

**THE COALITION FOR EVIDENCE-BASED POLICY:
ITS ROLE IN ADVANCING EVIDENCE-BASED REFORM, 2004-2009**

Submitted to

The William T. Grant Foundation

**By
Monica Herk, PhD**

October 5, 2009

EXECUTIVE SUMMARY

This study attempted to assess the Coalition for Evidence-Based Policy's role in advancing evidence-based reforms in the five years since 2004 through semi-structured interviews with 33 policy-makers at the federal, state, and local levels. Over the past five years, the Coalition has successfully influenced legislative language, increased funding for evidence-based evaluations and programs, helped shape the Office of Management and Budget's Program Assessment Rating Tool [PART], and raised the level of debate in the policy process regarding standards of evidence. The Coalition has established a generally positive reputation as a rigorous, responsive, honest, and impartial advocate for evidence-based approaches, primarily at the federal level. Respondents saw the Coalition's ability to explain difficult concepts and its energy and action-orientation as additional strengths. Asked to name weaknesses, interviewees spoke of a possible need to expand the Coalition's focus beyond randomized controlled trials [RCTs], a need for improved communications, and possibly a need for a better understanding of the political and policy context in which decisions occur. The Coalition's accomplishments over the past five years leave the Coalition and its supporters in the positive position of deciding how best to build on its legacy of success.

PURPOSE AND METHODOLOGY OF THIS STUDY

The purpose of this study is to provide the philanthropic foundations that fund the Coalition for Evidence-Based Policy with an independent assessment of the Coalition's role in advancing evidence-based policy reforms in the five years since the last assessment was conducted in 2004.¹ In keeping with that aim, the Coalition contracted with Dr. Monica Herk, an independent consultant, to assess the Coalition's role in advancing such reforms and to present the results in a concise report emphasizing how the Coalition might improve what it does. Dr. Herk was instructed to conduct the evaluation independently and to submit her final report directly to the funders with a copy to the Coalition.

Consistent with those aims, Dr. Herk used the methodology of the 2004 evaluation. Between July 16 and September 10, 2009 she conducted semi-structured interviews with 33 individuals using a standard set of questions, which appear in Appendix A. Interviews averaged 30 minutes and were mostly conducted by phone, though 15% occurred in person.

The Sample

The 33 individuals interviewed represent a variety of policy-makers, as shown in the table below, who have either worked with the Coalition since 2004 or are familiar with its work. Initially the Coalition provided Dr. Herk with a list of 40 individuals whom the Coalition felt would be illuminating to interview. Dr. Herk was able to arrange interviews with 30 of those listed. One of those 30 turned out to be critical of the Coalition's

1. Bernard H. Martin, "The Coalition for Evidence-Based Policy: Its Impact on Policy and Practice," submitted to the William T. Grant Foundation, July 29, 2004. <http://coalition4evidence.org/wordpress/wp-content/uploads/2009/05/independent-evaluation-of-wt-grant-6-09.pdf>

work. In the interest of more fully understanding this perspective, Dr. Herk interviewed an additional three people who have been publicly critical of the Coalition’s activities.

| | n | % | n | % |
|--------------------|----|-----|----|------|
| Federal employees: | | | 22 | 67% |
| Executive branch | 17 | 52% | | |
| Legislative branch | 5 | 15% | | |
| All others: | | | 11 | 33% |
| Non-profits* | 4 | 12% | | |
| Universities | 3 | 9% | | |
| State government | 2 | 6% | | |
| Foundations | 2 | 6% | | |
| TOTAL: | | | 33 | 100% |

* Excluding universities and foundations.

Given the nature of the sample, it is not representative of any particular group. Nor can it be used to determine the relative frequency of supporters and critics among policy-makers who are familiar with the Coalition’s work. Answering that question would require a representative survey of policy-makers, and given the relatively low name recognition of the Coalition, such a survey would be costly to conduct. The assessment also does not touch on the attitudes of policy-makers that the Coalition might wish to influence but has not yet worked with. Rather this study is best thought of as an in-depth assessment of the opinions of the Coalition’s current “customers”, as well as a more limited examination of some voices critical of the Coalition.

