TRIO, THE WHAT WORKS CLEARINGHOUSE, AND THE COMPETITIVE PREFERENCE PRIORITIES (CPPS): AN IMPOSED STRUCTURED PRACTITIONER-RESEARCH COLLABORATION

BY MARGARET CAHALAN

Abstract

The "imposed structure" of incorporating competitive preference priorities (CPPs) for the use of evidence-based strategies as defined by the What Works Clearinghouse (WWC) has the potential to fundamentally redefine the relationship between practice and research within federal grants in education. Given the focus of this compilation on practitioner-research collaboration in this essay, we explore the implications of these new regulations and associated competitions for TRIO programs and service. We present an overview of the WWC history and the content of the reviews in the areas of relevance to TRIO. Then we review the results of the specific application in the SSS competitions. Finally we discuss the limitations of using the WWC and give some reflections on making the use of "evidence based research" more helpful.

Reflecting the articulated commitment of the Obama administration to foster the increased use of "evidence" based policies to achieve program goals, in 2013 the Department of Education regulations known as EDGAR were officially modified to allow the Secretary of Education to foster increased use of "evidence-based" strategies through the awarding of competitive financial advantages to programs that have "evidence of success" (Federal Register 2013-08-13). These modifications provided authority to the Secretary to use "absolute priorities" or "competitive preference priorities (CPPs)" in grant competitions to give priority to those projects that were using "evidence-based" practice, and that proposed to evaluate these practices using rigorous evaluation procedures. The definitions of and standards for "research evidence" specified in the EDGAR regulations were those adopted and put forth by the What Works Clearinghouse (WWC).

These new EDGAR regulations were consistent with the 2012 OMB circular calling for using Competitive Preference Priorities (CPPs) in competitions for existing programs that are "difficult to change." TRIO became an obvious target for applying the new regulations as the largest discretionary competitive grant program in the U.S. Department of Education which is intentionally structured to have a high level of ongoing funding of existing programs. Discretion was exercised by the Secretary to use the CPPs for the next TRIO competitions following the above noted formal modification of the EDGAR regulations that occurred in 2013. The Student Support Services (SSS) competition in 2015; the Talent Search competition of 2015-2016, and the Upward Bound

Reprinted from *Reflections on connecting research and practice in college access and success programs*, pp.104-130, by H. T. Rowan-Kenyon, M. Cahalan, & M. Yamashita, Eds., 2018,Washington, DC: Pell Institute.

competition of 2016-17 each utilized CPPs as a method of fostering the implementation of evidence-research based strategies. This "imposed structure" of incorporating competitive preference priorities (CPPs) for the use of evidence-based strategies as defined by the What Works Clearinghouse (WWC) has the potential to *fundamentally redefine the relationship between practice and research within federal grants in education*. Given the focus of this compilation on practitioner-research collaboration in this essay, we explore the implications of these new regulations and associated competitions for TRIO programs and services.

Specifically in this essay we:

1. Give a brief overview of the history and the definitions developed by the WWC;
2. Present an overview of the *Interventions, Single Study Reviews*, and *Practice Guides* in the WWC that are within the two content areas of most relevance to TRIO (*Pathways to Graduation* and *Postsecondary*);
3. Use SSS as an example and present some preliminary data on the impact of the CPP points on the 2015 SSS competition scores; and
4. Discuss the identified limitations to use of the WWC and give some reflections on making the "using evidence based research movement" more useful.

Overview of the What Works Clearinghouse (WWC) History and Growth as the Major Arbiter of Research Legitimacy for Federally Funded Education Programs

The WWC was established in 2002 by ED with a formal mission to "*be a resource for informed education decision making"* by identifying those interventions "that work" so that practitioners could have increased guidance to inform practice. The Clearinghouse was also charged with producing user-friendly guides for educators on effective instructional practices in order to understand what instructional programs have been shown to be effective. The WWC replaced the ERIC annotated bibliographies and research summaries that had been developed in the 1980-1990s to provide summaries of research in the field. The website asserts that for over a decade, the WWC has been "a central and trusted source of scientific evidence for what works in education to improve student outcomes."

Contract Organization and Costs. The work of the WWC is conducted under a set of contracts held by several leading firms with expertise in education and research methodology, and managed by the Institute for Education Sciences (IES), the unit in ED focused on research and evaluation. Since its inception in 2002, ED has held competitions at roughly five-year intervals for the major contract to implement and manage the WWC. Periodic smaller competitions for additional work have also been held. The major contractors in 2017 were Mathematica Policy Research (MPR), Development Services Group (DRS)—responsible for postsecondary reviews, Inc. and Sanametrix, Inc. The WWC is not an inexpensive operation. The costs of the WWC contracts since 2006 have averaged well over $10 million per year across the major contract and sub-contracts.

Enshrining the WWC in U.S. Department of Education Regulations. As noted, the *Notices of Inviting Applicants* for the most recent TRIO grant solicitations programs utilize the procedures, definitions and ratings developed by the What Works Clearinghouse (WWC 2014). Following the general guidance from OMB, ED has taken a tiered approach that has three levels that are summarized below. Figure 1 gives the definitions of Randomized Control Trial (RCTs) and Quasi-Experimental Design (QED) needed to meet the Moderate Evidence requirement of the CPPs.

- **Promising for developmental studies**—Strong theoretical conceptualization and correlational evidence

- **Moderate Evidence**—Meets WWC methods standards "with or without reservations." This involves interventions that have at least one well-designed and implemented Randomized Control Trail (RCT) or at least one well-designed and implemented Quasi Experimental Design (QED) that allows establishing a positive significant and substantive causal inference between the intervention and outcomes and no negative significant relationships. Based on this evidence, interventions studied are evaluated as "effective," thus "evidence-based." Well-executed RCTs can meet WWC standards "without reservations." Well-executed Quasi-experimental designs (QED) can meet WWC standards "with reservations."

- **Strong Evidence**—Meets WWC methods standards "without reservations" with significant positive outcome and substantive effect size. This involves interventions that have at least one well-designed and implemented RCT that establishes significant and substantive positive relationships and no negative relationships.

Rationale of the CPPs. While the Trump administration in its FY2018 budget proposals has seemingly returned to and extended the Bush Administration's tactic of selectively using research evidence to support its proposals for advancing, cutting or eliminating federal programs, the Obama Administration articulated a more systematic approach. This approach fostered evidence-based policy making and use of evidence as keys to "make government work effectively" (OMB, May 18, 2012; OMB 2016). The government-wide strategies sought to allocate more funds to programs that were supported by stronger evidence, and to require rigorous evaluations to produce "strong" evidence. This policy included supporting social innovation funds in which the government paid more to programs that could meet required outcomes. For existing grant programs, the OMB encouraged agencies to "provide points or significant competitive preference priorities to programs that are backed by strong evidence" (OMB, May 2012). These strategies were based on the idea that the government should invest taxpayer money on programs that are most likely to bring results (Haskins and Margolis, 2014; Orszag 2009).

Figure 1: Definitions of two methodologies that would meet the Moderate Level of Evidence as defined in the Department of Education *Student Support Services Applications for New Awards* notice

***Randomized controlled trial* means** a study that employs random assignment of, for example, students, teachers, classrooms, schools, or districts to receive the intervention being evaluated (the treatment group) or not to receive the intervention (the control group). The estimated effectiveness of the intervention is the difference between the average outcome for the treatment group and for the control group. These studies, depending on design and implementation, can meet WWC Evidence Standards without reservations.

***Quasi-experimental design study* means** a study using a design that attempts to approximate an experimental design by identifying a comparison group that is similar to the treatment group in important respects. These studies, depending on design and implementation, can meet WWC Evidence Standards with reservations (they cannot meet WWC Evidence Standards without reservations).

