Boruch R, Petrosino A, Chalmers I. The Campbell Collaboration: a proposal for systematic, multinational, and continuous reviews of evidence. In: Davies P, Petrosino A, Chalmers I, eds. The effects of social and educational interventions: developing an infrastructure for international collaboration to prepare, maintain and promote the accessibility of systematic reviews of relevant research. London: School of Public Policy, University College London, 1999. Available: http://www.ucl.ac.uk/spp/download/publications/Annexe7.pdf (accessed 17 June 2003)

The Campbell Collaboration: a proposal for systematic, multi-national, and continuous reviews of evidence

Robert Boruch, University of Pennsylvania
Anthony Petrosino, Harvard Children's Initiative
Iain Chalmers, UK Cochrane Centre

Background paper for the meeting at
The School of Public Policy, University College London, 15/16 July 1999

Contents

1. Introduction: The proposed Campbell Collaboration

The proposed Campbell Collaboration's object is to develop continuously updated, multi-national systematic reviews of studies on the effects of demonstration programs in the social and behavioral sectors, including education. Primary attention in the near term is dedicated to randomized field trials. Secondary attention, and in the longer term, is dedicated to non-randomized field trials.

Our focus is on the evidence on the effects of new programs, demonstrations, and new variations on programs. The areas targeted for reviews include (but are not limited to) preschool education and interventions such as Head Start, elementary, secondary and higher education, delinquency and criminal justice, mental health and substance use, social services including welfare, housing, and employment and training. Inasmuch as the quality of evidence

is an issue, we will attend also to methodological studies on enhancing the quality of study design, data, implementation, and analyses.

The main target audience for these reviews are people with a strong interest in high quality evidence on what works. These include professional researchers, practitioners, policy makers, educators and their students, and members of the public who want to keep abreast of the best evidence on the effectiveness of innovative programs and projects.

The proposed Campbell Collaboration transcends geo-political boundaries in that it will routinely engage people from multiple countries. Further, it will involve teams of both professional researchers and lay persons in reviewing the research. This inclusion principle is to assure that the products of the systematic reviews are relevant to diverse audiences. That is, we presume that understanding what works, for whom, and for how long, and what does not work has interest for many groups.

The proposed Campbell Collaboration will be world-wide web oriented in addition to relying on conventional methods of communication. This is because the effort is multi-national and designed for continuous access to reviews and their continuous updating. This orientation entails construction of web sites in cooperation with the Cochrane Collaboration in health.

These aims will be achieved through an effort that depends on its participants' agreement to eight basic principles (identified in what follows) for the proposed Campbell Collaboration. This Collaboration involves developing Review Groups that produce systematic reviews of high quality evidence, and developing international Campbell Collaboration Centers that are located in various countries and which will provide infrastructure support to the Collaboration.

In what follows, we describe the rationale for an effort of this kind, the aims of the proposed Campbell Collaboration, a set of operating principles for a Collaboration, and a description of the organizational ingredients that will drive the production of high quality systematic reviews of high quality evidence about the effectiveness of interventions. We also propose to create several Review Groups and Centers as pilot efforts to understand the feasibility of the proposed Collaboration over the next few years.

2. Rationale and justification

Why is it sensible to develop a multi-national collaboration to produce systematic reviews of studies on what innovations in the social and behavioral sectors and in education work and fail? Three justifications are given here: (a) a surge of interest in high quality evidence for public policy decisions, (b) the increased frequency in using randomized field trials to discern the effects of new programs or new variations on existing programs, and (c) the increased frequency of systematic reviews.

2.1 Surge of interest in better evidence for policy

Over the past three years, there has been a remarkable surge in concerns about the quality of evidence being produced on the effectiveness of various interventions and new social programs.

This surge is reflected in recent international conferences organized in Germany by Karl Schumann on the importance of randomized field trials in research in civil, criminal, and juvenile justice (Krisberg and Schumann, In Press). The conference contributors were from Germany, Switzerland, the United Kingdom, the United States, and Israel. The interest is also reflected in international meetings convened by Sweden's new national Center for Evaluation of Social Services. These involved researchers from Sweden, Denmark, Norway, Finland, Britain, Scotland, Canada, and the United States, and focused on how to enhance the quality of evidence through randomized trials and other approaches. Proceedings were published in the Scandinavian Journal of Social Work (Soydan, 1998). In the United Kingdom, a biennial series of international conferences has been inaugurated by Carol Fitz-Gibbon to focus on evidence-based policy, including fostering higher quality evidence. A cross sectoral meeting at St. Andrew's University was the basis for a special issue of the United Kingdom's accounting journal Public Money and Management (Volume 19) in 1999.

In the health arena, since its inauguration in 1993, the international Cochrane Collaboration has grown to consist of about 50 Collaborative Review Groups whose members are preparing and maintaining systematic reviews of the effects of healthcare interventions. The work of these groups is coordinated and supported in various ways by 15 Cochrane Centers scattered around the world, and there are now contributors to the Cochrane Collaboration in well over 50 countries (Chalmers et al., 1997; Cochrane Collaboration Web Site).

At least three efforts in the United States are important partly because they graded the quality of evidence that has been produced on new programs. Sherman et al. (1997), for instance, produced a report for the U.S. Congress on the effectiveness of juvenile and adult crime prevention programs in the U.S. and gave grades to each study based on the quality of evidence produced on each program. The Herman et al. (1999) report on evidence about the effectiveness of twenty-four "school-wide education reform" programs is important partly because it attempted to consolidate evidence to grade quality of evidence. The Academy on Welfare Reform is a third example (Besharov, Rossi, Germanis, 1997). This effort to review evidence on innovative welfare reform programs was critical and self-critical in appraising the quality of the evidence and in summarizing the implications of dependable evidence.