Decision-Making Latitude of the Sample

Given the sample’s nature as a set of the Coalition’s “customers”, one early question was the amount of actual decision-making power that these individuals wield: that is, assuming that the Coalition has influenced the interviewees in some way, is the Coalition’s leverage primarily with policy-makers with a significant amount of power or merely with individuals having limited autonomy? Most respondents were asked, “On a scale of 1 to 5, where 1 is very little latitude and 5 is complete latitude, how much latitude would you say you have to make... decisions or is the decision mostly in the hands of your superiors?”² A common response was, “It depends,” but among the 16 respondents that actually provided a rating, the average score was 4.0, indicating that these are individuals who are quite advanced in their careers and who have a good deal of decision-making autonomy. Interviewees often explained that the final decision rested with their superiors, but that their input or advice was rarely overridden.

2. 23 out of the 33 respondents were asked this question. The question was omitted from truncated interviews, which occurred with the busiest respondents, who tended also to be the most senior members of the sample. The question was also omitted from interviews where it did not make sense in the context of the interview; for example, in interviews of methodologists or university researchers who do not make policy decisions.

How Members of the Sample Approach Decision-Making and “Evidence”

Another initial question concerned how members of the sample approach decision-making and how important a role evidence, of the type that the Coalition promotes, plays in those decisions. For example, for these decision-makers, what is the relative importance of evidence versus politics? The Coalition might be strongly shaping policy-makers’ views on evidence, and yet if evidence matters little in the policy-making process, the net impact will be small.

Two sets of questions attempted to tap this issue. The first set asked the respondents what considerations they thought about when making a program or policy decision. Interviewees were also asked to score the characteristics they named on a scale of 1 to 5, where 1 was not important at all and 5 was extremely important. The second set of questions focused on the issue of evidence by asking, “When you’re thinking about program or policy effectiveness, what types of evidence do you consider?” For both sets of questions, various follow-up prompts were used to elicit whether the interviewee considered certain things (e.g., politics) or types of evidence (e.g., RCTs), if the respondent did not mention them in their initial answer.

Of the 25 respondents asked these questions,³ the most common consideration cited when making a program or policy decision was along the lines of “evidence of effectiveness” or “most likely to achieve the program’s objectives”, with 60% of the 25 giving this type of response. The second most common response was “the political environment”, “politics”, “the administration’s priorities” or “Congressional priorities”, with 44% of the 25 giving this type of answer. Thirty-six percent of those asked cited issues such as “feasibility”, “ease of implementation”, and “replicability” of the program. Another common response was “cost”, “budgetary consequences”, or “budgetary limits”, with 32% of the 25 giving this answer. Finally, 28% of the 25 respondents asked this question cited “cost-benefit analysis,” “cost effectiveness”, or “the best return on investment”. Interestingly four out of the seven respondents giving this last answer worked at the state or local level, rather than in the federal government. The average scores given to the importance of these considerations as well as a fuller list of responses is given in Appendix B.

When asked about the types of evidence they considered when thinking about program or policy effectiveness, 68% of the 22 respondents who were asked this question said they used RCTs when available, although most qualified their response by saying that RCTs were only rarely available. Almost a quarter mentioned considering other, non-RCT evaluations either in place of or in addition to RCT studies. No one volunteered that they looked at programs in other states and localities, but 36% of the 22 asked responded positively to this question when prompted. Although only 9% volunteered that they used cost-benefit analyses when thinking about program effectiveness, an additional 68% said, when prompted, that they did. Almost a third of the 22 interviewees who were asked about the types of evidence they considered either volunteered that they used academic research or a literature review, or responded positively when prompted.

3. As with the question about decision-making latitude, some interviewees were not asked this question, either for reasons of time or because it did not make sense in the context of the interview.

So, to answer the questions posed at the beginning of the section, the policy-makers in the study rate evidence of effectiveness quite highly when making program or policy decisions – even more highly than political considerations. (This is perhaps not surprising given that it is a sample of policy-makers who have been responsive to the Coalition’s work.) When looking for evidence of effectiveness, the sample uses and generally prefers RCT and cost-benefit evidence when they are available, but relevant studies only rarely exist.

THE EFFECT OF THE COALITION’S WORK SINCE 2004

As the Coalition itself would point out, it can be difficult to determine an organization’s impact in the absence of a controlled experiment. Since we cannot compare our world with a world in which the Coalition does not exist, this assessment falls back on other approaches to estimate the effect of the Coalition’s activities.

The Coalition claims a number of accomplishments on its website.⁴ Since the interviewees, including both supporters and critics of the Coalition, volunteered the same accomplishments and credited the Coalition with a causal role, it seems reasonable to attribute these duly corroborated accomplishments to the Coalition.