Applications for New Awards; Student Support Services Program A Notice by the Education Department on 12/18/2014)

The Movement from Studying Outputs and Outcomes to Using Research Evidence. Newcomer (2016) has done historical research tracing the growth of emphasis within government agencies from what she calls a stress on measuring "outputs and outcomes" to the use of "Demonstrated Evidence-Based Interventions" (DEBIs). She outlines the movement from reports such as the Hatry's (1967) Senate report on Measuring Program Effectiveness, and groups such as the World Bank calling for developing outcome measures of the 1990s and the Government Performance Results Act (GPRA) of 1994 to the Executive Order of the Bush Administration establishing the Program Assessment Rating Tool (PART). Faced with intense criticism of the Bush OMB PART order, the Obama Administration discontinued PART and subsequently issued new OMB Guidance on Tiers of Evidence that form the basis of the Obama Administration's approach. Newcomer also identifies key challenges faced with regard to the use of DEBIs and concludes that there is an overstating of the "ease of the flow of evidence to practice." She cites key challenges, including inadequate attention to support factors and understanding the way causal mechanisms need to work together to produce expected results, and the focus on replication with fidelity rather than adapting models to changing contexts. She also notes that measuring fidelity of implementation and impact are very expensive and often difficult ethically.

The Shift from Asking "Does this Program Work?" to Asking "What Strategies Work and Under What Circumstance?" Under the Obama administration there was a change to move away from the emphases on evaluating federal programs as a whole which had dominated the period of the 1970s to the early 2000s to a move to look at the efficacy of specific strategies within a program. Given the broad sweep of federal programs usually intentionally allowing a range of services to achieve goals, the overall program evaluations suffered from a lack of clear

specification of the actual "treatment" being studied and also a lack of control over the extent to which the control group or comparison group was also getting the services due to ethical considerations. They tended to be black-box evaluations with numerous study error issues, but they were often used to justify reduced or level funding, or in the case of the Upward Bound program, using what were later found to be "error filled" conclusions, to justify the "zero funding" recommendations of the Bush Administration's budgets in FY2005 and FY2006 (Cahalan and Goodwin, 2014). To some extent, this shift to focus on specific strategies offered a relief from the defensive position in which most federal programs increasingly found themselves at the end of the 20th century due to these overall evaluations.

Researcher-Practitioner Collaboration--The Logic Flow of the Goals of the CPPs. Figure 2 outlines the logic flow of the articulated goals of the use of the CPPs in the TRIO competitions held under the Obama Administration. As stated in the "Invitations to Apply", ED articulated the belief that use of evidence-based practices will result in both "*a better competition and in better results for students*." **ED also articulated its interest in possibly partnering in research with the grantees in the priority topic.** As the logic flow indicates, there is anticipated spillover impact toward increasing general evidence-based practice and research use throughout the government and in non-profit social programming. The 2015 SSS application stated:

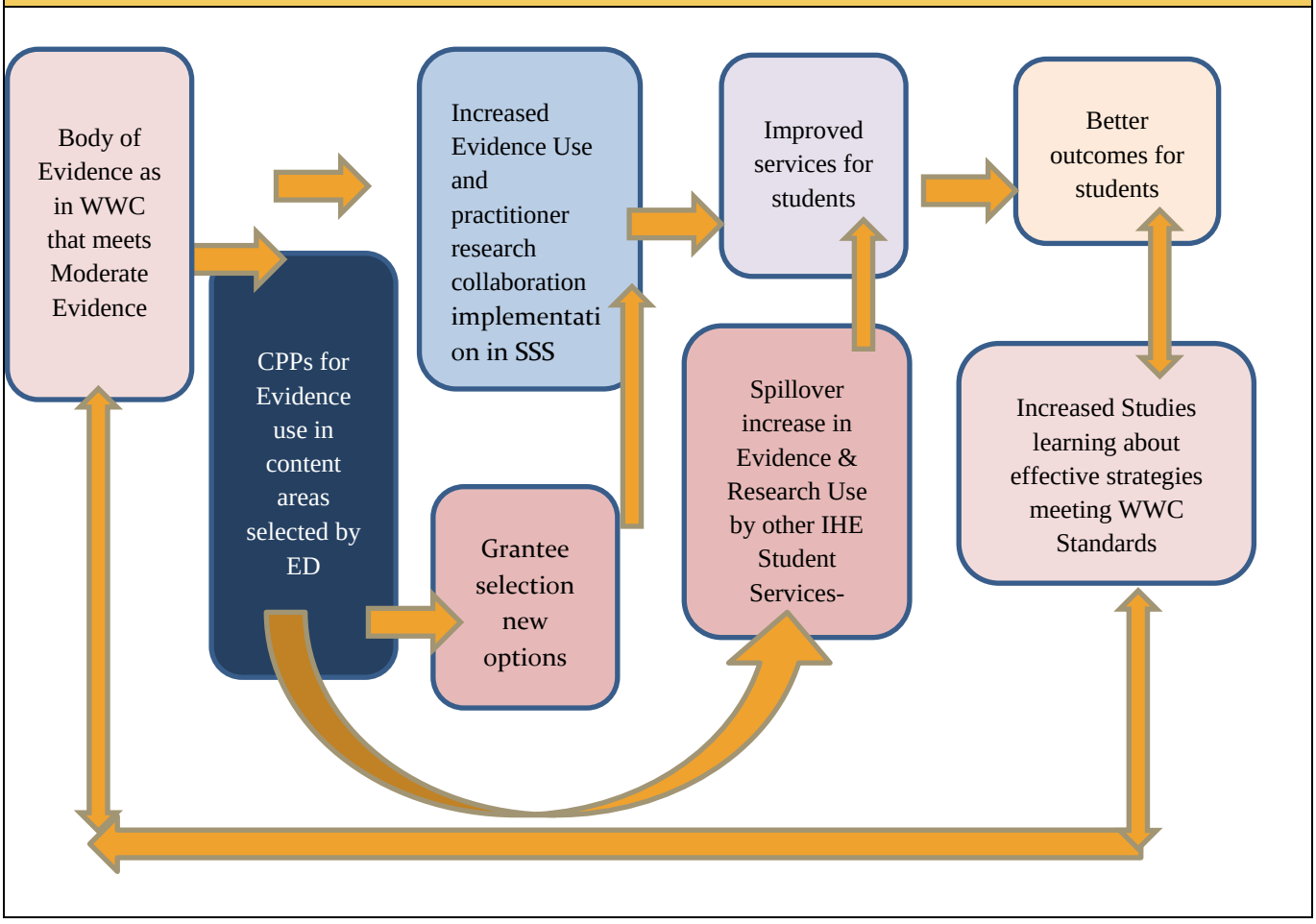
> *"In recent years, the Department has placed an increasing emphasis on promoting evidence-based practices through our grant competitions. We believe that encouraging applicants to focus on proven strategies can only enhance the quality of our competitions and the outcomes of students who participate in our programs.....**The Department is sufficiently interested in this priority topic that we may later seek to partner with successful applicants to conduct research and evaluation"** (2015 Invitation for Applications for Student Support Services Grants).*

Selected Related Research on Research Evidence Use. While the use of evidence-based practices within the U.S. Department of Education is too new to have a large body of information, scholars have generally been more cautious about the promise of evidence-based policymaking than government agencies. Sutherland and colleagues observed that "the normative claim that policy should be grounded in an evidence base is itself based on surprising weak evidence" (Sutherland et al; 2012, cited by NRC, 2012). A report entitled the *Science of Using Science* from the EPPI Center at the University of London (Langer et al., 2016) presents results from a comprehensive meta-analysis of the efficacy of various strategies to increase research use by decision makers.

