UNDER THE PRESSURE TO REPLICATE "EVIDENCE-BASED" INTERVENTION, TRIO PROGRAMS SHOULD BUILD THEIR LOGICS OF EVIDENCE USE *BY MIKA YAMASHITA*

Introduction

In 2014, when TRIO grantees saw the instructions for the grant application for the new grant cycle of Student Support Services (SSS), they saw a new requirement. Notice of Inviting Applications for Student Support Services (SSS) issued for the 2015 Competition included a Competitive Preference Points (CPP), which states the applicants can earn up to six points by proposing two interventions that are "supported by moderate evidence of effectiveness" (79. Fed. Reg. 2014, page 79575). To show that their proposed intervention is effective, applicants had to submit one supporting study that meets What Works Clearinghouse (WWC)'s moderate evidence standards. The similar requirement appeared in 2016 in the Talent Search grant competition. To earn six points, applicants had to demonstrate that each of proposed "strategy is based on research that meets the Moderate Evidence of Effectiveness standard" (80. Fed. Reg. 2015, page 79575) and "demonstrate how the proposed project activities align with the cited study with sufficient fidelity" (80. Fed.Reg. 2015, p. 79575). It appears the competitive preference points (CPP) resulted in enforcement because the cut-off score of SSS grant application was around 104, indicating that no new program that did not have any prior experience could receive funding without earning the competitive preference points.

As a researcher for the Pell Institute, a research arm of the Council for Opportunity in Education (COE), a membership organization of TRIO and other college access and success programs, I have heard concerns about TRIO program evidence standards raised by COE and Pell staff and TRIO program directors and staff. These concerns include the requirement to replicate interventions that were developed somewhere else could

Abstract

This essay discusses the origin of the competitive preference priorities (CPP) that appeared in Student Support Services (SSS) and Talent Search grant applications in 2014 and 2015. In these competitions, CPP encouraged programs to propose evidence-based interventions. I discuss the CPP as a series of policy instruments that the federal government implemented to promote evidence-based policymaking. The idea underlying CPP is that by encouraging grantees to implement evidence-based interventions, the grant programs are likely to deliver results. One important element of this assumption is that randomized controlled trials (RCTs) are the best methods to ascertain if an intervention is effective. After describing limitations of applying findings from RCTs to improve program practice, I argue that TRIO programs need to articulate their own specific logics of evidence use that reflect how TRIO professionals work in practice. This is because RCTs alone can provide very limited information about actual interventions for TRIO professionals to improve their own programs. Research questions that reflect how TRIO programs actually work are needed.

discourage innovation among programs and whether or not there are sufficient interventions that meet WWC standards that can be implemented by SSS programs. They also cited challenges in coordinating the evidence-based services that are required by legislation within the limited program budget. The overall sentiment expressed was that CPP is another hoop TRIO programs must jump through to receive federal funding.

Like many federal education programs, TRIO programs are under pressure to show evidence that they are effective. The programs have been regularly monitored, and grantees must meet performance objectives. Prior experience points, which programs earn by meeting a set of performance objectives created at the beginning of each grant cycle, make up a significant portion of the grant scoring. This means a program must meet a set of performance objectives in order to be funded again.

In this essay, based on my reading of publicly-available documents, I discuss the CPPs as a part of policy instruments the federal government has been using to promote evidence-based policymaking. The government views requiring evidence-based intervention as one of their effective grant-making strategies. I also argue that the CPP reflects the idiosyncrasy of government policy making, where one aspect of idea is taken up and other aspects are marginalized or lost. By describing limitations of the assumptions underlying the CPPs, I discuss why it is critical for TRIO programs to identify a research agenda that reflects the actual work of TRIO professionals. While TRIO programs may view CPP as another hoop to jump through, it is important to understand the big picture because the CPP carries certain expectations about how TRIO programs should improve. The primary audience of this essay is TRIO professionals.

Origins of Competitive Preference Points Attached to Evidence-Based Interventions

The Education Department General Administrative Regulations (EDGAR), which governs grant programs, provides a legal justification of inserting CPPs in the most recent SSS and Talent Search grant competitions. Recent revision of EDGAR in 2013 is in line with the government-wide initiative to infuse evidence use in government grant making (McNail, 2013). The Office of Management and Budget (OMB) has issued guidance documents to "better integrate evidence and rigorous evaluation in budget, management, operational and policy decisions" and it recommended agencies to adopt "more evidence based structures for grant programs" (OMB, no date). A document that provided guidance on Fiscal Year (FY) 2014 budget outlines "evidence-based grants" (OMB, May 18, 2012, page 2), and one of the approaches recommended by OMB is as follows:

> *agencies can provide points or significant competitive preference to*
> *programs that the agency determines are backed by strong evidence,*
> *and can build the evidence base by embedding evaluation into*
> *programs. (OMB, May 18, 2012, p. 3)*

As described in "Strategic Plan for Fiscal Year 2014-2018" and in the most recent "FY16-17 Agency Priority Goals," the Department of Education continues to use competitive preference points on discretional grant programs to "enable evidence-based decision making" (U.S.

Department of Education, 2014, 2016). The Department of Education estimates the percentage of new competitive funding that supported evidence-based projects (including strong, moderate and evidence of promise as defined by What Works Clearinghouse) in FY 15 was 16 percent, and it aims to increase it to 20 percent by September 30, 2017 (U.S. Department of Education, no date). TRIO-Talent Search is listed as one of the contributing programs, along with "various Higher Education Act Title III programs" and" Investing in Innovation Fund (i3)" for FY 16-17(U.S. Department of Education, 2016).