Such efforts reflect (a) recent trans-national and national interests, (b) a focus on high quality evidence, and (c) diverse discipline areas. With the exception of the Cochrane Collaboration and perhaps the Sherman et al. work, however, they are episodic. Moreover, with some exceptions, the proceedings are not easily accessible to an interested general public, although there are some exceptions. The proposed Campbell

Collaboration is designed to capitalize on the enthusiasm that motivates this interest in evidence, to build on it, and to reiterate the need for quality in evidence through the Collaboration's reviews and other products. The collaboration will also foster continuous rather than episodic interest attention to systematic reviews and helps to legitimize, consolidate, and make more accessible the cross-national and cross-discipline aspects of such interest.

2.2 Increased use of Randomized Field Trials (RFTs)

Although they are not common, randomized field trials (RFTs) for evaluating the effects of programs and projects are increasing in frequency and scope to judge from available evidence (Figures 1, 2, 3, and 4). Figure 1 summarizes RFTs completed in social and behavioral sciences, education, and welfare during 1945-1975. The upward trend is obvious. Figure 2 summarizes the incidence of publications on RFTs in juvenile, civil, and criminal justice during 1955-1987. Again, the trend is obvious. Figure 3 summarizes the incidence of completed economic experiments, showing an escalation. Finally, Figure 4 shows the incidence of controlled trials that were randomized or that were probably randomized in laboratory and field settings in education, psychology and other areas (Petrosino et al., 1999). This chart is based on an archive that was developed partly to augment the proposed Campbell Collaboration. In particular, the SPECTR File (Social, Psychological, Educational, and Criminological Trials Register) was constructed partly for the Cochrane Collaboration Library and will be part of the Library's web site, a site that is accessible to the public. The 10,000 entries include randomized trials in the field and laboratory, possibly randomized trials, and other controlled studies. There are 10,000 entries based on a trained-hand searcher's judgements of articles that have appeared in 48 refereed journals and from specialized reference lists. The screening criteria were designed to assure high sensitivity (maximizing the number of controlled field trials that could be found) and specificity (minimizing the number of uncontrolled trials). The search strategy, though incomplete in ambit by some standards, was conscientious; the screening for controlled trials was also conscientious. See Petrosino et al. (1999) and the annexes for detail on computer-based code and key word searches.

Systematic reviews and syntheses need to be updated frequently because the newer randomized trials build on, get beyond, and sometimes contradict earlier trials. In any event, developing a continuous cumulative statement of findings is in the interest of science and the public. However, with the exception of the Cochrane Collaboration, no nation and no scientific or policy research organization of which we are aware has created a system for continuously updated reviews. The Campbell Collaboration that is proposed here can fill this gap in the social and behavioral sectors, including education.

In building a registry of RFTs, both the studies that are published in refereed journals and reports that are not published in journals are important. Focusing only on one type of distribution will lead to biases in the conclusions that are drawn. Locating studies of both kinds continuously, particularly the unpublished ones, is a demanding task beyond the capacity of a single individual or institution. It requires cooperation among many people

to develop complete registries and to keep them updated. Some people, especially in the United States, have done remarkable work in this arena nonetheless, e.g., Lipsey (1992), Hedges (1988), Cooper, Glass and Smith, among others. The proposed Campbell Collaboration would build on the experience of individual meta-analysts, get beyond it so as to make their task easier, and build an infrastructure to assure that their analyses are informed in new ways and are accessible to people beyond the research community who want to know about the effects of new interventions.

That there is a need for a multi-national perspective seems clear. RFTs on fertility control programs or programmatic efforts to reduce out-of-wedlock pregnancies, for instance, have been run in the U.S., Taiwan, India, African countries and elsewhere. Results of these and of evaluations of other kinds of programs are rarely reviewed in meta-analyses in ways that transcend geo-political boundaries. The proposed Campbell Collaboration would do so.

The Campbell Collaboration proposes to establish an infrastructure that will make it easier to register, track, and update RFTs on effects of programs. With experience and development, we expect the Collaboration to eventually enlarge its ambit to include non-randomized trials ("observational studies: or "quasi-experiments) in much the same spirit. This infrastructure, we expect, will make more efficient the task of building registries across discipline and country and, more important, help various multinational Review Groups to construct reviews that are helpful to the public and scientific communities. It will also help to foster Campbell Collaboration projects that advance methodology, for instance, in combining evidence from RFTs with evidence from passive surveillance systems (details are given later).

11.11Evidence and its quality

In some social, behavioral, and education areas, we have no evidence beyond anecdote on the effects of innovative demonstration projects and interventions. Recall, for instance, the Herman et al. (1999) findings that about 20% of new comprehensive school programs in the United States had no scientifically defensible evidence to support their claims that the interventions enhanced children's achievement.

Further, evidence on the effects of new interventions has varied considerably in quality. Recall, for instance, the University of Maryland's report to the U.S. Congress on what works, what does not, and what is promising in criminal and juvenile justice (Sherman et al., 1997). The report graded program evaluations on the quality of evidence that was produced using a simple grading scheme. Less than 30% of the studies that were reviewed had produced sufficiently high quality evidence to determine whether the program worked.

Recall also that in the welfare reform arena, the research literature is large, complex, and varies greatly in its quality. Partly because reviewing all of it might be impossible, the

Academy on Welfare Reform chose to focus heavily on high quality impact studies, especially on randomized field experiments (Besharov, Rossi, Germanis, 1997).