The EPPI report found some evidence of positive effects with the following interventions:

- Those that facilitated active access to research evidence through communication and evidence repositories;
- Those that built on decision-makers' skills to access and make sense of evidence, and
- Those that foster changes to the decision making structures, incentives, and processes.

Figure 2: Logic flow of goals and anticipated outcomes of Competitive Preference Priorities (CPPs) for Research Evidence Use and Increased Practitioner-Research Collaboration as articulated in U.S. Department of Education's 2015 *Invitation for Applications for Student Support Services Grants*:



Review of WWC Relevant Interventions, Single Studies and Practice Guides

The WWC does not purport to be a source for meta-analysis or systematic review of the literature on a given topic such as is the approach of groups such as the EPPI center at the University of London, or even in a less systematic manner to take the approach of the ERIC Topical Literature Reviews of the 1990s.

The WWC has evolved over time since its inception, but not in the direction of meta-analyses. Initially, there was more of a focus on named "Interventions," for which there could be a search of the literature for research with regard to a given named "intervention." This approach has a number of pitfalls, including the fact that the named "interventions" often were dynamic programs that changed over time, and that often few studies could be found that met the standards, and these were not necessarily representative of the project at the current time. More recently, there seems to be a greater focus on Single Study Reviews and Practice Guides. The three types of reports constituting the current searchable WWC data base are:

- Intervention Reports;
- Single Study Reviews, and
- Practice Guides.

The WWC initially was primarily focused on K-12 topics; however, prompted by the increased use of the WWC data base for postsecondary federal competitions, there has been additional focus in college access and success topics and strategies since 2012. These are categorized under the key words "Pathways to Graduation" and "Postsecondary" in the WWC data base. Figures 3-6 summarize the results of the WWC reviews for *Intervention Studies* and for *Single Studies* in these two areas.

The Category of "Intervention Studies" Reviewed. Figure 3 gives summary information and Figure 4 lists the names of the "Interventions" that have been reviewed by the WWC under the general categories of *Pathways to Graduation*, which may involve either high school or college graduation; and under the category *Postsecondary*. Interventions are cross-listed in the system so that all of the 7 Interventions listed under *Postsecondary* are also listed under *Pathways to Graduation* category. As of December 2016, there were 35 named "Interventions" that had been reviewed. Typically this means that more than one potential study has been found, or that the named program has been around for more time. Among those *Interventions* without any studies meeting standards there are a number of well-known programs. Of the 35 *Interventions* that WWC studied in these topic areas, 13, or 37 percent, had no study that met the methods standards for the WWC. Those without any studies meeting the methods' standards cannot be reviewed as to effectiveness by the WWC. Notably this listing of programs with "no studies" meeting the WWC methods standards includes well-known programs such as "I Have a Dream," "Puente," "Residential Learning Communities," "Communities in Schools," and several others (see those coded as 99 in Figure 4).

Figure 3: *Intervention Reports* (Typically includes a body of studies on a named intervention) for the topical categories of *Pathways to Graduation* and *Postsecondary* as of December 2016

Topical Classification	Number of interventions in data base under the topic (note there is double listing)	Number of interventions listed with no studies meeting WWC methods standards	Number with at least 1 study meeting WWC methods standards with or without reservations	Number with at least one significant positive result	Percent of Interventions reviewed with no positive results (either no studies met the methods standards or met and no impact)
Pathway to Graduation	35	13	22	14	60%
Postsecondary	7	1	6	3	57%

NOTE: all of the interventions listed under the topic of Postsecondary are also listed under the topic of Pathway to Graduation. Interventions listed on WWC website as of December 2016.

Figure 4: Listing of the "Interventions" reviewed in What Works Clearinghouse under the categories of "Pathways to Graduation" and "Postsecondary": December 2016 (Some programs are double counted under both categories)

Name of Intervention	Year of Review	Methods rating: 99 = none met standards 1 = Met without reservation 2 = Met with reservation 3 = mixed of 1 and 2	Effectiveness rating: 3 =Positive 2=Potentially positive 1 = Mixed 0 = No discernable impact - 1= Potentially negative -2 =Negative 99= No studies met standards	Highest effect size reported	Number of studies meeting methods standards of those reviewed
Pathways to Graduation section (35 Interventions Reviewed)					
1. **ACT/SAT Test Preparation and Coaching Programs**	2016	1	3	8	6 of 27
2. **First Year Experience Courses for Students in Developmental Education**	2016	1	0	0	1 of 19
3. **Career Academies**	2015	1	1	11	1 of 9
4. **Check & Connect**	2015	1	1	30	2 of 3
5. **Credit Recovery Programs**	2015	99	99	99	NA
6. **Reconnecting Youth**	2015	99	99	99	NA
7. **Linked Learning Communities**	2015	1	0	0	4 of 16
8. **Residential Learning Communities**	2015	99	99	99	NA
9. **National Guard Youth ChalleNGe Program**	2015	1	2	23	1 of 4
10. **Service and Conservation Corps**	2010	2	0	0	1 of 1
11. **Communities in Schools**	2010	99	99	99	NA
12. **Youth Build**	2009	99	99	99	NA
13. **High School Puente Program**	2009	99	99	99	NA
14. **Coca-Cola Valued Youth Program**	2009	99	99	99	NA
15. **Summer Training and Education Program (STEP)**	2009	1	0	0	1 of 1
16. **I Have a Dream**	2009	99	99	99	NA
17. **Middle College High School**	2009	1	0	0	1 of 1
18. **Talent Development Middle**	2009	99	99	99	NA

Grades Program (TDMG)					
19. **Wyman Teen Outreach Program (TOP)**	2009	99	99	99	NA
20. **New Century High Schools**	2008	99	99	99	NA
21. **Accelerated middle schools**	2008	1	2	35	3 of 3
22. **Job Corps**	2008	1	1	13	1 of 1
23. **JOBSTART**	2008	1	2	14	1 of 1
24. **First Things First**	2008	1	0	0	1 of 3
25. **New Chance,**	2008	1	1	8	1 of 1
26. **Belief Academy**	2007	99	99	99	NA
27. **Project COFFEE**	2007	0	99	99	NA
28. **Project GRAD**	2007	1	0	0	1 of 3
29. **Quantum Opportunity Program**	2007	1	0	0	1 of 2
30. **Talent Development High Schools**	2007	1	2	0	1 of 2
31. **High School Redirection**	2007	1	1	3	3 of 3
32. **Twelve Together**	2007	1	1	13	1 of 2
33. **Financial Incentives for Teen Parents to Stay in School**	2007	1	1	6	2 of 2
34. **Talent Search**	2006	1	2	17	2 of 3
35. **ALAS**	2006	1	2	42	1 of 1
Postsecondary category (7 Interventions Reviewed)					
1. **ACT/SAT Test Preparation and Coaching Programs**	2016	1	3	8	6 of 27
2. **First Year Experience Courses**	2016	1	3	9	4 of 97
3. **Summer Bridge Programs**	2016	1	2	3	1 of 1
4. **First Year Experience Courses for Students in Developmental Education**	2016	1	0	0	1 of 19
5. **Developmental summer bridge programs**	2015	1	0	0	1 of 10
6. **Linked Learning Communities**	2014	1	0	0	4 of 16
7. **Residential Learning Communities**	2014	99	99	99	NA

The Category of "Single Studies" Reviews. Over time the WWC developed more emphasis on what they call" Single Study Reviews." Of 160 single studies under *Pathways to Graduation* for which methods were rated, there were 36 studies that met the methods standards, and 23 of the 160 also had a least one positive result—about 14 percent. Among those categorized as

Postsecondary, about 212 were reviewed for methods, and of these, 62 met the methods standards. Of the 212 studies, there were 44 that both met the methods standards and had at least one positive result (21 percent of total reviewed). Note that *Single Studies* like the *Interventions* may be doubly classified under *Postsecondary* and *Pathways to Graduation*. Considering all topics, there are about 2,349 total Single Studies that have had methods reviewed and are classified in the WWC data base.