The Obama Administration used grant-making as a strategy to promote evidence use to drive outcomes (Berlin, 2016; Haskins & Margolis, 2015; Nussele & Orszag, 2015). Based on interviews with over 130 individuals who were involved in the Obama Administration, Haskins and Margolis (2015) describe how the administration moved from formula grant-making to competitive grant-making. They report "the administration wanted to take advantage of the pressure created by competition to spend the money on projects that meet strong evidence criteria" (pp.29-30) and the administration assumed that competition helps to ensure "that only the highest quality applicants win funding" (p.30). The Administration viewed it as "the smarter investment" to invest in programs that can "deliver results" (Orszag, 2009).

The Obama Administration introduced tiered evidence, where programs that have "strong evidence" received more funding than programs that have "some evidence" in its signature social programs, such as Home Visitation Programs at the Department of Health and Human Services and Investing in Innovation (i3) at the Department of Education. This approach was to ensure that funded programs would deliver results and to promote innovation, as OMB described below:

> *For these two very different subjects (referring to Home Visitation Program and Teen Pregnancy Prevention Program), we are using a similar, two-tiered approach. First we're providing more money to programs that generate results backed up by strong evidence. That's the top tier. Then, for an additional group of programs, with some supportive evidence but no as much, we've said: Let's try those too, but rigorously evaluate them and see whether they work. Over time, we hope that some of those programs will move into the top tier—but, if not, we'll redirect their funds to other, more promising efforts. This design differs from the typical approach. We haven't simply created a block grant and told states they can do whatever they want, nor have we dictated a particular program design and told everyone to follow it. Instead, we've said that we are flexible about the details of the program; we only insist that most of the money go toward the programs backed by the best available evidence, and the rest to programs that are promising and willing to test their mettle. (Orzag, June 8, 2009)*

Other grant-making strategies OMB has recommended include use of evidence in formula grants and Pay for Success (OMB, May 18, 2012). Pay for Success forces grantees, often intermediary organizations, to have significant evidence that proposed programs will deliver results, as program success determines if applicants receive payment from the government. As it is, the CPPs included in TRIO competitions are a part of the grant-making strategy that was intended to

promote evidence-based policymaking in the federal government. The underlying assumption is by investing in programs that work, the government and taxpayers will get better results.

Spread of Evidence-Based Programs/Interventions

While the competitive preference priorities (CPPs) attached to evidence-based interventions was new for TRIO programs, there were grant programs with similar requirements prior to the Obama Administration. For example, previously "evidence-based" appears in laws related to social welfare programs, such as immunization, drug prevention, and care for mothers with new born babies, where grantees were to implement "evidence-based" assessment and practices. In education, in the late 1990s, Safe and Drug Free programs issued Principles of Effectiveness, which states programs should be "guided by research and best practices." This guidance was later incorporated in the No Child Left Behind Act in 2001 (Weiss, Murphy-Graham & Birkeland, 2005). The Department of Education's Safe and Drug-Free School Program gathered a panel of experts to create an evidence-based programs list. The panel reviewed programs on four criteria: efficacy, quality, educational significance, and usefulness to others, and compiled a list of exemplary and promising programs (Department of Education, 2001).

Subsequently, the federal government began initiatives to identify effective interventions, which created inventories of programs that are evidence-based. What Works Clearinghouse (WWC), created in 2002, is one of them. Some others are:

- Evidence –Based Practice Centers (EPC) by Agency for Healthcare Research and Quality (AHRQ), created in 1997, focuses on evidence on relative benefits and risks of healthcare interventions.

- The Guide to Community Prevention Services by Health and Human Services (HHS), created in 1996, provides evidence-based recommendations and findings about public health interventions and policies to improve health and promote safety.

- National Registry of Evidence Based Programs and Practices (NREPP) by the Substance Abuse and Mental Health Services Administration (SAMHSA), created in 1997, provides for the scientific basis and practicality of interventions that prevent or treat mental health and substance abuse disorders. (Government Accountability Office, 2009)

These evidence review initiatives were to provide information to professionals, policymakers, and researchers about effective interventions by reviewing studies. Each initiative created its own evidence standards, reporting template, and review guidelines. According to a report published by the Government Accountability Office (GAO) in 2009, the review process was in general similar across the initiatives, but there were some variances. While randomized controlled trials (RCTs) and quasi-experimental studies are the most commonly reviewed study designs across the initiatives, Evidence-Based Practice Centers (AHRQ) and Guide to Community Preventive Services (CDC) also review observational studies, such as time series and case control studies. While the What Works Clearinghouse (WWC) did not assess implementation fidelity, four other initiatives explicitly assess intervention fidelity "through either describing in detail the intervention's components or measuring participants' level of exposure" (GAO, 2009, p. 15).

These resources, procedures, and experiences of identifying, reviewing, and rating studies and interventions since the late 90s contributed to the Obama Administration's evidence-based initiatives.

Spread of the Idea of Causality as Effectiveness

The underlying idea of attaching competitive preference points (CPP) is that the programs will produce better results by implementing evidence-based interventions. Evidence standards set by WWC determine if an intervention is evidence-based because the study design must meet WWC's design standards and also show positive results. According to the WWC, only RCTs potentially meet design standards without reservation and certain RCTs and quasi-experimental studies with comparison groups potentially meet WWC design standards with reservations (What Works Clearinghouse, 2014). According to the Notice of Inviting Applicants of the SSS grant, the study has to show "statistically significant favorable impact" (79. Fed. Reg. 2014, page 75724). This view of effectiveness as determined by causal relationship between an intervention and outcomes has been in place for over a decade in the Department of Education and the federal government at large.