- ***Work with the Office of Management & Budget [OMB], especially the Program Assessment Rating Tool [PART]***

A third of the respondents, both in and out of OMB, spoke of the Coalition’s impact on the agency, and particularly on the PART, which was put in place in 2004 (although supporters and critics of the Coalition diverged on whether the Coalition’s effect on the PART was positive or negative). A number of interviewees cited the Coalition’s influence on OMB’s general thinking about evidence, with one saying, “The Coalition’s major impact has been in getting OMB to acknowledge the importance of variations in standards of evidence.” More recently, OMB Director Peter Orszag positively cited the Coalition’s work in his blog.⁵ Other respondents pointed to the Coalition’s help to OMB in embedding rigorous evaluation requirements into the solicitation for projects under the American Recovery and Reinvestment Act of 2009 as additional evidence of the Coalition’s impact on OMB’s thinking.

- ***Increased spending on rigorous research at the Department of Education [ED]***

A number of respondents mentioned the Coalition’s impact on research at the federal Department of Education, although there was a little confusion concerning the Coalition’s precise role. Having explored the issue in greater depth, it seems accurate to say that the Coalition assisted the Institute for Education Sciences [IES] in increasing its research budget one year, helped to prevent a significant cut in IES’s research funding in another year, assisted in having certain research funds and programs housed in IES rather than elsewhere in the Department, and

4. http://coalition4evidence.org/wordpress/?page_id=12

5. <http://www.whitehouse.gov/omb/blog/09/06/08/BuildingRigorousEvidencetoDrivePolicy/>. Hyperlink at “smarter” links to the Coalition’s website.

in general, helped IES by having language supportive of RCTs placed in the Senate and House Appropriations Committee reports over several years.⁶

- ***Impact on other legislative language and appropriations***

Interviewees cited the Coalition's role in the successful inclusion of legislative language that will either increase funding for rigorous evaluations, create a preference for such evaluations or both. These include:

- A requirement that certain World Bank projects be rigorously evaluated with a preference for RCT designs where appropriate and feasible.⁷
- A 2% set-aside for rigorous evaluations of prisoner re-entry programs as part of the Second Chance Act.⁸
- \$110 million in funding for evidence-based teen pregnancy prevention (proposed).
- Evaluation funding in the National and Community Service appropriation (proposed).
- Increased evaluation funding for ED and the Department of Labor (proposed).

- ***Impact on the expansion of the federal home visitation program***

A number of respondents spoke of the Coalition's impact on the pending expansion of the federal home visitation program. Original drafts of the proposal specified that only the Nurse Family Partnership would be funded. Other home visitation programs, including ones that used paraprofessionals in place of nurses, lobbied for a wider range of programs to be included. The Coalition helped make the case on behalf of programs that had been evaluated using strong research methodologies and contributed to the current compromise, which provides preferential funding for programs that meet this higher evidentiary standard.

- ***Work outside Washington, DC***

One notable change since the Coalition's last assessment in 2004 is that the Coalition has established a greater presence in the states/localities, a presence that was almost non-existent at the time of the 2004 report. Although this presence is currently limited to approximately six states or localities, it does represent an expansion of the Coalition's scope. Policy-makers interviewed in these areas generally could not yet point to concrete impacts of the Coalition in their states, with some exceptions. In Rochester, New York, Jeff Kaczorowski, M.D., Executive Director of The Children's Agenda, says the Coalition played an important role in helping his organization advocate for public funding for early childhood programs and for the initiation and expansion of the Nurse Family Partnership by promoting the strong research base behind these interventions. In other states and localities, those who had worked with the Coalition were appreciative of the Coalition's efforts to help them raise the level of dialogue in their states around evidence. For example, in North Carolina the Alliance for

6. E-mail and phone conversations with Grover "Russ" Whitehurst and Jon Baron respectively on 9/21/09.

7. Public Law 111-32, June 24, 2009.

8. Public Law 110-199, April 9, 2008.

Evidence-Based Family Strengthening is a coalition of public and private funders that is using the Coalition's guidelines for evidence-based approaches.⁹

- ***More general impact on the policy process and debate***
Although harder to document definitively, many interviewees spoke of the Coalition's effect on the level of debate and beliefs about standards of evidence in the policy process through the Coalition's advocacy and training with agencies and Congress. One OMB staffer stated that the Coalition had helped make program funding more driven by evidence. A respondent who works outside the federal government said, "There's little question that they've raised the level of discourse about evidence. It's very evident in the last five years, that even if a program doesn't have good evidence, everyone is *claiming* to have evidence. And that's a step in the right direction." A critic of the Coalition, who works for the federal government, agreed that the Coalition had had an impact on government policy: "They've pressured agencies to get more rigorous program impact information. That's a good thing." (However, this individual clearly disagreed with the Coalition's general approach to program evaluation.) Similarly another critic of the Coalition stated that the Coalition had been successful in getting its language about evidence-based approaches and evaluation adopted by OMB and Congress. Indeed, in the interviews for this study, at least six respondents both inside and outside the federal government and across both branches of government spontaneously used the term "gold standard" when speaking of RCTs.