A review of the single studies (both impactful and not impactful) reveals that many of the strategies examined in the *Single Study* reviews (and also the *Intervention* reviews) have similar conceptual frameworks and seemingly are implementing similar practices. This suggests that differential judgements as to whether there is positive impact may be more about the study structure, the context, and comparison group-counterfactual differences rather than the differences in the strategies. Additional work is needed to assess if there is a pattern of differences between those strategies that show positive impact and those that show no impact.

One also notices that the role of the WWC in the federal grant competitions, especially the TRIO postsecondary grants (SSS), Talent Search (TS), Upward Bound (UB) which have required use of *Single Study* reviews, have resulted in an increase in both the number and the percentage of studies in the WWC in these areas that have positive impacts. Thus the WWC's role of doing somewhat systematic reviews of an "Intervention" has evolved to focus on *Single Studies* and Practice Guides focusing on those studies with positive results.

The Category of "Practice Guides." Since its inception, the WWC has published 22 Practice Guides. Of these two are classified under *Pathways to Graduation* and one under *Postsecondary*. These are: *Helping Students Navigate the Path to College: What High Schools Can Do*, published in 2009; *Dropout Prevention* published in 2008; and *Strategies for Postsecondary Students in Developmental Education – A Practice Guide for College and University Administrators, Advisors, and Faculty* published in 2016. The 2016 guide lists six recommendations that range from minimal to moderate in evidence rankings by the WWC. The recommendations are as follows:

1. Use multiple measures to assess postsecondary readiness and place students.
2. Require or incentivize regular participation in enhanced advising activities
3. Offer students performance-based monetary incentives.
4. Compress or mainstream developmental education with course redesign.
5. Teach students how to become self-regulated learners.
6. Implement comprehensive, integrated, and long-lasting support programs

Figure 5: Results of Searching WWC *Single Study Review* Category under Path to Graduation and Postsecondary: December 2016								
Type of Report	**Number of Single Studies with methods rated**	**Number meeting methods standards with or without reservations**	**RCT**	**Quasi-Experimental**	**At least one positive RCT**	**At least one Quasi Experimental positive**	**At least one positive result**	**Percent that met standard and had at least one positive result**
All topics in WWC	2349	877 (37% with or without) --572 without --322 with	576	597	145	87	280	12%
Pathways to graduation	160	36 (29% with or without) --24 without; --15 with;	33	23	20	3	23	14%
Post-secondary	212	62 (29% with or without) --37 without; --25 with	38	24	26	17	44	21%
NOTE: Numbers do not sum to totals due to double listing and multiple methods								

The Special Case of the National Evaluation of Upward Bound. The 2009 Practice Guide entitled *Helping Students Navigate the Path to College: What High Schools Can Do,* prepared by Mathematica Policy Research, included a review of the then recently-released National Evaluation of Upward Bound final report (Sefter et.al. 2009) and also the Mathematica UB evaluation Third Follow up report released in 2004 (Myers et.al. 2004). The review ignored the fact that the original and final two technical monitors for the long-running study in the U.S. Department of Education had, after a Quality Assurance Review that involved both external and internal analysis of all the data files from the study, found that the Mathematica reports in 2004 and 2009 were seriously flawed in their conclusions that the UB program had no discernible impact on postsecondary entrance or degree attainment other than the attainment of vocational certificates. In contrast, the analysis done by the Department of Education technical monitors had found that Upward Bound had statistically significant and substantial positive impacts on postsecondary entrance, financial aid award, and degree attainment (Cahalan and Goodwin 2014; Cahalan, 2009). Despite replication of the technical monitors' findings by Nathan, 2013; and Harris, Nathan, Marksteiner, in 2014, the Practice Guide has not been corrected in its reporting

concerning the Upward Bound Program. The flawed Mathematica study has had a negative impact on Upward Bound program's reputation and identity as an effective college access program. The corresponding undeserved "ineffectual PART rating" led to zero funding requests in the FY05 and FY06 Bush Administration budget requests. This experience has led to a profound mistrust among the TRIO practitioners to contract evaluation research studies that is only now beginning to be overcome.

Using the 2015 Student Support Services (SSS) Competition as an Example: Preliminary Data on the Impact of the CPP Points in the Competition Scores

As the data above indicates, there are not a large number of strategies that meet the WWC methods standards, have positive results, and also are feasible within the legislatively-defined service and funding options of the SSS program. Moreover, those strategies that are applicable and have positive impacts often have relatively small effect sizes. For the 2015 SSS competition, the Department of Education identified two priority areas. The substantive areas in which the Department invited CPP Submissions were:

- Strategies to Influence the Development of Non-Cognitive Factors Supported by Moderate Evidence of Effectiveness (up to 3 points)
- Providing Individualized Counseling for Personal, Career, and Academic Matters Supported by Moderate Evidence of Effectiveness (up to 3 points)

As part of the SSS competition, the Department provided to applicants citations for three specific research articles applicable to the priority areas (See Figure 6) that the WWC had determined both met the methods criteria and had at least one positive result.

Figure 6: Relevant articles provided by the Department of Education to SSS applicants as already reviewed by the What Works Clearinghouse (WWC)

- **Social belongings intervention**: Walton, G. M. & Cohen, G. L. (2011). *A brief social-belonging intervention improves academic and health outcomes of minority students*. Science 331, 1447-1453 Online Available: http://www.sciencemag.org/content/331/6023/1447.full
- **Difference education intervention:** Stephens, N M., Hamedani, M G., & Destin (2014). *Closing the social-class achievement gap: A difference-education intervention improves first-generation students' academic performance and all students' college transition*. Psychological Science 1-11. Online Available: http://www.psychology.northwestern.edu/documents/destin-achievement.pdf
- **Individualized Counseling**-Inside Track: Bettinger, E. & Baker, R. (2011). *The effects of student coaching in college: An evaluation of a randomized experiment in student mentoring.* Online Available: https://ed.stanford.edu/sites/default/files/bettinger_baker_030711.pdf

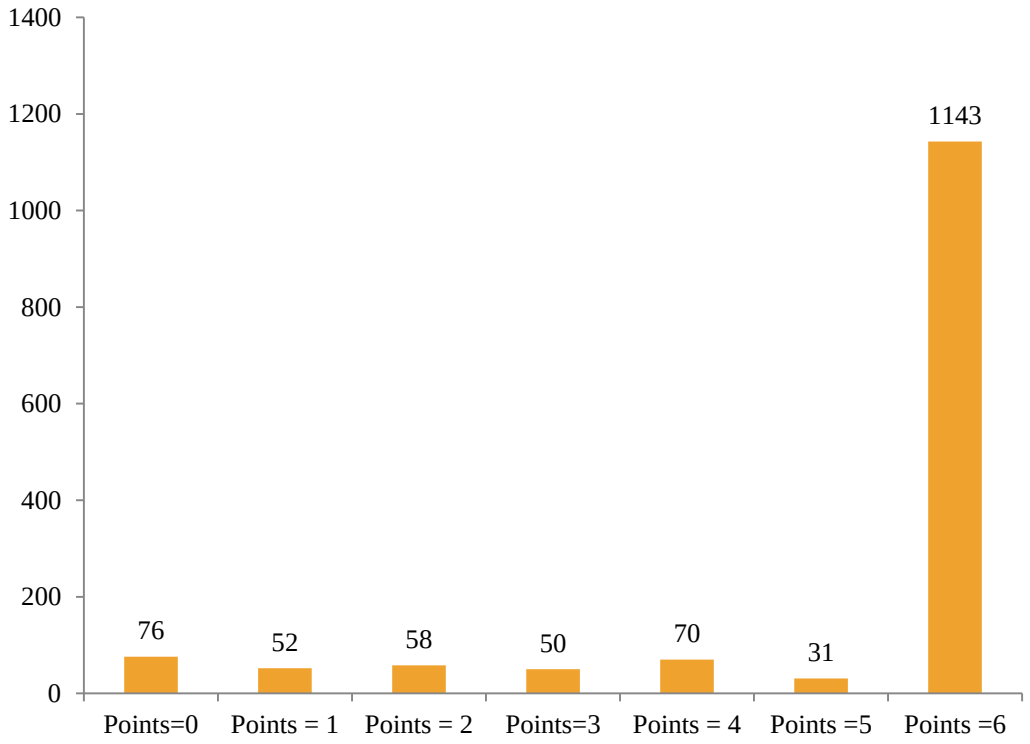
SOURCE: U.S. Department of Education Application for 2015 SSS Grant Cycle.