The best of contemporary meta-analyses on impact studies have tried to take the quality of each study's design into account, usually by classifying the study's design into various categories. Some, perhaps many, meta-analyses have not done so, however, or do so in ways that do not make the standards of judgement transparent.

Complicating matters is the capacity of some individuals, institutions, and nations to make a study available quickly and with little effort through the world wide web, uncountable hard copy "briefs", and other media. A particular study may be worthless, relative to good standards of evidence. Nonetheless, it can be made accessible quickly to a large number of people.

The proposed Campbell Collaboration's intent is to assure that the standards for judging the quality of studies' designs transparent and uniformly applied, and that study search strategies also meet criteria for completeness and coverage.

11 Objectives for the proposed Campbell Collaboration

The Campbell Collaboration's objectives are to:

- produce high quality systematic reviews of research on the effects of new programs and program variations;
- make as transparent as possible the standards employed in assuring the quality of systematic reviews and the standards used to judge the quality of studies that are reviewed;
- assisting reviewers by developing useful raw materials, such as registries, of high quality studies;
- foster criticism of review standards and the improvement of standards;
- assure that reviews of research are intelligible and relevant to people who are interested in evidence on the effects of interventions; and
- make the reviews easily accessible through user friendly world-wide-web sites and other media.

These objectives should address the need to foster better evidence of the effectiveness of interventions in the social and behavioral sciences, and in education research.

12 Operating principles for the proposed Campbell Collaboration

We rely partly on the Cochrane Collaboration precedent in health research to enunciate eight principle values that will guide the development of the proposed Campbell Collaboration in the social and behavioral sciences and education research:

- collaboration, apart from competition, among scientists and practitioners, across nations,
- building on the enthusiasm of individuals with an interest in the objective,
- avoiding duplication of efforts by researchers,
- minimizing bias in searches of literature, structure of systematic reviews, publications, etc. through the use of transparent standards for each,
- keeping up-to-date, through plans for continuous updating,
- assuring relevance to contemporary science, practice, and policy, through the engagement of lay persons,
- assuring access to results of systematic reviews through world-wide-web sites and CD-ROMS as well as journal articles,
- continually improving the quality of the work, by eliciting criticism continuously, through web-based reports, methodological studies on quality of evidence and reviews of the evidence, and in other ways
- continuity, by ensuring that responsibility for reviews, editorial processes and key functions is maintained and renewed

These principles appear to have served the Cochrane Collaboration well. We will depend tentatively on them in this proposed Campbell Collaboration. In some measure, the principles have served well in other efforts to collaborate in the synthesizing of research and in making judgements about the quality of research. Actualizing these principles is what concerns us next.

13 Operations: proposed Campbell Collaboration Review Groups

Campbell Collaboration Review Groups will be responsible for producing systematic reviews of studies in areas that are of interest to the Group. The members of each Review Group will be (a) international and (b) include people from the research community and other communities with an interest in evidence on the effects of interventions. This composition rule is intended to assure that the products of Review Group efforts have value in the scientific, professional, and public forums and that they are not limited by professional, cultural, or geopolitical boundaries. It also helps to increase the possibility that multiple versions of a systematic review will be produced.

Each Campbell Collaboration Review Group will be responsible for:

- identifying the target area for a systematic review of studies on interventions, and
- developing a plan outlining how the systematic review would contribute to the Campbell Collaboration's overall aims.

A multi-national Review Group might, for example, focus on randomized field trials of "charter schools" in education, on studies of class size, or on the effect of scholarships or other incentive systems for high achieving children from poor families. Or, a Review

Group might build on, augment, and reanalyze the results produced in the Sherman et al. report to Congress on crime prevention programs. Other Review Groups might be cross-cutting in focusing on all high quality randomized field trials that are of a particular outcome variable regardless of program type, e.g., children's achievement on standardized test.

A Group's plan would normally be based on prior communications among the prospective Review Group's members. The plans would outline responsibilities for an identified coordinating editor and the identified team of reviewers, and for developing a registry of studies and plan for its updating.

Part of the technical assistance for the Campbell Collaboration Review Groups can be provided through the Cochrane Collaboration workshops and materials. This is to avoid duplication and enhance efficiency. The Cochrane Collaboration materials include (a) protocols on uniform methods for world wide-web-based searches, hand searching research journals and other sources, and (b) software (RevMan) for facilitating the organization, preparation, analysis, and presentation of reviews. These resources are unique and should be exploited by the proposed Campbell Collaboration.

Start-up Review Group proposals

At this point, we believe that at least two Review Groups should be created to serve as test beds for this feature of the proposed Campbell Collaboration. Candidate review groups will doubtless emerge at the London meeting. To provoke conversation, we offer two options. The first concerns randomized field trials on approaches to preventing and handling delinquency. The second concerns randomized field trials on teaching mathematics and science.

The focus of the Review Group on Delinquency Prevention and Treatment is on studies of prevention and treatment programs that may have origins in the crime prevention traditions as well as studies of housing, social services, mental health and substance use. This effort will also be multi-national, of course. The object is to (a) exploit the SPECTR File and to augment it with auxiliary literature and hand searches of journals, (b) pilot test a review protocol, RevMan Software, and communications system for review, and (c) develop the systematic review in forms accessible on the web and other media. This effort will extend Petrosino's earlier work in this arena, build on the Gottfredson (1998) and Sherman et al. (1997) reports to the U.S. Congress, and work by others (e.g., Elliot et al., 1998; Lipsey, 1992; Durlak, 1997; Lipton, 1995).

The second option focuses on randomized trials in teaching mathematics and science. The aims of this pilot effort are to exploit and augment the SPECTR archive, use and troubleshoot protocols and RevMan Software for constructing a review group, planning a review and completing the systematic review, and developing web-based and other systems for making the review accessible and useable.