HOW THE COALITION IS SEEN: THE COALITION "BRAND"

In order to tap respondents' most basic reactions to the Coalition, respondents were asked, "When you think about the Coalition, what comes to mind?" A remarkably consistent set of themes surfaced. Those who interact with the Coalition associate it primarily with evaluation, evidence, and the methodology related to those areas, particularly RCTs. The words "rigor" or "rigorous" arose frequently. Interviewees saw the Coalition as an honest and credible voice. They spoke of its impartial and unbiased approach to research and methodology. Related to this, a number spoke positively about their perception of the Coalition as being apolitical. Many spoke of the Coalition's knowledge and expertise. (One respondent went so far as to characterize the Coalition, not entirely positively it seems, as "technocrats...pushing principles through math".)

Respondents associate the Coalition primarily with its work at the federal level with the executive branch and Congress. Some spoke of its advocacy role. Others mentioned its work in training and dissemination of evidence-based approaches. Some used words such as "selfless", "passionate", or "sincere" in describing the Coalition. A number of respondents mentioned the Coalition as being particularly, and effectively, focused on its objectives.

Given the nature of the sample, most of those interviewed were very positive about the Coalition. However, in keeping with John Stuart Mill's argument that "a

9. A description of the Alliance for Evidence-Based Family Strengthening is available at http://www.nciom.org/projects/adolescent/Hughes_5-8-09.pdf

THE COALITION IN THE RESPONDENTS' OWN WORDS

The following are abbreviated "quotations" from interviewees' responses to the question, "When you think about the Coalition, what comes to mind?" The "quotations" listed are the real words used by the respondents as written down by Dr. Herk. However, the respondents typically spoke in full sentences, rather than the telegraphic phrases listed below. That said, the phrases listed were actually said by individuals, and Dr. Herk was careful to note down their phrases as exactly as possible. Quotations were chosen as representative of a particular theme. The percentage following the heading indicates the percentage of interviewees expressing this theme in their response.

- ***Evaluation/Evidence/Methodology – 33% of interviewees***
"Evaluation. How to structure evaluations. How to design evaluations. How to interpret evaluations. How to assess their rigor." – Federal executive branch employee
- ***Rigor – 21%***
"Rigor. Integrity. Pushing or shifting the culture of decision-making at the federal level. Working with government agencies and the Hill..." – Individual who works with programs outside of Washington, DC
- ***Responsiveness/Helpfulness/User-friendly resources – 18%***
"Timely and relevant advice. One of the few interactions where I can consider the advice without worrying about the ulterior motive... Really responsive." – A Congressional staffer
- ***Honesty/Integrity/Credibility/Trust – 15%***
"[The Coalition has the] highest level of integrity around evidence."—Individual who works with programs outside of Washington, DC
- ***Impartiality/Lack of bias/Objectivity – 15%***
"Courage. Disregard for the politics of endorsing one program over another. Willingness to call it as they see it..."—Foundation staff person
- ***Knowledge/Expertise – 15%***
"A repository of knowledge and expertise. Unbiased. I can shoot Jon Baron an e-mail, and someone will give me an unbiased response. I trust their expertise." – A Congressional staffer
- ***Advocacy – 15%***
"An advocacy group for encouraging groups to use rigorous evidence for decisions and to advocate for increased use of rigorous evaluations to generate that evidence." – Staffer for an organization that works with programs for the disadvantaged
- ***Work with federal Executive Branch & Congress – 15%***
"Smart. Honest. Good at working with Congress and the Executive Branch." – Federal executive branch staffer
- ***And in the view of critics... – 12%***
"Unyielding focus on RCTs." – Federal executive branch staffer

person whose judgment is really deserving of confidence,” becomes that way “because he has kept his mind open to criticism of his opinions and conduct,”¹⁰ four critics of the Coalition were interviewed in order to learn more about their perspectives. A brief summary of their critiques is that, in their view, the Coalition overstates and oversimplifies the benefits of RCTs, thereby leading to the application of RCTs to inappropriate situations and ultimately hurting the cause of program evaluation. They would like to see the Coalition more fully address 1) the situations in which RCTs are inappropriate and 2) the value of non-RCT methodologies, both in these situations and more generally as a complement to RCTs. Some critics call their position “methodological pluralism”.¹¹