One of the two studies in the non-cognitive topic area was an experimental intervention that took about one hour to implement (Walton, G. M. & Cohen, G. L., 2011). The small sample from a selective university was 37 Blacks and 33 Whites. Modest positive impacts were observed for Blacks but not Whites when compared to a comparison group, which was given a one-hour session that did not involve the social belonging element. The other study in the "non-cognitive" priority area was an orientation experiment for first-generation students that involved about three hours (Stephens, N M., Hamedani, M G., & Destin, 2014).

The one study citation given by ED for the individual counseling priority area involved a randomized control trial with 3,527 students for an intervention of the use of a commercial on-line software tool for college coaching and mentoring –Inside Track (Bettinger, E. & Baker, R. (2011). This research found a 4 percentage point difference in 12-month persistence, but no differences in completion of a degree in 4 years. One might expect that in the case of SSS, the expected impact would depend on how different this was from what the projects were already doing.

SSS applicants were free to find and submit other articles within the priority areas that met the WWC criteria and had at least one positive result; however, out of some 1,500 applicants, only 17 applicants identified another intervention study that met the criteria. The competition for TRIO grants is very intense, and in order to be competitive, it is necessary to successfully address the competitive priorities. The data displayed in Exhibit 5 confirms that 95 percent of the applicants chose to address the CPPs and over three-fourths (77 percent) were able to obtain the full 6 points, which means that they proposed to implement the specific research-based programming in both of the areas (non-cognitive and student coaching). A review of their status as "new" or "previously funded" applicants indicates that most of the applicants who did not choose to address the CPPs were new applicants. Given SSS scoring, it was not possible for new applicants to win unless they addressed the CPPs. The cut-off scores each cycle depend on the amount of funding available relative to the number of applicants. For 2015, there were 1,480 applicants and 1,071 awards. The cut-off score was 104, meaning that a new applicant would need to get a perfect or near perfect score on the Project Design section (worth 100 points) and also obtain 4 of the 6 CPP points. In the TRIO competitions held every five years, existing projects may earn up to 15 extra points for meeting their objectives specified in the previous cycle years. These points are called Prior Experience (PE) and are tabulated by an independent contractor based on submitted performance reports. These points give existing projects an edge, and there is a high degree of project continuance. The experience of past competitions funding cycles, in which about 2/3 of applicants are funded and 1/3 are not funded, led COE to advise applicants to try for the CPP points even if they believed they would get most or all of the Prior Experience (PE) points for the cycle.

Figure 7: Total Number of SSS IHE Applicants Getting Various Numbers of CPP Points in 2015 Competition



NOTE: SSS Applicant Score File Prepared by Office of Federal TRIO Programs for this Submission January 2017.

Discussion of the Identified Limitations to the Use of the WWC and Reflections on Making the "Evidence Based Practice" Movement" more Useful

As discussed, the Department of Education has chosen the WWC Standards as the way to provide an authoritative independent judgment concerning what constitutes "research evidence" for implementation of its CPPs. *The use of WWC contractor ratings has the advantage of removing ED from technical decision making in this regard, and in having published standards available to all.* However, throughout its history, since the early 2000s, IES and the WWC have not been without criticism, and these limitations can be expected to influence the efforts at implementation of the increased evidence use in the federal education programs.

Below is a summary of some of the critiques of the WWC, followed by some ideas moving forward that might lead to a more meaningful use of research to foster program improvement.

- ***Favoring Research Capability over Need for Services.*** One of the criticisms of the CPPs for 2015, in the public comment period, was that the CPPs would promote undue preference to research institutions and put smaller, less resourced, rural, and minority-serving

institutions at a disadvantage in the grant-writing competition process. Over time, the larger IHEs with grant-writing staff increasingly submit grants written by professional grant writers other than the TRIO staff at the institution. For example, several of the larger IHE's now have multiple SSS grants--(Regular, STEM, Disabled, Veterans). Commenters questioned whether the Department was not putting "Research Capability" over the "Need for Services." Additional research is needed to observe if this was indeed the case.

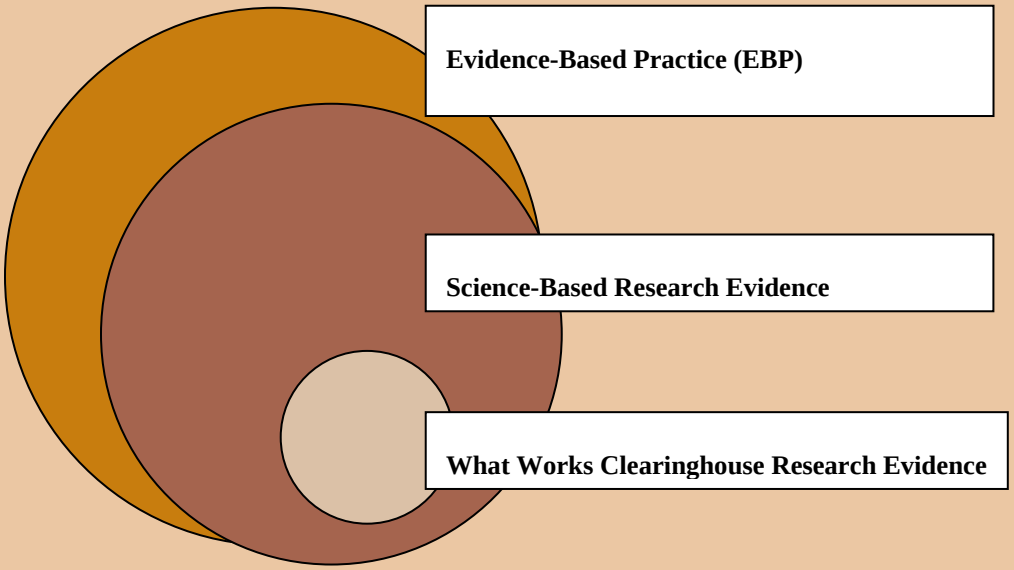
- ***Narrowness of WWC Compared to Wider Conceptualizations of Evidence to the Exclusion of Those Practices that Might be of Most Benefit.*** Another criticism of the WWC as the arbiter of which practices are to be promoted concerns the narrowness of the standards for what constitutes "evidence." The methods favored by WWC ignore a wide body of non-experimental observational science. Those of us who have attempted to implement ethical random assignment studies in the time-sensitive area of college access know that it is only feasible to implement a very narrowly-focused random assignment in which almost equally resourced interventions are offered and then it becomes very difficult to assess impact.