The Institute of Education Sciences (IES) established What Works Clearinghouse (WWC) in 2002. The goal of WWC is:

> *to be a resource for informed educational decision making. To reach this goal, the WWC identifies studies that provide credible and reliable evidence of the effectiveness of a given practice, program or policy (referred to as "interventions"). (WWC, no date)*

WWC defines "effectiveness" as follows:

> *an intervention demonstrates effectiveness if the research has shown that it caused an improvement in outcomes. (WWC, no date)*

The Department of Education promoted this idea of effectiveness since early 2000. IES Director Grover (Ross) Whitehurst's presentation at the American Educational Research Association (AERA), after the passage of Education Science Reform Act of 2002, which created the Institute of Education Sciences (IES), states IES's position: "Randomized trials are the only sure method for determining the effectiveness of education programs and practices" (Institute of Education Sciences, no date).

A report by Congressional Research Service (CRS) published in 2006 illustrates the spread of the idea that RCTs is the best way to determine program effectiveness within the Department of Education and at the Office of Budget and Management (OMB). The Department of Education published "Notice of Final Priority" in November 2003 to clarify "scientifically based research" stated in the No Child Left Behind Act (NCLB) and to make it a priority to be used for all programs beyond those funded by NCLB, which became effective in February, 2005. The final notice states:

> *The definition of scientifically based research in section 9201(37) of NCLB includes other research designs in addition to the random*

assignment and quasi-experimental designs that are the subject of this priority. However, the Secretary considers random assignment and quasi-experimental designs to be the most rigorous methods to address the question of project effectiveness. While this action is of particular importance for programs authorized by NCLB, it is also an important tool for other programs and, for this reason, is being established for all Department programs. Establishing the priority on a Department-wide basis will permit any office to use the priority for a program for which it is appropriate. (70. Fed. Reg. January 25, 2005, p. 3586)

A similar message appeared in another publication by the Department of Education, "Identifying and Implementing Educational Practices Supported by Rigorous Evidence: A User Friendly Guide," which was to "provide educational practitioners with user friendly tools to distinguish practices supported by rigorous evidence from those that are not" (Brass, Nunez-Neto & Williams, March 7, 2006, page iii). CRS summarizes:

guidance asserted that evaluation methods other than RCT and certain quasi experiments (1) have no meaningful evidence' to contribute to establishing whether an intervention was 'effective' and (2) cannot be considered 'scientifically-rigorous evidence' or 'rigorous evidence' to support using an educational practice to 'improve educational and life outcomes for children. (Brass, Nunez-Neto & Williams, March 7, 2006, p. 25)

The Office of Management and Budget (OMB) also promoted this conceptualization of program effectiveness during Bush Administration with the technical assistance from the Coalition for Evidence-Based Policy, a private non-profit organization (Brass, Nunez-Neto & Williams, March 7, 2006, Coalition for Evidence Based Policy, no date). "What Constitutes Strong Evidence of a Program's Effectiveness" in 2004 is a guidance document to support agencies with reporting on The Program Assessment Rating Tool (PART), a set of questions asked to agencies to draw conclusions about program benefits and recommendations to improve programs. PART asked if an evaluation conducted for its program is sufficient in scope and quality and if the evaluation indicates the program is effective. The guidance document states:

The most significant aspect of program effectiveness is impact—the outcome of the program, which otherwise would not have occurred without the program intervention. While it is feasible to measure the impact of the program, RCTs are generally the highest quality, unbiased evaluation to demonstrate the actual impact of the program. (Office of Management and Budget, 2004)

Similarly, a guidance document for FY 2007 PART reporting describes:

most significant aspect of program effectiveness in impact—the outcome of the program, which otherwise would not have occurred without the program intervention. A number of evaluation methodologies are available to measure the effectiveness of programs. Some, such as randomized controlled trials, are particularly well suited

to measuring impacts. …. Quasi-experimental studies should be scrutinized given the increased possibility of an erroneous conclusion. (Office of Management and Budget, January 29, 2007, pp 30-31)

The promotion of these specific conceptualizations of "effectiveness" and "evidence" by the executive branch has drawn attention from legislative offices. Over a decade ago, the Congressional Research Service (CRS) produced a report that describes design features of RCTs and potential issues the Congress might encounter in the future (Brass, Nunez-Neto & Williams, March 7, 2006). GAO's report in 2009 on the Coalition for Evidence Based Policy's Top Tier Evidence Initiative, which identifies interventions that meet Coalition's top tier evidence standards, compared the Coalition's review process with the review process taken by six federally-funded evidence review initiatives, such as What Works Clearinghouse (WWC) and National Registry of Evidence Based Programs and Practices (GAO, 2009). These reports presented limitations of the RCTs' research design and for making policy decisions. For example, one of the conclusions by GAO report was:

> *Requiring evidence from randomized studies as sole proof of effectiveness will likely exclude many potentially effective and worthwhile practices. (GAO, 2009, cover page)*

Prior to these reports, the research community raised concerns about a specifying research design for the purpose of improving knowledge about education because scientific inquiry uses multiple types of research design. In 2001, during the time Education Science Reform Act was in the Congress, the National Research Council gathered a Committee on Scientific Principles for Education Research, which produced guiding principles for scientific inquiry. In the report, after acknowledging that a part of the Committee charge was to provide recommendations to improve education research, the report states:

> *attempting to boost the scientific basis of federally funded education research by mandating a list of "valid" scientific method is a problematic strategy. (National Research Council, 2002, p. 130)*

It appears these concerns were not reflected in the guidance and regulations developed by the federal government. It may be possible that policy makers i.e., government officials, legislative staff, and advocacy groups who were involved in policymaking, limited their frame of reference to what they saw as manageable (National Research Council, 2002). From reading Haskins and Magolis (2015)'s report on how Obama Administration's transition team, government officials, legislative staff, and advocacy groups crafted evidence-based initiatives and how grant competitions and selection criteria were developed, it seems the notion of effectiveness equals causal relationship had been institutionalized. Disagreements discussed in their report centered around whether evidence-based programs should be limited to programs supported by RCTs only or should be expanded to programs supported by quasi-experimental studies, as it determines which model programs would be replicated by grantees. It appears there was no question about the assumption that replicating evidence based intervention is most likely to deliver results.