THE COALITION’S STRENGTHS

Following the initial broad question about the Coalition, respondents were asked to think specifically about the Coalition’s strengths and weaknesses. Many of the interviewees had already touched on strengths in their answers to the previous question, so many of the same themes arose. The single most common answer was that the Coalition was *knowledgeable and smart* about research methodologies and evidence-based approaches (10 respondents). Another very common answer (9 respondents) focused on the Coalition’s *responsiveness and helpfulness*. Many interviewees began with the words, “I sent Jon an e-mail...” and then recounted how they received a timely, helpful, and very specific response to their question. (Often they added just how rare a resource that is.) One interviewee who works outside Washington, DC said, “The Coalition is very customer-oriented and responsive. They’ve gone above and beyond to find answers to my questions.”

Another strength commonly cited by respondents was the Coalition’s skill at *communicating and explaining difficult concepts* to policy-makers, thereby creating a bridge between the policy and research worlds. One interviewee told the following story:

[An administrator at a federal agency] understood the evidence-based approach but was wondering how to get his staff to understand. He scheduled a brown bag. John gave a great presentation. Afterwards there were nodding heads. The brown bag led to an agreement that grant competitions would be run differently. After that the grant competition wasn’t perfect, but it was better. Jon leaves people with understanding. He seems accessible from so many different places.

10. John Stuart Mill, *On Liberty*, (Indianapolis: Hackett Publishing Company, Inc., 1978), p.19. Later Mill goes on to say, “Nor is it enough that he should hear the arguments of adversaries from his own teachers... He must be able to hear them from persons who actually believe them, who defend them in earnest and do their very utmost for them. He must know them in their most plausible and persuasive form...” (p.35)

11. This report was written without reference to the 2004 assessment. However, it is interesting to note that the 2004 assessment contains a remarkably similar passage concerning “reservations” about the Coalition’s approach. Martin, p.4-5. Similarly the 2004 report’s comments about “preaching to the already converted” echo statements in the “Weaknesses” section of this report. See Martin, p.4.

A Congressional staffer said, “The Coalition is able to convey what they know about the data in a way that staffers can understand.” Even a critic of the Coalition stated, “The Coalition has used the right rhetoric to help the idea of evidence-based policy-making gain currency.”

Echoing the themes expressed in the section on the Coalition’s “brand”, many respondents cited the Coalition’s *independent, objective, non-ideological, apolitical, credible, and honest voice* as one of its strengths. Others spoke of the Coalition’s *focus*, especially on its core issue of science and the quality of evidence, as a strength.

Finally, not as many respondents raised the issue of the Coalition’s *energy and action-orientation* as a strength, but the comments are interesting enough to warrant inclusion. A federal agency employee said, “Jon shows up in person, and that counts for a lot.”¹² A Congressional staffer said, “The Coalition is the first group in 11 years to come to me and ask to talk about these issues in a coherent way. There was only an occasional voice prior to them.” If anything, critics of the Coalition accord it even greater impact than some of its supporters. One such critic said, “When I went and talked with people at OMB, they said that the Coalition were the only ones really talking with them about evaluation methodology. That speaks very well for the Coalition. It speaks to the effectiveness of Jon Baron. And it speaks to the ineffectiveness of other organizations.”

WEAKNESSES OF THE COALITION

Respondents were specifically asked to name the Coalition’s weaknesses. Fifteen percent of the sample could not think of one. Among the other responses, common themes fell into three main categories: a need to expand beyond RCTs; a need for improved communications, especially with regard to the Coalition’s website; and a need for a better understanding of the political and policy context in which program decisions occur.

The most common weakness mentioned, by both supporters and critics, was the Coalition’s focus on RCT evaluation designs. (Note that this same focus was also cited as a strength by some respondents.) In the critics’ view, the Coalition overstates the applicability of RCT designs. One individual with a background in program evaluation said, “The Coalition’s over-advocacy of RCTs has significantly harmed the cause [of promoting them] even though in the short run they’ve been successful. In the long run, their efforts will hurt because they’ve over-advocated RCTs in situations that are inappropriate.” Even some strong supporters of the Coalition feel that the Coalition’s exclusive focus on RCTs has costs. One federal executive branch employee said, “Jon is so smart, and he listens well. Still, there is something of an all-or-nothing aspect to the Coalition’s support of RCTs that makes the Coalition have less impact than it could.”