11 Operations: proposed Campbell Collaboration Centers

The Campbell Collaboration Review Groups that have just been described are responsible for the actual production of systematic reviews. The Campbell Collaboration Centers would have been a different mission. It is to help to coordinate the Campbell Collaboration and to support the efforts of the Review Groups and the Collaboration.

A Campbell Collaboration Center may focus on statistical methods for summarizing diverse studies of program effects. Or, such a Center might review research on the use of the internet for doing conventional surveys in social, behavioral, and educational research on the use of the internet in randomized field trials that assess the value of different ways to disseminate scientific information. Or, methodologically oriented Centers might attend to regression-discontinuity designs, or the uses of qualitative studies embedded in large-scale evaluations of program effects, or coupling randomized field trials to systematic observational (non-randomized) studies.

More generally, Campbell Collaboration Centers would be responsible for:

- maintaining a registry of people who are willing to collaborate on their research interests,
- helping to establish Campbell Collaboration Review Groups,
- coordinating and/or assisting in searches of the research literature in different languages and areas,
- contributing to the Campbell Collaboration's registries of high quality evaluations, especially RFTs, and to updating registries,
- helping to systematize and update guidelines on reporting and software that facilitates good reporting,
- exploring ways of helping the public, service providers, journalists, policy-makers to make full use of Campbell Collaboration Reviews,
- organizing workshops, seminars, and colloquia to support and guide the development of the Campbell Collaboration, and
- exchanging information with the multi-national Cochrane Collaboration in the health sector.

Start-up Center proposals

To initiate the Centers' development, we propose to create two Centers.

(a) **Center on macro level Randomized Field Trials (RFTs).** Macro-level RFTs involve the random assignment of a set of entire organizations or geopolitical jurisdictions to one of two or more treatment programs. They are also referred to cluster trials or group-wide experiments, depending on discipline area. Boruch and Foley (1999), identified over 50 such RFTs in which entire schools, police jurisdictions, hospitals, and other entities have been assigned. These RFTs are

infrequent but important. Moreover, they engender problems that are common across disciplines, political—institutional, managerial, scientific, and ethical boundaries. A Center would focus continuously on problems and their solutions.

(b) **Cross-design synthesis and analyses.** At times, RFTs can be designed so as to make it easy to combine RFT results with the results of analyses of data from surveillance systems, administrative record systems, non-randomized trials and other less quantitative approaches such as organizational process studies, and ethnographic work. The combination of analyses from RFTs and surveillance systems, for example, carries the promise of estimating program effects that are unbiased (i.e., based on RFTs) and generalizable (i.e., based on national surveillance systems). See, for instance, Boruch and Terhanian (1998) and references therein on the possibilities, and Kunz and Oxman (1998) on serious potential problems.

7. Operations: Relations between the proposed Campbell Collaboration and the existing Cochrane Collaboration.

The Cochrane Collaboration has had considerable experience in understanding how to produce high quality, publicly accessible, and systematic reviews of evidence on the effects of interventions in the health sector. That experience includes understanding the infrastructure that is needed to do the job, the incentives, and disincentives for collaborators, and other matters.

The proposed Campbell Collaboration is in the fortunate position of being able to build on and at times get beyond the current Cochrane Collaboration. Moreover, there is a nice opportunity to enhance communications among diverse communities of researchers by connecting the Cochrane Collaboration and the proposed Campbell Collaboration.

The proposed Campbell Collaboration will capitalize on the Cochrane Collaboration in obvious ways. These include:

- efforts by the Cochrane Collaboration to explore new arenas, represented by meetings on the proposed Campbell Collaboration,
- the Cochrane Collaboration's development of protocols for systematic reviews of studies on the effects of interventions and protocols that can be exploited or adapted in a Campbell Collaboration,
- the Cochrane Collaboration's agreement on principles for guiding the intellectual effort , principles that can be adapted to a Campbell Collaboration, and
- the Cochrane Collaboration's work on organizational structure that could be emulated by the proposed Campbell Collaboration.

There are also fine opportunities for other partnership arrangements between the proposed Campbell Collaboration and the existing Cochrane Collaboration. There are at least two options that we propose be pursued. First, one or more current Review Groups in the Cochrane Collaboration might be made as joint Review Groups with the proposed

Campbell Collaboration. The Cochrane Developmental, Psychosocial and Learning Problems Group, which is coordinated by a professor of social work (Geraldine Macdonald), is an example. It relates to both the Cochrane Collaboration and the proposed Campbell Collaboration.

The second option for partnership between the Cochrane Collaboration and the proposed Campbell Collaboration hinges on each effort's common interest in evidence and the methods of producing good evidence. The Cochrane Collaboration has about a dozen Methods Groups whose members are assessing and developing ways of improving methodologies for preparing systematic reviews. These Groups are confronting issues such as statistical methods, meta-analyses using individual patient data and control of bias in systematic reviews, and they are beginning to prepare systematic reviews of empirical methodological studies. This work is coordinated by the current chair of the Collaboration, Dr. Andy Oxman, and one of the associate directors at the U.K. Cochrane Centre, Mike Clarke. Both are keen to explore how the Cochrane and Campbell Collaborations could work together to tackle these cross-domain methodological issues.

11 Basic expected products

We expect the proposed Campbell Collaboration to produce at least four basic kinds of products. The first kind includes high quality systematic reviews of studies on the effects of new programs in education and in the social and behavioral sectors. These reviews will be produced for people who want to understand the effects of interventions based on high quality evidence from different countries. They will provide people with easily accessible and up-to-date evidence about what works, for whom, for how long, and about what does not work.