Limitations of Evidence-Based Interventions

TRIO programs raised concerns when they learned about the competitive preference points (CPP). The concerns derive from the limitations of RCTs for informing practice, which I describe below.

Lack of External Validity. Even when RCTs are well-designed and well-implemented, they are limited in external validity. An intervention that showed positive results in one site does not guarantee the intervention will produce the same effect in other sites. The study tells the intervention worked in a certain context in which the experiment happened. Sometimes grant programs assign higher scores to interventions that are found to be effective in multiple sites. For example, the Corporation for National and Community Services' AmeriCorps grant competition provided more points to a proposal that proposed an intervention that had been tested nationally or in multi sites, while if applicant submitted a study that was conducted in one site, they received lower points (Corporation for National and Community Services, 2015). However, as of now, there are not many replication studies in education. The Department of Education asked SSS applicants to submit one study that includes "a sample that overlaps with the populations or settings" that meets the definition of moderate evidence of effectiveness; however, there was no further definition of what it means by "population" and "settings."

Another difficulty surrounding multiple replication studies seems to be that programs improve. The excerpt below, by a CEO of Building Educated Leaders for Life (BELL), a program that provides after school and summer learning experience for high risk students, describes these limitations of RCTs.

> *Even when we are successful generating good evidence, the story of what works is still incomplete. Even when we have well-conducted, highly rigorous studies, there are still limitations in what they tell us about the real nature of social problems, interventions, and changes in the lives of those we reach. As an example, even when randomized controlled trials tell us the story of an organization's work at one point in time, five years later most organizations' work has evolved. So at its best, good evidence is still incomplete."*

> *Obtaining evidence is risky. ... So you do a randomized controlled trial, and it says your program works. It's not the end of the story. It worked for some group of kids in some context at some point in time, and so we should celebrate that and make sure that's understood. And your randomized-controlled-trial study two years later might say, 'It doesn't work for kids or families.' And so there, too, we don't want to unfairly punish an organization without understanding that in context. So this risk, I think, is something we'll have to pay more attention to across sectors (Nussle & Orszag, 2015)*

Meta-analysis compiles results from multiple RCTs to generate an overall effect of an intervention implemented across different populations or settings. Systematic review could include other study designs and narrative synthesis to inform conditions that might have led to better results, such as minimum duration of intervention exposure and conditions. However, users of the review also need to determine the applicability of the results of these reviews (Brass, Nunez-Neto & Williams, March 7, 2006). For example, the guidance document by Cochrane Review, which conducts systematic reviews of evidence on healthcare treatments and products, describes the importance of readers to make informed decisions about applicability of findings of systematic review and meta- analysis as follows:

> *Another decision users must make is whether the patients before them are so different from those included in the studies that they cannot use the result of the systematic review and meta-analysis at all…. Authors can sometimes help clinical decision makers by identifying important variation where divergence might limit the applicability of results, including: biologic and cultural variation, variation in adherence to an intervention. In addressing these issues, authors cannot be aware of, or address, the myriad of differences in circumstances around the world. They can, however, address differences of known importance to many people, and importantly, they should avoid assuming that other people's circumstances are the same as their won in discussing the results and drawing conclusion. (Higgins & Green, 2011).*

Complexity of Intervention. Another limitation is that social service and education interventions studied are actually complex with high "causal density" (Manzi, 2010, 2012). An intervention studied includes many causal relationships that researchers cannot separate out. For example, if an after school tutoring program was an intervention, the intervention happens in a context that was enabled by tutors and students who may have different characteristics and intentions. Tutors may use various strategies, and they may have various levels of knowledge, including tacit understanding of how things work in a specific school, and they may approach their work differently. So, when this after school tutoring showed effectiveness in improving students' math skills, it is difficult to know how the combinations of components led to effectiveness. As it is, "RCTs typically do not assess how and why impacts occur, how a program might be modified to improve program results." (Brass, Nunez-Neto & Williams, March 7, 2006, page 16). Thus, even if the evidence-based intervention is identified, programs have to find ways to make the intervention work. To support implementation, some research reports include contact information for implementation support, logic model, theoretical framework, and minimum dosage; however, replication is not straightforward.

Furthermore, because the result of an RCT stems from one intervention, an important question for programs is how interventions can be coordinated so the programs can produce outcomes.

> *RCTs also typically do not provide a full picture of whether unintended consequences may have resulted from a program or indicate whether a study is using valid measures or concepts for judging a program's success (e.g., assessing a study's or a measure's construct validity). Many of these kinds of questions have been considered to be more*

appropriately addressed with observational or qualitative designs.
(Brass, Nunez-Neto & Williams, March 7, 2006, p. 16)

This limitation applies to research syntheses because they often focus on outcomes that are determined by the reviewers. Finally, there is a gap between "effective intervention" and decision-to-adopt intervention because "deciding to adopt an intervention involves other considerations in addition to effectiveness, such as cost and suitability to the local community" (GAO, 2009, cover page).

Taken together, although the federal government intended to improve TRIO programs by implementing evidence-based interventions, because of the design limitations of RCTs, TRIO cannot simply replicate the interventions. The way the Competitive Preference Priority (CPP) policy instrument was presented to TRIO programs minimized one critical aspect that is necessary for delivering results, i.e., consideration of what it takes for professionals to make decisions on how to implement the "evidence-based" interventions.