A related theme had to do with the extremely small number of programs on the “top tier” list of interventions that have met the Coalition’s highest standard of evidence. Another federal executive branch staffer said, “So few things meet the top tier that the Coalition’s message rings hollow with some.” She added, “Right now there’s a push for innovation,” implying that those with an interest in “innovation” might find themselves

12. Or as Woody Allen famously put it, “Eighty percent of success is showing up.”
http://www.quotationspage.com/quotes/Woody_Allen

opposed to the Coalition's approach.¹³ Another executive branch staffer gave the example of a particular program area in which only five interventions reach the top tier: a mayor who wants to do something in that area would not find enough on the list. She said, "There's a purist aspect to it that comes across as rigid sometimes. We need to alert communities to what's promising and point them to those things." Others echoed the view that they would like to see more on what to do in the common situation in which there is not any top-tier research in a program area.

Another related issue was the desire by some respondents for the Coalition to pay more attention to replicability and implementation issues. One interviewee who works with programs outside of Washington, DC gave an example of a case where the developers of a program on the top-tier list were not interested in having their program replicated. The local community leaders who had reached out to the program developers became discouraged at that point. "It's great to promote science," this individual said, "but if no one can actually use it, it's a problem." He suggested that the Coalition include information on replication readiness and contact information on its top tier website. Another respondent, who also works with programs outside Washington, DC, stated, "The Coalition has a very strong understanding of research methodology. They're very good at reading research. They have less experience in the field. They're not as good at knowing what will happen when the rubber hits the road."

Communication

The next most common weakness mentioned, cited by 18% of the respondents, was the need to raise the Coalition's visibility by improving its communications and outreach. (But once again note that as many respondents praised the Coalition's communications and website as suggested that they needed improvement.) A few of the interviewees spoke of "moving beyond preaching to the choir" or "reaching out beyond the true believers". One executive branch staffer said that the Coalition could use media training and that their website is not good. He then qualified his comments by saying that the site was "good" in the sense that it obviously invested in content rather than glossiness, but added, "There's a reason why Washington groups invest in glossiness."

Political and Policy Context

Finally, a relatively small number of respondents (12%) expressed the view that the Coalition was insufficiently sensitive to the political or policy or programmatic context in which decisions are made. (Note again that this view is in contrast to those who see Jon Baron and the Coalition as being highly effective, consummate Washington insiders.) Although the numbers are small, this was the only other weakness that more than 10% of the sample agreed upon. Also, since all those giving this response work in the federal government, among federal respondents the percentage rises to 18%.

13. As examples of this point of view, see Lisbeth B. Schorr, "Innovative Reforms Require Innovative Scorekeeping," *Education Week*, 8/26/09, and Lisbeth B. Schorr, "To Judge What Will Best Help Society's Neediest, Let's Use a Broad Array of Evaluation Techniques," *The Chronicle of Philanthropy*, 8/20/09.

THE BOTTOM LINE: HOW THE COALITION CAN IMPROVE WHAT IT DOES – SOME POINTS FOR DISCUSSION

Based on these interviews, it seems reasonable to conclude that the Coalition has had a strong impact on government policy, especially for an organization of its size. Supporters, and even critics, agree on this point. The Coalition should take care to protect the very positive aspects of its “brand”, which it has developed through its hard work and successes over the past five years or more. The remaining question, set at the beginning of this process, is how the Coalition can further improve what it does. The following is structured as a set of discussion points, since there are tensions pulling in both directions for each item. These points are best thought of as questions that the Coalition might consider in dialogue with its supporters.

Discussion Point 1: To What Extent Should the Coalition Extend its Focus Beyond RCTs?

The question here has to do with the extent to which the Coalition might reach out beyond its traditional focus on RCT evaluation methodology to address other aspects of evidence-based interventions that some respondents are asking for. This could encompass a variety of efforts:

- Should the Coalition increase its effort to identify and disseminate “promising practices” in addition to its top tier programs?
- Should the Coalition expend more effort on implementation and replicability by providing more information on these issues in its materials, training, and website? Should the Coalition expend more of its own effort on exploring the implementation and replication experience of particular interventions?
- Related to the issue of promising interventions, should the Coalition either develop or adopt a model for the systematic development of a research base that would allow policy-makers to go from a situation in which there are many “good ideas” and little evidence (the situation characterizing many policy domains) to the desired state of truly evidence-based interventions? What would such a research cycle look like and what types of methodologies are appropriate at various stages? What would be the mix of studies and funding devoted to research at various levels (e.g., small, exploratory studies of promising ideas versus large-scale RCTs)?
- To what extent does the Coalition wish to talk about evaluation methodologies that are appropriate at less mature stages of the research cycle? Or does this diminish the strength of the Coalition’s focused message about the value of RCTs? One respondent said, “I would love to see the Coalition, in their push for rigor, take the lead in developing and publishing rigorous standards for criteria that must be met *before* an RCT should be mounted.”
- To what extent should the Coalition devote more resources to encouraging comparative efficacy designs that help policy-makers choose among competing interventions rather than all-or-nothing effectiveness studies, given that the government is unlikely to completely cease all interventions in an area where a program is found to be ineffective? To what extent should research focus on program improvement as opposed to program effectiveness?