These systematic reviews will be accessible on a Campbell Collaboration Web Site along with transparent standards on how studies were chosen and reviewed, and the standards used to judge the quality of evidence that the studies produced. The Campbell Collaboration Web Site will be cross referenced to the Cochrane Collaboration Site. Further, systematic reviews that are produced by the Campbell Collaboration would be issued in peer reviewed hard-copy journals or world wide web-based electronic journals. We also expect the reviews to be incorporated into continuing education courses, workshops, and conferences.

These systematic reviews would, moreover, be subject to criticism by anybody. This opportunity for criticism is important and will be actively fostered through the creation of a Campbell Collaboration.

A second broad class of products issued by the Campbell Collaboration includes listings of studies. A good example of one kind of listing is the SPECTR Database on randomized and non-randomized field trials in the social and behavioral sciences and education (Petrosino, et al., 1999). Such listings are, in effect, registries or archives on the studies that fall within various Review Group ambits. This registry then allows anyone—

researcher, practitioner, civil servant, or member of the public—to get back to an original study that was included in a review, and to exploit it, or to call attention to its virtues or flaws. It permits other reviews of related information to be done easily, independent of the Campbell Collaboration or with the Campbell Collaboration's support. This class of products, registries, will also be produced in electronic form, available on the World Wide Web and CD ROM, for instance.

A third class of products is the identification of important areas in which no studies have been undertaken to produce good evidence about the effects of particular programs. Until 1997, for example, only one randomized field trial in the United States had been done on school choice/voucher programs. There appear to have been no high quality field trials on the effect of small scholarship awards on the academic achievement of children from poor families. The Campbell Collaboration can make a distinctive contribution, we believe, to mapping the areas in which high quality randomized field trials have not been done, as well as mapping the areas in which such studies have been done.

A fourth class of products hinges on indicators of scientific progress as reflected in the updated systematic reviews. In particular, the production of these reviews will enhance our ability to track progress in different areas, permit us to learn about, and provide some of the raw material for thinking strategically about research and evaluation policy in the social, behavioral, and education research.

12 Operations: A node at the University of Pennsylvania

The development work on the proposed Campbell Collaboration has to be located in places where there are people with interest and expertise. There has already been encouraging progress in establishing the basis for this within the UK, particularly in the field of education (Davies 1999). However, there is wide acknowledgement that a high proportion of the data and expertise relevant to the tasks outlined above are located in North America. It is important, therefore, to establish a strong base there, early in the development of the Campbell Collaboration.

For the years 2000-2005, we suggest that the development work in the North and South America and Asia be located first at the University of Pennsylvania in the United States. During this period, and afterwards, other nodes in the form of Centers and Review Groups will (we expect) be developed elsewhere in the United States and more important, in other countries.

Depending on the University of Pennsylvania as a starting node has several justifications. First, a remarkable number of people at the University have contributed substantially to mounting randomized field trials to estimate the effects of new programs. They include: Phyllis Solomon (School of Social Work), Trevor Hadley and Jeffrey Draine (Center for Mental Health Policy Research) on programs for the seriously mentally ill, Lawrence Sherman (Department of Sociology and Director of the Fels Institute) on criminal and juvenile justice programs, Rebecca Maynard (Graduate School of Education) on welfare

reform and teenage parent/pregnancy prevention, Robert Boruch (Graduate School of Education, Wharton School of Business) on randomized field trials in all social sectors, John Fantuzzo (Psychology in Education, GSE, and Psychology) on early childhood programs, Margaret Spencer (Graduate School of Education and Psychology Department) on scholarship programs for low-income children, Linda Aiken (School of Nursing) on behavioral aspects of health interventions in AIDS and other diseases, Burt Cohen and Neil Weiner (Center for Youth Policy) juvenile and adult criminal justice.

Other people at Penn have contributed seriously to the design of observational studies (Cochran, 1983). These are studies that can be designed well before the introduction of new programs so as to estimate the program's effect. Colleagues who are expert in this arena and who also understand the benefits of randomized field trials include Paul Rosenbaum (1995) in the Department of Statistics, and Robert Hornick and Martin Fishbein at the Annenberg School of Communications.

To reiterate a point made earlier, the University of Pennsylvania is one proposed node in the United States. Other talented people are engaged in related work in the U.S.; recall experts on meta-analysis that include Rosenthal (Harvard), Cook (Northwestern), Cooper (Missouri), Hedges (Chicago). Still other talented people are located at potential institutional nodes in Canada, Colombia and Brazil, and in some Pacific Rim countries.

11 Resource commitments for reconnaissance of the Campbell Collaboration

Between May 1998 and July 1999, UK Cochrane Centre (which is supported by the National Health Service Research and Development Programme) and the School of Public Policy at University College London invested resources in exploring the possibility that a sibling organization to the Cochrane Collaboration might be established to prepare, maintain and disseminate systematic reviews of social and educational interventions. These resources included the organization of meetings at the School of Public Policy in May 1998, and January and July 1999, and support for Anthony Petrosino to attend the meetings in 1999; the time of UK Cochrane Centre staff, and other resources, for consultations about a possible Campbell Collaboration, and to establish an electronic register of over 10,000 references to reports of controlled trials of social, psychological, educational and criminological studies (Petrosino et al., 1999). Anthony Petrosino's reconnaissance was also supported by the Harvard Children's Initiative and the American Academy of Arts and Sciences.