Need for Logics of Evidence Use in TRIO Programs

Discussions in evidence-based practices in medicine, social services, and education have highlighted the complexity of professionals' work, partly because there is a tension between research knowledge and practitioner knowledge when evidence-based practice policies are put in practice. Drawing upon this line of discussion, I argue that TRIO programs need to draft logics of evidence use and demand research that aligns with how TRIO professionals work.

Over the decades, clinical knowledge has been considered to be a critical component for practicing evidence-based medicine, as the definition of evidence-based medicine below indicates.

> *Evidence based medicine is the conscientious, explicit and judicious use of current best evidence in making decisions about the care of individual patients. The practice of evidence based medicine means integrating individual clinical expertise with the best available external clinical evidence from systematic research. (Sackett, et al, 1996, p. 71)*

In evidence-based medicine, it is widely accepted that "any clinician who feels restricted to behave only as the evidence dictates is missing the concept of evidence based medicine" (Stevens & Hootman, 2004, page 84), and it requires "bottom-up approach" (Sackett, et al. 1996, page 72). Similarly, researchers have discussed the importance of professional knowledge in evidence-based practice in education. For example, Davis (1999) states in his article titled, "What is evidence based education?" as follows:

> *Establishing best practice, in both education and health care, is more than a matter of simply accessing, critically appraising, and implementing research findings. It also involves integrating such knowledge with professional judgment and experience (Davis, 1999, pp.116-7)*

However, as exemplified in the CPP and other guidelines, there is little acknowledgement about the role of professional knowledge in policies that promote evidence-based practice in education and social services. Policy instruments project an image that professionals should replicate evidence-based interventions. As Thomas Archibald (2016) reported in his study, there is a risk that this simplified view of replication is manifest in the implementation of evidence-based interventions in the federally funded programs. His study found status differentials of knowledge produced scientists and local programs, which consequently challenges the ideal of evidence-based intervention movement. During implementation, the program developer makes a tacit adaptation of interventions based on their "more privileged position as expert knowers while community educators feel limits placed on their agency to respond to and adapt to the needs and realities in their local communities" (Archibald, 2016). I argue it is important for TRIO programs to examine the idea and practice of "evidence-based," because this issue involves "welfare of very many people, not just the egos of a few" (Scriven, 2008, p.24)

Practitioner-Researcher Differences in Clinical Needs and Notions of Causality. To argue against the simplified notion of evidence-based practice, researchers have discussed that knowledge which comes from research and knowledge for the needs of professionals are different. Knowledge from external research, i.e., RCTs or synthesis of RCTs, does not fully meet the knowledge needs of professionals. This line of discussion suggests that for TRIO programs to improve practice, they need to articulate how TRIO professionals work and they need to identify research questions that reflect TRIO professionals' work.

Burton and Chapman (2004) reviewed the concept of "causality" on which "evidence-based practices" hold. They argue that causality in RCTs and the causality professionals use in their work differ, thus, over-reliance to external clinical evidence, which derive from RCTs, does not reflect how health care and social services professionals make decisions about what intervention they should provide to their clients. They argue that theory or conception of causality by healthcare and social service professionals is a "generative model of causality" where "causation can be identified in the activation of an internal potential of a system or substance, under particular conditions, to generate particular outcomes." Thus, useful knowledge for professionals should "capture the complexity of the context of application via provision of an account of the relationships among at least the following elements: context, participants, practices, causal mechanisms, regularities and outcomes" (Burton and Chapman, p. 9). On the contrary, the concept of causality underlying evidence identified through RCTs "associationist model," which "assumes causation is invisible, so the only way to establish what causes what is to rule out 'spurious associations' between independent and dependent variables." (p. 13). Similar to the challenge associated with high "causal density" in an intervention, the external clinical evidence does not necessarily provide all the information professionals need to make causal inferences to do their work.

Professional decision-making includes more than complex causality. Biesta (2007) argues that professionals sometimes do not do the most effective practice because of the consequence of one effect on another. Thus, "in education means and ends are not linked in a technical or external way but that they are related internally or constituently" (Biesta, 2007, p. 10), and value and ethics play part. Hammersley (2001) adds another level of complexity. Professionals do not always deal with an individual student, but a group of students with different levels of motivation and different goals, so the same action of a professional can have multiple consequences, which

are differentially distributed across students. Hammersley (2001) discusses conceptualizing educators work as a series of linear process from information to practice has a limitation (Hammersley, 2001).

These discussions on how professionals make decisions explain the challenges TRIO professionals face when they are asked to replicate evidence-based interventions. TRIO program professionals must decide how to implement evidence-based interventions in relation to context, value, other interventions, and what is seen as best for students who are in a particular high school or college setting that promotes certain values and behaviors. The complexity of professionals' decision making includes the fact that they have to deal with not only multiple causality between action and outcomes, but also causalities operating in different time frames, layers of context, beliefs, and values.

Professional Learning. While the above discussion highlights what constitutes as needed information for professionals, discussion about how professionals improve their practice highlights the importance of clinical knowledge and the role of external research knowledge in the integration with professional knowledge. Biesta (2007) discusses, by drawing upon John Dewey's conceptualization of learning, that experimentation and reflections are important part for professionals to change their actions. According to Dewey's theory, professionals continuously build knowledge as they interact with the world. Thus, knowledge produced in the past, including knowledge from research and experiences (including lessons learned from trial and error), helps professionals by providing hypothesis for intelligent problem solving, but it does not tell them what they should do. For professionals, "the only way to use this knowledge is as an instrument for undertaking intelligent professional action" (Biesta, 2007, p. 17). Since problem-solving for professionals is context-specific and the context changes constantly, "for Dewey, professional action is not about following tried and tested recipes, but about addressing concrete and in a sense, always unique problems" (Beista, 2007, p. 16). If we follow this notion of how professionals work, professionals have to be reflective of what they know, how they know, and how their knowledge, belief, and actions are shaped by surroundings, and they are to continue to refine their working knowledge of what works and how to make it work, so they can take "intelligent professional action" (Beista, 2007, p.17).