- Should the Coalition devote more resources to helping organizations with “front end” program evaluation design? (The Coalition already operates the Evidence-Based Policy Help Desk, but some respondents felt that more or different technical assistance was desirable.)

Discussion Point 2: Drop “The Gold Standard” Language?

On the surface, this seems like a relatively trivial question, namely, whether the Coalition should drop the “gold standard” language when discussing RCTs, but it actually applies to a broader range of issues around how to educate the Coalition’s audience about evidence and research methodology. The reason for dropping the “gold standard” terminology, is that more than any other single thing, it acts as a red flag for a certain group of critics who feel that the Coalition lacks nuance in its approach and inappropriately disparages all other evaluation methodologies. As the Coalition’s impact has grown, so has the number of its opponents. These include professional evaluation methodologists, both inside and outside the government, who use evaluation methodologies in addition to or in place of RCTs, as well as program implementers who face the loss of funding as various funders require higher standards of evidence. These two camps differ in their motivations. Yet it is possible that their combined force may take a growing toll on the Coalition’s aims. The Coalition might consider whether there is any possibility of or value to a partial rapprochement with at least some of its critics among methodologists.

The cost of dropping “the gold standard” is that it is a handy catch phrase that helps many policy-makers remember the benefits of RCT evidence – for example, almost a fifth of the sample volunteered the phrase in their interviews. The Coalition could move towards a more nuanced, academic explication of evaluation methodology in its materials, but many of the policy-makers in this study valued the Coalition’s explanation of complex topics in language that they, as people with backgrounds in law or other non-technical areas, could understand. There may be a fundamental tension between “Evaluation 101” that some policy-makers need and the more nuanced approaches (“Evaluation 401”) that some professional methodologists desire. Or there may be carefully chosen language that can bridge the gap.

Discussion Point 3: Should the Coalition Become More Institutionalized? – the Bus Question

Finally, fewer than 10% of respondents brought up the Coalition’s small size as a weakness. Nevertheless, the Coalition and its supporters might consider the question of the Coalition’s size and resilience, and whether there should be investment in greater institutionalization of the organization. Although the Coalition consists of more than just Jon Baron (and a few respondents mentioned David Anderson), it was clear from the interviews that most of the respondents equate the Coalition with Baron. In many cases interviewees used “the Coalition” and “Jon Baron” as essential equivalents.

The Coalition clearly has had a large impact, especially in light of its size. The question then arises, if the proverbial bus accident should occur, what happens to the Coalition and its efforts? Does it all go away?

More broadly, many of the possible initiatives posed under Discussion Point 1, would require additional resources, primarily in the form of staff or their equivalents.

Clearly expanding the Coalition so that its work and vision is institutionalized and spread across a broader range of individuals will require thought from the Coalition and its supporters regarding whether this outcome is desired, and if so, how it can be supported.

Discussion Point 4: Should the Coalition Reach Beyond “the Choir”?

A few respondents suggested that the Coalition’s current work primarily involves individuals who are already convinced by the Coalition’s evidence-based approach (although it may have taken some effort to get them to that point). If the organization chooses to grow and extend its reach, then an important next step would be to develop and implement a targeted outreach and marketing plan that would bring its message to a broader range of key decision-makers. Of course, the implementation of such a plan would also require the Coalition to grow as described in Discussion Point 3 in order to meet the added demand for its services.

Through the Coalition for Evidence-Based Policy, the philanthropic foundations supporting it have greatly advanced evidence-based policy-making since 2004 for a relatively small investment. In light of the Coalition’s substantial success to date, the next question facing the funders and the Coalition is how best to build on that groundwork moving forward.