Between January 1999 and May 1999, modest commitments were sought by Boruch in the United States to explore the feasibility and desirability of the proposed Campbell Collaboration and to lay some groundwork for its inauguration. The following commitments were sought and then made during this period. The U.S. Department of Education's Planning Evaluation Service, headed by Dr. Alan Ginsburg was brisk in providing $25,000 (U.S.) for reconnaissance. Dean of the School of Social Work, Dr. Ira Schwartz, and Dean of the Graduate School of Education, Dr. Susan Fuhrman, at the University of Pennsylvania will commit resources at the University in FY 2000 for the Collaboration's inaugural meetings.

Inasmuch as education is a target for systematic reviews, and science and mathematics education is a potential target for a Campbell Collaboration Review Group, part of Boruch's time is sponsored. This is through a National Science Foundation Grant to the University of Pennsylvania and the Pennsylvania State University for analyses of data from the Third International Mathematics and Science Study and research on related topics. The proposed Campbell Collaboration is one such related topic.

12 Tasks to which resources must be allocated

The successful development of the proposed Campbell Collaboration requires that a number of tasks be done well. Each task, of course, requires resources. The tasks are identified briefly in what follows.

11.1 Facilitating development of the proposed Campbell Collaboration

In October 1998, Boruch agreed with colleagues in the United Kingdom to invest time in reconnaissance and development of the Campbell Collaboration. Iain Chalmers, Philip Davies, and other members of the Cochrane Collaboration have dedicated time to exploring a Campbell Collaboration. Fred Mosteller of Harvard's Center for Children's Initiatives has dedicated time to reconnaissance. Anthony Petrosino, a Spencer Fellow at the Center, has been industrious in his exploratory work.

Over the next few years, Boruch is willing to dedicate 50% of his time to further development of the proposed Campbell Collaboration. Anthony Petrosino is willing to dedicate 80% of his time to doing so, Iain Chalmers and other members of the Cochrane Collaboration in health have agreed to assist in the effort. The cost of these individual's time has to be subsidized.

11.11Knowledge-base development

Uncovering published and unpublished reports from around the world on the effects of interventions in education and the social behavioral sectors is not easy. It is essential, however, to do so especially in the early stages of the proposed Campbell Collaboration (2000-2005). The physical form this will take is the SPECTR data base that has been compiled on randomized field trials and possibly, on randomized field trials whose current contents are represented in Figure 4 (Petrosino et al., 1999). Continuous updating of registries and archives is essential for the continuous updating of reviews. The continued development of the Database of Systematic Reviews in the Social Sciences is also important to the knowledge base in the Campbell Collaboration. These will help others to augment reviews, develop better reviews, and enhance the value of reviews for informing science and public policy.

The development of the registry of studies on the effects of interventions, notably the SPECTR database, and of a registry of meta-analyses, depends heavily on people who know how to do this well. Anthony Petrosino at Harvard and his colleagues at the Cochrane Center have the ability. The time that they dedicate to this has to be paid for.

11.3 Web-Site Development and integration of information systems

Setting up and maintaining one or more world-wide web sites for the proposed Campbell Collaboration will take able people. Any such web set-up ought be integrated with the archive, reproduction, scanning, distribution system at local institutional Campbell Collaboration Review Groups or Centers. This set-up also ought to be integrated with telephone systems. And it must be integrated with any technical support system to assist others in downloads, etc.

The Cochrane Collaboration people are a valuable resource in all this. The proposed Campbell Collaboration will get beyond Cochrane Collaboration systems, assuming that resources are available for web-site development and technical assistance on development, and will provide information back to the Cochrane community.

Web-site development, and especially the electronic integration of this with other information systems, requires resources. These must be paid for.

11.11Hard-copy/electronic interface

Not all studies that should be reviewed by the Campbell Collaboration Review Groups or Centers will be available in machine readable form e.g., diskette, web-site text files) during 2000-2005. The capacity to put hard copies into machine readable form is in the interest of brisk transfer of information. This capacity will vary across nation.

To reduce the burden of hard-copy reports, high quality reproduction and electronic scanning capacity is essential in the long term. The high quality hard-copy reproduction must be subsidized. The cost of high quality scanning equipment must be paid for. And the technical support, especially for interpretation of systems and scanning equipment, must be underwritten.

11.12Review Group and Center development

Part of the cost of Review Groups' and Centers' development is tied up in time and resources dedicated for the Collaboration's infrastructure. Beyond this, it seems sensible to be able to allocate resources to prospective Review Groups or Centers. The resources might include technical assistance or seed money or both.

11.13Language translation interface

Software exists for crude translation of basic reports and, more important, the reviews of studies that would be produced on the WWW. This software is insufficient for excellent translations of both original studies and the reviews of such studies. All this is despite the increasing use of the English language in multi-national forums and studies.

It seems sensible to develop resources to support translators with special knowledge of terminology in social scientific evidence and research.

11.14Travel

Travel can be minimized through the worldwide web/internet, teleconferencing, telephone and mail. Each is excellent for discussions of facts. Nonetheless, face-to-face meetings are essential at times. We confine such meetings to topics that focus partly on fact and, more important, address issues of judgement, standards, and mutual education and partnership.

To develop the Campbell Collaboration over 2000-2005, principals in North America will have to make cross-continental trips between North America and Europe, North America to North and South Asia, North America to South America, North America to the Middle East and North/South Africa. Resources will be needed for this travel.

11.15Secretarial & staff support

Developing systematic reviews involves people from different countries. Communications and correspondence through emailing, faxing, and the mail must be accurate and brisk, especially in coordination of calendars and tasks. Sophisticated internet-oriented assistance is essential. These resources must be paid for.

11.16Information distribution channels

High quality systematic reviews depend on each Review Group's and Center's members having the information readily at-hand in each member's country. The primary distribution channel will be the world wide web. The secondary channels, depending on the country, will include expedited mail systems and overland telephone lines.