This point that professionals' learning involves experimentation and theory building is important. Studies have shown that when teachers change their instructional practices, they need to make sense of why and how a specific instruction strategy works, and to change their strongly held belief of what works, they need to experience success as they try new instruction in their classrooms (Coburn, 2001; Spillane, Reiser & Reimer, 2002). Their sense-making shapes how they implement it, and the change of practice happens when research evidence is connected closely with the situation where they improve practice (Simons, et al., 2003).

Burton and Chapman (2004) propose to take a notion of practitioner as theorists. They present exhaustive list of types of information (including research and non-research information) that shape various sub-decisions that lead to a decision of what intervention to provide to a client. They make a point that knowledge from RCT "always leave gaps" in information and professionals have to "fill in the gaps" by using various sources of information (Burton & Chapman, p. 20). Similar to Biesta's (2007) argument about how professionals develop "intelligent professional action," they argue that " 'evidence' that does not engage with these

practical theories is likely to have little impact on practice, and hence on the outcomes experienced by the people served" (Burton and Chapman, p. 20).

Varieties of Implementation of Evidence in TRIO. Based on brief conversations with TRIO programs, I developed the impression that TRIO programs are implementing evidence-based interventions in various ways. There are theories programs used to select a specific evidence-based intervention and decide when and how to replicate the evidence-based interventions in their programs. I think it is important for TRIO programs to document their theories and identify questions that will help them to build and refine these theories. I think TRIO programs should demand these questions be researched. Of course, it is unlikely that each research question will be studied, but I think the research question should start from the work of professionals by articulating why and how that can be a guide for professional action. And we should be very clear that the research does not provide us with a recipe, but it is professionals who decide how external evidence can be integrated into their work, if at all. Rather than being subordinate to the notion that they have to implement evidence-based interventions, I think TRIO programs should step back and intentionally observe and articulate what is the logic of evidence use in TRIO programs.

But how can we move ourselves to articulate and address important questions? How can we articulate our logic? After experiencing programs that required grantees to use "evidence-based" intervention, I think the grant-making field is becoming more aware of the limitations of an overreliance on evidence-based intervention. For example, Wandersman and colleagues (2016) report that the 2015-2020 Teen Pregnancy Prevention grant program by the Office of Adolescent Health at Health and Human Services "has taken lessons" (p.14) from the previous grant cycle of 2010-2015 that required applicants to implement evidence-based interventions from its list. The new grant cycle still requires grantees to implement evidence-based interventions from the list, but it also includes funds to provide capacity-building assistance to organizations on implementing evidence-based interventions. I think it is a hopeful sign because it opens up a space for a discussion on how to make evidence-based intervention work and how to take up local knowledge. I hope the grant-making of TRIO programs will be affected by this policy shift.

Going back to the question of how we can go about articulating our logic and how we can generate questions, I list ongoing activities in the TRIO community and resources that could contribute to supporting each program. As I discussed above, because the professional's work involves complex decision making, what is important could vary, and it is the professionals who have to make intelligent decisions. There is not one size fits all approach for professionals to articulate logics for evidence use. So, I list some of the resources and tools that may be helpful. Regional meetings and conferences can be a place where TRIO practitioners can compile research questions and look for knowledge resources. TRIO program personnel view these meetings as helpful and they get ideas for improving their practices. It would be helpful as a community to know how this learning process happens, and if there are better ways to promote knowledge sharing and transfer.

- **Logic model:** As a part of online community of practice activity, TRIO programs have shared tools that can be used for better program planning and evaluation. As described during the webinar of the COE's Research, Evaluation, and Data Use Community of

Practice (CoP), a logic model is a helpful tool to document theories and gap in knowledge (Wahl, April 20, 2016). Logic models could be also helpful for assessing if intervention is ready to be researched (Epstein & Klerman,2012).

- **Development of evaluation capacity:** Program evaluators have developed tools, assessments, and guiding questions that can be used in the program meetings and to develop evaluation capacity (for example, Buckley, et al, 2015 on teaching of evaluative thinking, which provides strategies and examples of activities for promoting evaluative thinking). Action research, as shared in the CoP webinar, can provide contextualized knowledge to fill in the gap between what is known from external research and how it is rolled out in a program (Arreola, Pena, et al., 2015; Rohse & Kaplan, January 28, 2016). Empowerment Evaluation has been used in public health programs and there are tools developed to support programs bridging the gap between evidence-based interventions and practice. For example, Getting to Outcomes (GTO) provides 10 steps for organizations to plan, implement, and evaluate their programs (Rand Corporation, no date; Wandersman, et al., 2016) to educational organizations. Both action research and empowerment evaluation take the importance of context in which the program unfolds, which TRIO practitioners also need to be able to articulate when they develop the logic of their work.

Final Thoughts

In this essay, I portrayed CPP as a reflection of idiosyncrasy in policymaking, which consequently minimized the considerations of how professionals learn and improve practices. I think this is a lost opportunity for both the federal government and TRIO programs. Some TRIO programs took CPP as a hoop to jump through, which will possibly lead to symbolic adoption of a new practice.

