuroticism and vigilance revisited: A Transcranial Doppler investigation. *Consciousness and Cognition*.

Mandell, A., Becker, A., Nelson A. A., & Shaw, T. H. (2014). The effect of neuroticism on vigilance performance: A Transcranial Doppler investigation. *Proceedings of the Human Factors and Ergonomics Society,* USA, 58.

Nelson, A. A., Grahe, J. E., Ramseyer, F., & Serier, K. (2014). Psychological data from an exploration of the rapport / synchrony interplay using Motion Energy Analysis. *The Journal of Open Psychology Data.*

**5) Typed Name of Student**

Andrew A. Nelson