The cost of assuring that each channel is exploited well must be paid for. In effect, many of the costs for this would be paid for through investments in web site development and other tasks described elsewhere. Nonetheless, residual costs and costs of alternatives must be sustained.

11.10 Space

The proposed Campbell Collaboration emphasizes the exploitation of the world wide web and on electronic archiving and communications. Nonetheless, we expect that physical space will be necessary for archiving hard-copy materials and for equipment. People who maintain these will also need space.

The cost of hard-copy archive space and equipment space, and space for people will have to be paid for.

11.11 Start-up (pilot) Review Groups and Centers

In an earlier section, we proposed several options for new Review Groups and Centers. Recall that the aim at this point is to use these options as test beds for further development of the Campbell Collaboration.

The two proposed Review Groups will focus respectively on Delinquency Programs and on Mathematics and Science Education. The two proposed Centers will concern respectively Macro-level Randomized Field Trials and on Cross-Design Synthesis and Analysis. The London meetings will cover other possibilities that may be better or more attractive in a multi-national context.

Regardless of which particular Groups or Centers are chosen, the seed money to engage the interest of prospective participants would be very helpful.

12. Bibliography and web sites

Besharov, D., Germanis, P. and Rossi, P. (1997) Evaluating Welfare Reform. College Park, MD: University of Maryland and American Enterprize Institute (dbesharov@puafmail.umd.edu)

Boruch, R. F. (1997) Randomized Experiments for Planning and Evaluation: A Practical Guide. Thousand Oaks, CA: Sage Publications.

Boruch, R. F. and Terhanian, G. (1998) Cross-Design Synthesis. In H. Walberg and A. Reynolds (Eds.) Advances in Educational Productivity, Volume 7, Greenwich, CT: JAI Press, pp. 59-85.

Boruch, R. F. and Foley, E. (1999) Using Organizations, Schools, and other Entities as the Units of Assignment and Analysis in Randomized Field Trials. In L. Bickman (Ed.) The Legacy of Donald T. Campbell, Thousand Oaks, CA: Sage Publications, In Press.

Chalmers, I., Sackett, D. and Silagy, C. (1997) The Cochrane Collaboration. In A. Maynard and I. Chalmers (Eds.) Nonrandom Reflections on Health Services Research. London: The British Medical Journal (BMJ) Publishing Group, pp. 231-249.

Cochran, W. G. (1983) Planning and Analysis of Observational Studies. New York: Wiley.

Cochrane Collaboration (www.cochrane.org)

Cook, T. D., Cooper, H., Cordray, D., Hartmann, H., Hedges, L., Light, R. A., Louis, T. A. and Mosteller, F. (Eds) (1992) Meta-analysis for Explanation. New York: Russell Sage Foundation.

Cook, T. D. and Campbell, D. T. (1979) Quasi-experimentation: Design and Analysis Issues for Experimentation: Design and Analysis Issues for Field Settings. Boston: Houghton-Mifflin.

Cook, T. D. and Reichardt (1979) Qualitative and Quantitative Methods in Evaluation. Thousand Oaks, CA: Sage Publication (www.sagepub.com)

Cooper, H. (1988) Synthesizing Research. Thousand Oaks, CA: Sage Publications. (www.sagepub.com)

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

Durlak, J. A. (1997) Successful Prevention Programs for Children and Adolescents. New York: Plenum Press.

Elliott, D.S., Hamburg, B.A., and Williams, K. R. (1998) Violence in American Schools. Cambridge, MA. Cambridge University Press.

Gottfredson, D.C. (1997) Chapter 5. School-based Crime Prevention. In: University of Maryland, Department of Criminology and Criminal Justice. Preventing Crime: What Works, What Doesn't, What's Promising. (Report to the U.S. Congress, prepared for the national Institute of Justice). College Park, Maryland: University of Maryland, pp. 5-1 to 5-74. (Also at: http://www.ncjrs.org)

Herman, R. et al. (1999) The Effects of Comprehensive School-wide Education Reforms. Washington, D.C.: American Institute for Research.

Krisberg, B. and Schumann, K. (Eds.) (In Press) Special Issue of Crime and Delinquency.

Kunz, R. and Oxman, A. D. (1998) The Unpredictability Paradox: A Review of Empirical Comparisons of randomized and Non-randomized Clinical Trials. British Medical Journal, 317, 1185-11909.

Lipsey, M. W. (1992) Juvenile Delinquency Treatment: A Meta-analytic Inquiry into the Variability of Effects. In T. Cook and others (Eds) Meta-analysis for Explanation. New York: Russell Sage Foundation, pp. 83-127.

Lipton, D. S. (1995) CDATE: Updating the Effectiveness of Correctional Treatment 25 Years Later. Journal of Offender Rehabilitation, 22, 1-2, 1-20.

Mosteller, F., Light, R. J. and Sachs, J. A. (1996) Sustained Inquiry in Education: Lessons for Skill Grouping and Class Size. Harvard Educational Review, 66, 797-842.

Nave, B., Miech, E. and Mosteller, F. (1999) A Rare Design: The Role of Field Trials in Evaluating School Practices. Research Report. Cambridge, MA: Harvard University.

Petrosino, A. J. and others (1999) A Social, Psychological, Educational and Criminal Trials Register (SPECTR) to Facilitate the Preparation and Maintenance of Systematic Reviews of Social and Educational Interventions. Report prepared for Meetings on Evidence-Based Policy and Indicators, University of Durham and Planning Meeting on Research Synthesis and Public Policy, University College London, U.K. July 1999.