I argued that TRIO programs should draft logics of evidence use that align with their work and generate research questions. How to develop organizational capacity to use evidence is an important element for program improvement, but policy instruments that TRIO programs came across did not highlight it, and it is often told at the end of the program cycle. I am throwing out a proposal that this often-invisible process be documented, shared, and discussed, so programs can use their logics of evidence use as levers to ask for knowledge that they need for program improvement.

REFERENCES

Archibald, T. (2015). "They just know": The epistemological politics of "evidence-based" non-formal education. *Evaluation and Program Planning*, *48*, 137-148.

Arreola Pena, M., Gebremedhin, Y., Kadima, G.K., Kaplan, R., Kirshner, B., & Raitz, K. (2015). *Students of color are motivated agents of change: Why aren't we joining your programs?* Boulder, CO: CU Engage: Center for Community-Based Learning and Research.

Berlin, G.L. (May 10, 2016*). Using evidence as the driver of policy change: The next steps in supporting innovation, continuous improvement, and accountability.* Testimony before the Senate Finance Committee hearing "Can evidence based practices improve outcomes for vulnerable individuals and families?" Retrieved from http://www.finance.senate.gov/imo/media/doc/10MAY2016Berlin.pdf

Biesta, G. (2007). Why "what works" won't work: Evidence-based practice and the democratic deficit in educational research. *Educational Theory*, *57* (1) 1-22.

Brass, C.T., Nunez-Neto, B., & Williams, E.D. (March 7, 2006). *Congress and program evaluation: An overview of Randomized Controlled Trials (RCTs) and related Issues.* Congressional Research Service.

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.

Burton, M., & Chapman, M.J. (2004). Problems of evidence based practice in community based services. *Journal of Learning Disabilities, 8*(1). 56-70.

Coalition for Evidence-Based Policy. (n.d.). *Mission/Activities.* Retrieved from http://coalition4evidence.org/mission-activities/

Coburn, C.E. (2001) .Collective sense making about reading: How teachers mediate reading policy in their professional communities. *Educational Evaluation and Policy Analysis*, *23*(2). 145-170.

Corporation for National and Community Services. (2015). *Notice of Federal funding opportunity, AmeriCorps State and National Grants FY2016*. CFDA Number: 94.006. Retrieved from http://www.nationalservice.gov/sites/default/files/documents/2016_AmeriCorps_State_and_National_Notice_100915.pdf

Davis, P. (1999). What is evidence-based education? *British Journal of Educational Studies*, *47*(2), 108-121.

Epstein, D., & Klerman, J.A. (2012). When is a program ready for rigorous impact evaluation? The role of a falsifiable logic model. *Evaluation Review, 36(*5). 375-401.

Haskins, R., & Margolis, G. (2015). *Show me the evidence: Obama's fight for rigor and evidence in social policy.* Washington DC: The Brookings Institution.

Hammersley, M. (2001). *Some questions about evidence-based practice in education.* Paper presented at the symposium on "Evidence-based practice in education" at the Annual Conference on the British Educational Research Association, University of Leeds,

England, September 13-15, 2001. Retrieved from http://www.leeds.ac.uk/educol/documents/00001819.htm

Higgins J.P.T., & Green, S. (Eds.). *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Retrieved from www.handbook.cochrane.org.

Institute of Education Sciences. (n.d.). *The Institute of Education Sciences: New wine, new bottle.* A presentation by IES Director Grover (Russ) Whitehurst. Retrieved from:
http://ies.ed.gov/director/speeches2003/04_22/2003_04_22.asp
http://ies.ed.gov/director/speeches2003/04_22/2003_04_22a.asp
http://ies.ed.gov/director/speeches2003/04_22/2003_04_22b.asp

Manzi, J. (2010). What social science does – and doesn't --- know: Our scientific ignorance of the human condition remains profound. *City Journal.* Summer 2010. Retrieved from http://www.city-journal.org/html/what-social-science-does%E2%80%94and-doesn%E2%80%99t%E2%80%94know-13297.html

Manzi, J. (2012). *Uncontrolled: The surprising payoff of trial-and-error for business, politics, and society.* New York: Basic Books.

McNeil. M. (January 29, 2013). *Ed. Dept. raises evidence, research ante in grant awards.* Education Week. Retrieved from http://www.edweek.org/ew/articles/2013/01/30/19i3.h32.html?tkn=WSXFkOv3StWOdCTkwrgJHtidAXUfyGccC5fm&print=1

National Research Council. (2002). Scientific research in education. Committee on scientific Principles for Education Research. Shavelson, R. J., and Towne, L Editors. Center for Education. Division of Behavioral and Social Sciences and Education. Washington, DC: National Academy Press.

Nussle, J., & Orszag, P. (Eds). (2015). *Moneyball for government*. 2nd Edition. Results for America.

Office of Federal Register. Scientifically Based Evaluation Methods: Department of Education. Notice of Final Priority. 70. Fed. Reg. 3586-3589. Retrieved from https://www2.ed.gov/legislation/FedRegister/finrule/2005-1/012505a.html

Office of Federal Register. Applications for New Awards; Student Support Services Program; Office of Postsecondary Education, Department of Education. 79. Fed. Reg. 75722-75730. Retrieved from https://www.gpo.gov/fdsys/pkg/FR-2014-12-18/pdf/2014-29769.pdf

Office of Federal Register. Applications for New Awards; Talent Search Program; Office of Postsecondary Education, Department of Education Notice, 80. Fed. Reg. 79574-79580. Retrieved from https://www.gpo.gov/fdsys/pkg/FR-2015-12-22/pdf/2015-32089.pdf

Office of Management and Budget. (n.d.). *Evidence and evaluation.* Retrieved from https://www.whitehouse.gov/omb/evidence

Office of Management and Budget. (2004). *What constitutes strong evidence of a program's effectiveness?* Retrieved from https://www.whitehouse.gov/sites/default/files/omb/part/2004_program_eval.pdf