  We can distinguish three concepts relating to research evidence of which the WWC definition is the narrowest:

  1. WWC Standards Based Research Evidence;
  2. Science-Based Research Evidence, and
  3. Evidence-Based Practice (EBP).

Figure 8 illustrates our conceptualization of these three types of research based evidence in terms of the narrowness of the concepts. *Evidence-Based Practice (EBP)* is an encompassing concept that takes into account the "built up clinical knowledge," client interactions, implementation findings, as well as *Science-Based Research Evidence* (Soyden and Palinkas 2014). *Science-Based Research* is defined more broadly than the WWC ---*as research activity that employs systematic, empirical methods to address a specific question*. The *WWC Standards Based Definition* is the most narrowly defined of the three. As the tables above demonstrate, the WWC Standards Based determination of research evidence is comprised of far fewer studies. By the end of 2016, there were about 44 single studies in the WWC that were in the area of Postsecondary-- and that met the WWC Standards ---and that had at least one statistically significant positive finding.

Figure 8: Three Nested Concepts of Evidence



- ***Questions about Utility and Validity of Research Rankings of WWC.*** In the early days of the WWC most of the focus was on K-12 rather than postsecondary education. Hence, most of the critiques have addressed the ratings of the studies and conclusions with regard to the interventions at the K-12 level. Alan Ginsberg and Marshall Smith (2016) conducted a review of 27 RCT studies of math curricula in the WWC, and found that 26 of the 27 had serious threats to validity or usefulness other than that of the selection bias that the RCT method controls. Moreover, they note that the magnitude of the error generated by the threats was often greater than the average effect size of the RCT treatment. They recommended a panel of experts and users be established to consider how to improve the WWC criteria and standards for review.

- ***High Level of Changes in Ratings.*** Other critiques have noted the high rate of changes in WWC ratings following a request for a re-review of the findings, usually by study authors. For example, the National Institute for Direct Instruction (NIFDI) filed a Freedom of Information Act request for data on the results of the Requests for Reconsideration made to the WWC. NIFDI found that of the 69 requests reviewed 54 resulted in changes in ratings. The NIFDI reviewers also pointed to the high rate of exclusion of research studies as not meeting the WWC criteria. NIFDI has been among the most severe of the critiques of the WWC, especially related to the WWCs conclusions with regard to their highly studied reading curriculum. Siegfried Engelmann, founder of Direct Instruction, notes: "*Unfortunately, the WWC has failed to live up to its promise. The WWC's reports promote curricula that the scientific community has found to be ineffective and inefficient and denigrate those that the scientific community has found to be highly effective*."

Englemann concluded that the WWC was so "irreparably biased that it would have to be thoroughly reoriented and reorganized under different management rules to perform the function of providing reliable, accurate information about what works."
http://www.nifdi.org/research/reviews-of-di/what-works-clearinghouse.

After filing the FOIA request from the WWC, NIFDI researcher Wood (2014) summarized the issues:

> *With the information provided from the FOIA request and the publicly available information, three conclusions appear clear: 1) The WWC suffers from a lack of transparency in their policies and guidelines, 2) the conclusions they create in their reports can be misleading, and 3) the reports are potentially damaging to program developers and ultimately the success of students (Wood, 2014).*

- ***Small Percentage of Studies Meeting WWC Methods Criteria and Low Effect Sizes.*** A recent review (Malouf D. and Taymans 2016) of the total WWC "Interventions" reviewed points out (as have others) the small percentage of studies meeting the methods criteria. They also noted the small effect sizes of those studies that were listed as meeting criteria and having positive results. The authors concluded:

  > *Most interventions were found to have little or no support from technically adequate research studies, and intervention effect sizes were of questionable magnitude to meet education policy goals. These findings painted a dim picture of evidence based on education interventions and indicated the need for new approaches including a reexaminiation of federal reliance on experimental impact research as the basis for gauging intervention effectiveness (Malouf D. and Taymans, 2016).*

- ***Lack of Concern for the Differences in "Treatment Contrast."*** Other researchers, outside and inside the context of WWC criticism, have noted the importance of the "treatment contrast" and point out that the differential results of RCTs may often have more to do with how the study is structured and different counterfactuals than with the differences in the interventions. These critiques stress the importance of measurement of both the treatment and control groups' level of receipt of the intervention or its equivalents (Weiss, Brock, and Bloom, 2013). The "business as usual" practices for comparison may also include substantial practices that are similar to the interventions being studied (Heckman, 2002). To understand the treatment contrast a researcher must examine the intervention related experiences of both treatment and control group members.

- ***Unwise Linking of Funding Decisions to Research Findings.*** Political rhetoric and policy of the last 40 years has tended to support the linking of funding decisions to evaluation outcomes. This is a dangerous policy that risks failure to serve the most needy and vulnerable persons, and also mediates against honest and fair evaluations that lead to improved services. As Haskins and Margolis (2014) point out:

> *In our view, an important part of a comprehensive, evidence-based strategy will be continuing the funding of programs with initially discouraging evaluations. Part of the federal evidence-based culture should be that federal agencies will work with programs, and continue their funding, as long as they are using evidence to improve their outcomes and are showing some progress (Haskins & Margolis, 2014).*

- ***Promotion of Formulaic Implementation of Handful of Superficial, Standardized One Size Fits All Strategies.*** As the review of single studies that have both met the WWC standards requirements and also have achieved some statistical significance revealed, there is a limited range of strategies that can be studied with an RCT design that does not involve IRB-prohibited denial of services. The tendency for most TRIO programs to attempt to replicate the same studies can lead to a less creative response to changing situations and co-learning that takes into account the differing situations and needs of students. There is insufficient attention to what works and for whom and under what circumstances (Imai and Ratkovic, 2013; Weis, Bloom, and Brock, 2013; Gamoran 2014).

- ***Demonstrating the Need for Research over Service Needs.*** Related to the above comments, practitioners fear that their service programs will be turned into permanent pilot demonstration grants, and that those sites that are not able to do this or do not wish to do this, will gradually be competed out of service. Researchers like to cite lack of positive findings as a justification for increased need for more research projects and may have a house bias towards study designs unlikely to produce positive outcomes. Researchers such as Jon Barron of the Coalition for Evidence Policy states that: "The better the evaluations, the more likely they are to show that some programs do not produce significant impacts." Baron goes on to report that, out of 90 interventions evaluated by randomized control trials (RCTs) paid for by IES since 2002, 88 percent were found to have weak or no positive effects. Similar results are produced by RCTs of clinical interventions in medicine (Coalition for Evidence-Based Policy).

- ***Danger of Favoring of One Intervention over Another with Sparse Evidence.*** There are serious issues in favoring one type of intervention based on a single study that does not usually make precise comparisons across strategies or programs. The counterfactual is typically simply "business as usual," which can mean a variety of unmeasured contrasts. There could also be potential conflicts with the authorizing legislation for the federal programs which allow a wide range of services. Potentially an unfair advantage is given to those projects for which the "favored" intervention is appropriate or even feasible.

Ideas Moving Forward

Given these criticisms, the following are six ideas concerning use of evidence moving forward.

1. ***Decouple Funding Decisions From Results of the Evaluations.*** This coupling discourages innovation and thoughtful implementation and mediates against on-going learning from evaluations.

2. ***Address the Serious Validity and Reliability Concerns That Have Been Raised Concerning Validity of WWC Results.*** The WWC provides a formulaic way to address OMB requirements, because it seemingly is clear and definitive, but like the standardized tests, it is unclear that its overall impact is positive in terms of best practice in education.

3. ***Align the WWC Standards with Other National Standards,*** such as those of the Joint Committee on Education Program Evaluation Standards; NCES Statistical Standards, and AERA Standards for Research Publications.

4. ***Broaden the Definition of Research Evidence*** to include Evidence-Based Practice (EBP) and Science-Based Research Evidence and invite thoughtful consideration of how best to serve students. Consider meta-analyses and weight of evidence, more nuanced approaches.

5. ***Consider the Implications for Students if Most Projects Implement the Same "Cookie Cutter" Strategies*** regardless of context and the range of services available at the institutions. Consider the students' real needs and the program's own "niche" in the institution or the community.