Appendix A

SEMI-STRUCTURED QUESTIONNAIRE CEBP & Cost-Benefit Assessment

I am conducting this study under contract with the Coalition for Evidence-Based Policy for the W.T. Grant Foundation and the MacArthur Foundation. The purpose of this study is to improve the effectiveness of the Coalition and of other efforts being made by the foundations. Our conversation today will be on a *not for attribution* basis. Only I will see my notes from this interview. Any information from this conversation that appears in my final report will be presented in such a way that it is impossible to identify the person making the comment. [Give estimate of how long the interview will last.] Do you have any questions for me?

1. When you are making or advising on a program or policy decision, what considerations do you think about?

[Have them score those considerations:] On a scale of 1 to 5 where 1 is not important at all and 5 is extremely important, how important are each of those considerations? [Then have them score any of the following that they left out:]

- a. Political considerations
 - b. Popular support for the policy
 - c. The effectiveness of the policy – how likely the policy is to achieve its goals
 - d. What other considerations do you think about? [Get ratings of their importance.]
2. On a scale of 1 to 5, where 1 is very little latitude and 5 is complete latitude, how much latitude would you say you have to make these decisions or is the decision mostly in the hands of your superiors? [If low, press for how much difference they think their advice makes.]
 3. When you're thinking about program or policy effectiveness, what types of evidence do you think about?
 - a. If necessary, prompt for types of evidence they might use:
 - i. Results from state & local level; pilot programs
 - ii. Research studies (type, from where)
 - iii. Cost-benefit studies
 - iv. Experiments with randomized-control groups [probe a bit to see if they really understand this]
 - b. [Prompt for the organizations/individuals that they turn to for advice on program effectiveness]
 - c. [If they don't mention CEBP, ask if they've heard of it.]

4. [If they've heard of CEBP] When you think about CEBP, what comes to mind?
 - a. Do you think they've had an impact on government policy? If so, how and why? If not, why not?
 - b. Do you/would you turn to them for help in determining policy effectiveness? Why or why not?
 - c. What do you think are CEBP's strengths? Weaknesses?
 - d. Do you think there are any ways that CEBP could change what it does to make it more helpful to you?
5. When you hear the words "cost-benefit analysis" what comes to mind?
 - a. Do you use cost-benefit analysis in your work?
 - b. What are the advantages of cost-benefit analysis? Disadvantages?
 - c. How convincing do you think policy-makers find cost-benefit analysis? [Probe for reasons.]
 - d. Do you have suggestions for improving how cost-benefit analysis is conducted or how the results are presented to policy-makers that would increase its impact on decisions?
6. Anything else that you would like to add? Would you like to see a copy of the final report? Thank you for your time.

Appendix B

Responses to the question, “When you are making or advising on a program or policy decision, what considerations do you think about?” that were rated on a scale of 1 to 5 for importance by the respondent

The following list consists of all the responses in which the interviewee actually rated how important the consideration was on a scale of 1 to 5 where 1 was “not important at all” and 5 was “extremely important”. Considerations are listed in order of frequency of response. For considerations which three or more respondents rated, the average rating is given following the consideration. The ‘n’ given is the number of respondents who provided a numerical rating of importance for that consideration. (Additional interviewees may have given this response, but they are counted in this list only if they rated the consideration for importance.)

- Political considerations, 3.3 (n=8)
- Effectiveness of the intervention in meeting its goals, 4.6 (n=7)
- Feasibility/ease of implementation, 4.4 (n=4)
- Cost of the intervention/budgetary constraints, 4.0 (n=4)
- Popular support for the intervention, 2.0 (n=3)
- Cost-benefit analysis (n=2)

(All remaining responses were rated by only one person and are listed to give a sense of the range of responses.)

- The distribution of benefits of the intervention
- Whether resources are available for the intervention
- Need for the intervention
- Popularity of the intervention with Congress members
- What is known about best practices and gaps
- Whether the intervention fits the interests of the respondent’s direct boss
- Whether it is possible to measure results of the intervention
- Whether it is possible to track the funds associated with the intervention
- Social/environmental impact of the intervention
- Ability to implement the intervention within the required timeline
- Respondent’s authority to take the action
- Replicability of the intervention
- The respondent’s experience with the person to whom he/she is making the recommendation
- Whether the intervention has the support of its intended beneficiaries
- If the intervention is discretionary, whether it requires hiring new personnel
- Whether evidence of the intervention’s effectiveness exists*

* This respondent distinguished whether evidence of effectiveness existed from whether the intervention was in fact effective.

- The acceptability of the intervention (i.e., whether it violated community or ethical norms)
- The flexibility of the intervention (i.e., whether it could be implemented in a wide variety of jurisdictions)
- Whether the intervention was part of the President's budget