Office of Management and Budget. (January 29, 2007). Program Assessment Rating Tool Guidance No. 2007-02. Subject: Guidance for Completing 2007 PARTs. Retrieved from https://www.whitehouse.gov/sites/default/files/omb/part/fy2007/2007_guidance_final.pdf

Office of Management and Budget. (May 18, 2012). Memorandum to the heads of executive departments and agencies. Subject: Use of evidence and evaluation in the 2014 budget. Retrieved from https://www.whitehouse.gov/sites/default/files/omb/memoranda/2012/m-12-14.pdf

Orzag, P. (June 8, 2009). *Building rigorous evidence to drive policy.* Retrieved from https://www.whitehouse.gov/omb/blog/09/06/08/BuildingRigorousEvidencetoDrivePolicy/

Rand Corporation. (n.d.). *Getting to outcomes: Improving community-based prevention.* Retrieved from http://www.rand.org/health/projects/getting-to-outcomes.html

Rohse, C.S., & Kaplan, R.G. (2016). *Be an action hero: How to use participatory action research in your TRIO program evaluation.* A webinar on January 28, 2016 at COE's Research, Evaluation and Data Use Community of Practice. Retrieved from https://higherlogicdownload.s3-external-1.amazonaws.com/COENET/COEWebinar_28January2016.pdf?AWSAccessKeyId=AKIAJH5D4I4FWRALBOUA&Expires=1468945442&Signature=AenhsiYbFxXEQWm%2BK0vWZdiL%2BZA%3D

Sackett, D.L., Rosenberg, W.M.C., Gray, J.A.M., Haynes, R.B., & Richardson, W.S. (1996). Evidence based medicine: What it is and what it isn't. *BMJ. 312*(13). 71-72.

Scriven, M. (2008). A summative evaluation of RCT methodology: An alternative approach to causal research. *Journal of Multidisciplinary Evaluation*, *5*(9), 11-24.

Simmons, H., Kushner, S., Jones, K., & James, D. (2003). From evidence-based practice to practice-based evidence: The idea of situated generalisation. *Research Papers in Education, 18*(4). 347-364.

Spillane, J.P., Reiser, B.J., & Reimer, T. (2002). Policy implementation and cognition: Reframing and refocusing implementation research. *Review of Educational Research*, *72*(3), 387-431.

Steves, R., & Hootman J.M. (2004). Evidence-based medicine: What is it and how does it apply to athletic training? *Journal of Athletic Training*, *39*(1), 83-87.

U.S. Department of Education. (2001). Exemplary & Promising Safe, Disciplined and Drug Free Schools Programs 2001. Safe, Disciplined and Drug Free Schools Expert Panel. Retrieved from https://www2.ed.gov/admins/lead/safety/exemplary01/exemplary01.pdf

U.S. Department of Education. (2003). Identifying and implementing educational practices supported by rigorous evidence: A user friendly guide. Prepared for the Institute of Education Sciences by the Coalition of Evidence-Based Policy. Retrieved from https://www2.ed.gov/rschstat/research/pubs/rigorousevid/rigorousevid.pdf

U.S. Department of Education. (2014). *U.S. Department of Education strategic plan for Fiscal Years 2014-2018.* Retrieved from https://www2.ed.gov/about/reports/strat/plan2014-18/strategic-plan.pdf

U.S. Department of Education. (2016). U.S. *Department of Education FY 2015 Annual Performance Report and FY 2017 Annual Performance Plan.* Retrieved from http://www2.ed.gov/about/reports/annual/2017plan/2015-2017-apr-app.pdf

U.S. Government Accountability Office. (2009). *A variety of rigorous methods can help identify effective interventions.* Washington DC. Retrieved from http://www.gao.gov/new.items/d1030.pdf

Wahl, L. (April 20, 2016). *What is logic model?* A webinar on April 20, 2016 at COE's Research, Evaluation and Data Use Community of Practice. Retrieved from https://higherlogicdownload.s3-external-1.amazonaws.com/COENET/COE%20Logic%20Model%20Presentation%20Apr2016.pdf?AWSAccessKeyId=AKIAJH5D4I4FWRALBOUA&Expires=1468946413&Signature=MhUHtjTlAHo1kiq4ZRL4BtTc2%2Bc%3D

Wandersman, A., Alia, K., Cook, B.S., Hsu, L.L., & Ramaswamy, R. (2016). Evidence-based interventions are necessary but not sufficient for achieving outcomes in each setting in a complex world: Empowerment evaluation, Getting to Outcomes, and demonstrating accountability. *American Journal of Evaluation*, 1-18.

What Works Clearinghouse (2014). *What Works Clearinghouse Procedures and Standards Handbook Version 3.0.* Retrieved from http://ies.ed.gov/ncee/wwc/pdf/reference_resources/wwc_procedures_v3_0_standards_handbook.pdf

What Works Clearinghouse. (n.d.). *About the WWC*. Retrieved from https://ies.ed.gov/ncee/wwc/aboutus.aspx

Weiss, C.H., Murphy-Graham, E., & Birkeland, S. (2005). An alternative route to policy influence: How evaluations affect D.A.R.E. *American Journal of Evaluation*, *26*(12), 12-30.

About the Author:

Dr. Mika Yamashita is the Associate Director of the Pell Institute for the Study of Opportunity in Higher Education, Council for Opportunity in Education. She is a program evaluator and worked as both internal and external evaluator of college access and other social services programs.

Contact Information: MIKA YAMASHITA, ASSOCIATE DIRECTOR, The Pell Institute for the Study of Opportunity in Higher Education, mika.yamashita@pellinstitute.org