6. ***Consider New Evaluation Methods Such As Empowerment, Participatory and Collaborative Evaluation*** (Federman et. al. 1996; Chinman et.al, 2004; Chinman et.al. 2008, CDC Cox et.al., 2010). Consider methods that stress a deeper understanding of context and researcher-practitioner-client interactions. As A. Gamoran (2014) notes:

> *To date, many rigorous studies treat programs as if they were black boxes, seeking a positive or negative judgment without aiming to understand how the outcomes are reached. The next generation of policy research in education will advance if it offers more evidence on mechanisms so that the key elements of programs can be supported, and the key problems in programs that fail to reach their goals can be repaired (Gamoran, 2014).*

REFERENCES

Ahmad, A.R., Adi, M., Nazir, M., Md Noor, H., Ng, K.S., & Azizan, A.Z. (2013). *An assessment of agency theory in the context of higher education institutions.*

Bailey, T., Bashford, J., Boatman, A., Squires, J., Weiss, M., Doyle, W., Valen-tine, J.C., LaSota, R., Polanin, J. R., Spinney, E., Wilson, W., Yeide, M., & Young, S.H. (2016). *Strategies for postsecondary students in developmental education – A practice guide for college and university administrators, advisors, and faculty.* Washington, DC: Institute of Education Sciences, What Works Clearinghouse.

Bettinger, E.P., & Baker, R. (2011). *The effects of student coaching in college: An evaluation of a randomized experiment in student mentoring*. Retrieved from https://cepa.stanford.edu/sites/default/files/bettinger_baker_030711.pdf

Bourdieu, P. (1977). *Outline of a Theory of Practice*. Cambridge University Press: London.

Bourdieu, P. (1990). *The Logic of Practice*. Polity Press: Cambridge.

Bourdieu, P., & Wacquant, L.J.D. (1992). *An Invitation to Reflexive Sociology*. University of Chicago Press: Chicago.

Cahalan, M. (2009). *Addressing Study Error in the Random Assignment National Evaluation of Upward Bound: Do the Conclusions Change?* Washington, D.C.: The Pell Institute for the Study of Opportunity in Education. Retrieved from http://www.pellinstitute.org/publications-Do_the_Conclusions_ Change_2009.shtml.

Cahalan, M., & Goodwin, D. (2014). *Setting the Record Straight: Strong Positive Impacts Found from the National Evaluation of Upward Bound, Re-Analysis Documents Significant Positive Impacts Masked by Errors in Flawed Contractor Reports*. Washington, D.C. : The Pell Institute for the Study of Opportunity in Higher Education. Retrieved from http://www.pellinstitute.org/downloads/ publications Setting_the_Record_Straight_June_2014.pdf

Chinman, M., Hunter, S.B., Ebener, P., Paddock, S., Stillman, L., Imm, P., & Wandersman, A. (2008). The getting to outcomes demonstration and evaluation: An illustration of the prevention support system. *American Journal of Community Psychology*. *41*: 306–224. doi:10.1007/s10464-008-9163-2

Chinman, M., Imm, P., & Wandersman, A. (2004). *Getting To Outcomes: Promoting Accountability Through Methods and Tools for Planning, Implementation and Evaluation.* Santa Monica, CA: RAND Corporation.

Cox, P.J., Keener, D., Woodard, T., & Wandersman, A. (2010). *Evaluation for Improvement: A seven step empowerment evaluation approach for violence.* Center for Disease Control (CDC). Retrieved from https://www.cdc.gov/violenceprevention/pdf/evaluation_improvement-a.pdf

Crosse, S., Williams, B., Hagen, C.A., Harmon, M., Ristow, L., DiGaetano, R., Broene, P., Alexander, D., Tseng, M., & Derzon, J.H. (2011). *Prevalence and implementation fidelity of research-based prevention programs in public schools: Final report.* U.S. Department of Education. Retrieved from https://www2.ed.gov/rschstat/eval/other/research-based-prevention.pdf

Department of Education: Applications for new awards; Student Support Services Program," 79 Federal Register, 243, (December 18, 2014), p. 75723. Retrieved from http://www.gpo.gov/fdsys/pkg/FR-2014-12-18/pdf/2014-29769.pdf

Donaldson, S., Christie, C., & Mark, M. (Eds.). (2015). *Credible and Actionable Evidence, 2nd Edition.* Thousand Oaks: Sage Publications.

Eisenhardt, K.M. (1989). Agency theory: An assessment and review. *Academy of management review, 14*(1), 57-74.

Eisenhardt, K.M. (1989). *Building theories from case study research.* The Academy of Management Review; Oct 1989; 14, 4; ABI/INFORM Global

Epstein, D., & Klerman, J.A. (2013). When is a program ready for rigorous impact evaluation? The role of falsifiable logic model. *Evaluation Review, 36* (5), 375-401.

Engelmann, S. (2013). Machinations of What Works Clearinghouse. Retrieved from http://eppi.ioe.ac.uk/cms/ The Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI-Centre) is part of the Social Science Research Unit at the UCL Institute of Education.

Fetterman, D., Kaftarian, S.J., & Wandersman, A. (1996). *Empowerment Evaluation: Knowledge and Tools for Self-Assessment and Accountability*. Thousand Oaks: Sage Publications.

Ginsberg, A., & Smith, M. (2016). *Do Randomized Controlled Trials meet the "Gold Standard"? A study of the usefulness of RCTs in the What Works Clearinghouse*. Washington, D.C.: American Enterprise Institute.

Gamoran, A. (2013). *Educational inequality in the wake of No Child Left Behind.* William T. Grant Foundation Spencer Foundation Award Lecture to the Association for Public Policy and Management, November 7, 2013.

Hallfors, D., & Godette, D. (2002). Will the "Principles of Effectiveness" improve prevention practice? Early findings from a diffusion study. *Health Education Research: Theory & Practice, 17* (4), 461-470.

Haskins, R., & Margolis, G. (2014). Obama's promise: Using evidence to fight the nation's social problems. Finnigan, K S. & Daly, A J. (Eds). Using research evidence in education: From the schoolhouse door to Capitol Hill. New York: Springer. 117-132

Hatry, H. (1967). Report of the Senate Committee on Measures of Effectiveness for Social Programs. Retrieved from https://www.jstor.org/stable/244890?seq=1#page_scan_tab_contents

Heckman, J., Hohmann, N., Smith, J., & Khoo, M. (2000). Substitution and dropout bias in social experiments: A study of an influential social experiment. *The Quarterly Journal of Economics*, May 2000.

Honing, M.I. (n.d.). Complexity and policy implementation: Challenges and opportunities for the field. Honing, M. I. (Ed). New directions in education policy implementation: Confronting complexity. pp. 1-24. Albany: State University of New York Press.

Kivistö, J. (2007). *Agency theory as a framework for the government-university relationship*: Tampere University Press.

Kivistö, J. (2008). An assessment of agency theory as a framework for the government–university relationship. *Journal of Higher Education Policy and Management, 30*(4), 339-350.

Laura and John Arnold Foundation. (2016*). Demonstrating the power of evidence-based programs to "move the needle" on major U.S. social programs: A competition & funding opportunity*. Retrieved from http://www.arnoldfoundation.org/rfp-demonstrating-power-evidence-based-programs-move-needle-major-u-s-social-problems/

Langer, L., Tripney, J., & Gough, D. (2016). *The Science of Using Science: Researching the Use of Research Evidence in Decision-Making*. London: EPPI-Centre, Social Science Research Unit,UCL Institute of Education, University College London. ISBN: 978-1-907345-88-3

Liefner, I. (2003). Funding, resource allocation, and performance in higher education systems. *Higher Education, 46*(4), 469-489

Leake, M., & Lesik, S. (2007). Do remedial English programs impact first-year success in college? An illustration of the regression-discontinuity design. *International Journal of Research and Method in Education, 30*(1), 89-99.

Lesik, S. (2011). New Perspectives in Applying the Regression-Discontinuity Design for Program Evaluation: A Simulation Analysis. *Journal of Modern Applied Statistical Methods, 10*(1), 198-206.

Lesik, S. (2007). Do developmental mathematics programs have a causal impact on student retention? An application of discrete-time survival and regression-discontinuity analysis. *Research in Higher Education, 48*(5), 583-608.

Mayne, J. (2010). Building an evaluative culture: The key to effective evaluation and results management. *Canadian Journal of Program Evaluation, 24*(2), 1-30.

Luyten, H. (2006). An empirical assessment of the absolute effect of schooling: regression discontinuity applied to TIMSS-95. *Oxford Review of Education, 32*(3), 397-429.

Metz, A. J.R., Blasé, K., & Bowie, L. (2007). *Implementing evidence-based practices: Six drivers of success.* Retrieved from http://www.childtrends.org/wp-content/uploads/2013/07/2007-29EVPSuccess.pdf

Myers, D., Olsen, R., Seftor, N., Young, J., & Tuttle, C. (2004). *The impacts of regular Upward Bound: Results from the third follow-up data collection.* Report submitted to the U.S. Department of Education. Washington, DC: Mathematica Policy Research, Inc.

Malouf, D., & Taymans, J. (2016). Anatomy of an evidence base. *Educational Researcher, 45*(8), 454-459

National Center for Research in Policy and Practice. (2016). Survey of practitioners' use of research. National Research Council. (2012). *Using Science as Evidence in Public Policy.* Committee on the Use of Social Science Knowledge in Public Policy, K. Prewitt, T.A. Schwandt, and M.L. Straf, Editors.

Notice of Final Supplemental Priorities and Definitions for Discretionary Grant Programs, published in the Federal Register on December 10, 2014 (79 FR 73426) and from 34 CFR 77.1. 34 CFR 77.1

Newcomer, K., & Brass, C. (2016). Forging a strategic and comprehensive approach to evaluation within public and nonprofit organizations: integrating measurement and analytics within evaluation. *American Journal of Evaluation, 37* (1), 80-99.

Newcomer, K. (2016). *Evidence-based interventions and programs: What does this mean for us?* Presentation to Council for Opportunity in Education Annual Conference, San Diego California.

National Center for Research in Policy and Practice. (2016). *Survey of Practitioners' Use of Research.* Retrieved from http://ncrpp.org/assets/documents/NCRPP_SPUR-Instrument.pdf

Olejniczak, K., Raimondo, E., & Kupiec, T. (2016). Evaluation units as knowledge brokers: testing and calibrating an innovative framework. *Evaluation, 22* (2), 168-189.

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

Office of Management and Budget. (2016). Building and using evidence to improve results: Meeting our greatest Challenges: The President's Fiscal year 2017 budget. Retrieved from

https://www.whitehouse.gov/sites/default/files/omb/budget/fy2017/assets/fact_sheets/Building%20and%20Using%20Evidence.pdf

Office of Management and Budget. (May 18, 2012). Memorandum on 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

Office of Management and Budget .(July 26, 2013). Memorandum to the heads of departments and agencies. Next Steps in the Evidence and Innovation Agenda. Retrieved from https://www.whitehouse.gov/sites/default/files/omb/memoranda/2013/m-13-17.pdf

Palinkas, L., Garcia, A.R., Aarons, G.A., Finno-Velasquez, M., Holloway, I.W., Mackie, T.I., & Chamberlain, P. (2014). *Measuring Use of Research Evidence: The Structured Interview for Evidence Use, Research on Social Work Practice*. Thousand Oaks: Sage Publications.

Palinkas, L.A., Garcia, A.R., Aarons, G.A., Finno-Velasquez, M., Holloway, I.A., Mackie, T.I., Leslie, L.K., & Chamberlain, P. (2015). Measuring use of research evidence: The structured interview for evidence use. *Research on Social Work Practice*, 1-15.

Ross, S. (1973). The economic theory of agency: the principals' problem. *American Economic Review*, 63, 134-130. Results for America. (2015). Federal evidence-based innovation programs. Retrieved from http://results4america.org/wpcontent/uploads/2015/10/Innovation-fact-sheet.pdf

Seftor, N.S., Arif, M., & Schrim, A. (2009). *The Impacts of regular Upward Bound on postsecondary outcomes 7-9 Years after scheduled high school graduation.* Report submitted to the U.S. Department of Education. Washington, DC: Mathematica Policy Research, Inc.

Seastrom, M. *NCES Statistical Standards* (NCES 2003–601). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office, 2002.

Shadish, W.R., Cook, T.D., & Campbell, D.T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Cengage Learning.

Stephens, N.M., Hamedani, M.G., & Destin, M. (2014). *Closing the social-class achievement gap: A difference-education intervention improves first-generation students' academic performance and all students' college transition.* Psychological Science. Retrieved from http://www.psychology.northwestern.edu/documents/destin-achievement.pdf

Stockard, J., & Wood, T.W. (2013). Does the What Works Clearinghouse Work? A NIFDI White Paper National Institute for Direct Instruction and University of Oregon and National Institute for Direct Instruction.

Soydan, H., & Palinkas, L.A. (2014). *Evidence-Based Practice in Social Work: Development of a New Professional Culture.* New York, NY: Routledge.

Supplee, L.H., & Metz, A. (2015). Opportunities and challenges in evidence-based social policy. *Social Policy Report 28*(4). Retrieved from http://www.srcd.org/sites/default/files/documents/spr_28_4.pdf

Sutherland, W.J., Bellingan, L., Bellingham, J.R., Blackstock, J.J., Bloomfield, R.M., Bravo, M., et al. (2012). A collaboratively-derived science-policy research agenda. *PLoS ONE 7(3):* e31824. doi:10.1371/journal.pone.0031824.

Tseng, V. (2013). *Studying the use of research evidence in policy and practice.* William T Grant Foundation. Retrieved from http://wtgrantfoundation.org/library/uploads/2015/09/Studying-the-Use-of-Research-Evidence-in-Policy-and-Practice.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

Walton, G.M., & Cohen, G.L. (2011). A brief social-belonging intervention improves academic and health outcomes of minority students. *Science, 331*, 1447-1451. http://web.stanford.edu/∼gwalton/home/Publications_files/Walton_Cohen_2011_Science_1.pdf

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

Weiss, M.J., Howard, S.B., & Brock, T. (2013). *A conceptual framework for studying the sources of variation in program effects.* MDRC Working Papers on Research Methodology

Wood, T.W. (2014). Examining the inaccuracies and mystifying policies and standards of the What Works Clearinghouse: Findings from a Freedom of Information Act Request. Technical Report 2014-15. National Institute for Direct Instruction. Retrieved from http://www.nifdi.org/docman/research/technical-reports/1241-2014-5-examining-the-inaccuracies-and-mystifying-policies-and-standards-of-the-what-works-clearinghouse-findings-from-a-foia-request/file

About the Author:

Dr. Margaret Cahalan is the Vice President for Research and Director of the Pell Institute for the Study of Opportunity in Higher Education, of the Council for Opportunity in Education (COE). Over a 30 year career she has directed numerous large sample surveys and evaluation studies. After working at Westat, Mathematica Policy Research and RTI, she joined the U.S. Department of Education from 2004 to 2011. In this role she served as the Leader for the Secondary, Postsecondary and Cross Cutting Division of the Policy and Program Studies Services (PPSS).

Contact Information: MARGARET CAHALAN, DIRECTOR VICE PRESIDENT FOR RESEARCH, The Pell Institute for the Study of Opportunity in Higher Education, margaret.cahalan@pellinstitute.org