Petrosino, A. J. (1998) Preliminary Findings from 150 Randomized Evaluations in Crime Reduction. JRSA Forum, 1, 5-9.

Petrosino, A. J. (1995) The Hunt for Experimental Reports: Document Search and Efforts for a "What Works?" Meta-analysis. Journal of Crime and Justice, 18, 63-80.

Petrosino, A. J. (1995) Specifying the Inclusion Criteria in Meta-Analysis. Evaluation Review, 19(3), 284-293.

Rosenbaum, P. (1995) Observational Studies. New York: Springer-Verlag.

Rosenthal, R. (1991) Meta-analytic Methods for Social Research. Thousand Oaks, CA: Sage Publications.

Schwandt, T. A. and Halpern, E. S. (1988) Linking Auditing and Meta-Evaluation. Thousand Oaks, CA: Sage Publications (www.sagepub.com)

Sherman, L. and others (1997) What Works, What Doesn't, and What's Promising? A Report to the U.S. Congress. Washington D.C.: National Institute of Justice, U.S. Department of Justice, (and College Park, Maryland, U.S.: University of Maryland).

Tashakori, A. and Teddlie, C. (1998) Mixed Methodology Combining Qualitative and Quantitative Approaches. Thousand Oaks, CA: Sage Publications.

Weisburd, D. L., Sherman, L. and Petrosino, A. J. (1990) Registry of Randomized Experiments in Criminal Sanction, 1950-1983. Los Altos, CA: Sociometrics Corporation.

Appendix

Questions and Answers about the Proposed Campbell Collaboration

Question: How would the proposed Campbell Collaboration differ from other efforts to produce "meta-analysis?"

Answer: The Collaboration would perforce, depend on contemporary technology in meta-analysis and its application by an individual or a small team. It would differ in several respects. First, the Collaboration Review Groups will include both lay persons and scientists, not just scientists, as is currently the case in most academic or institutional meta-analyses. Second, the Collaboration will be multi-national. Current meta-analytic work is done on studies within nation or are bounded by a single language such as English.

Third, the Collaboration would involve continuous updating of registries and reviews. Contemporary meta-analyses are usually "one-shot" efforts. Further, there is a need for brisk response to criticism of reviews and to the completion of new studies. Fourth, the Collaboration will depend on a plainly stated (transparent) and uniform protocol for the design of reviews and for assigning grades to the quality of evidence that was produced in each of the studies under review. In contrast, the standards used in judging quality of evidence are not uniform in contemporary meta-analyses as they are stated only in the best of meta-analyses, and they are often stated in ways that do not make clear to a lay public what the standards are.

Fifth, the results of most meta-analyses and systematic reviews of research are published in academic journals, using a particular language and presentation styles.
Although academic journals are in principle accessible to any eager reader, they are not accessible in practice as other media. In any case, the language used in such meta-analyses is not easily understood by non-researchers who would like to learn from such reviews. The proposed Campbell Collaboration in the social and education sectors will emulate and get beyond the Cochrane Collaboration in the health sectors to assure that reviews are (a) accessible on the world wide web and (b the reviews can be understood by both researchers and non-researchers.

Question: How would the Campbell Collaboration differ from efforts by, for instance, the National Academy of Sciences in the United States to review certain research literature?

Answer: The NAS efforts usually focus on one particular research area at the request of a federal agency or the Congress. In contrast, the Collaboration will focus on many research areas and the focus will be a function of the interests of the Collaboration, and it will have audience beyond the federal level, including practitioners. Certainly, the Collaboration's products could inform Academy deliberations and, to avoid duplication, can choose to offer to collaborate or to defer a study.

Further, the Collaboration will include data from many countries. The NAS often focuses solely on studies done in the United States, international work committees being an exception. Finally, the genesis of a systematic review includes the scientific community, the lay public, and service providers, local or state government, or (as in the NAS case), the national government. These reviews will be open to criticism and continuing revision; in doing so, they get beyond a typical single report.

Question: How would the Collaboration's effort differ from others' in producing training materials for implementing new programs?

Answer: Virtually no contemporary meta-analyses produce training materials as part of a meta-analysis. Nonetheless, valuable short-courses, graduate level courses, and workshops on meta-analysis have been produced. For instance, the Cochrane Collaboration organizes a programme of training workshops on developing protocols for systematic reviews, and using its Review Manager software. At this stage, we make no specific provisions for inventing training materials in the proposed Campbell Collaboration. We prefer to depend on and exploit the work already done by the Cochrane Collaboration. The proposed Campbell Collaboration will, nonetheless, compile a continuous roster of such training materials that will be made available on its web site.

Question: How would the Collaboration's registries differ from related efforts to develop clearinghouses of data-sets or studies?

Answer: Any conscientiously constructed bibliography on RFTs that is published in a peer-reviewed journal is, in effect, a registry of RFTs. The charts given in this proposal are based on such bibliographies. However, such bibliographies are often based on the idiosyncratic search strategies of the authors of a review, fragmentary resources, and imperfectly designed search engines. In contrast, the Campbell Collaboration will operate as follows.

First, the Campbell Collaboration's search strategies will be transparent and uniform. The emphasis on transparency and uniformity is based partly on successful precedent (notably the Cochrane Collaboration) and continuously improved. Second, the Collaboration's search strategies will focus on both reports published in peer reviewed journals and those which are not. Reports that are not published in refereed journals are informative and have to be recognized. Third, the Collaboration's registries are to be continuously updated; this in fact is a condition for researchers who want to participate in the Collaboration. Finally, the Collaboration's focus is on high quality studies, notably randomized field trials. This is in sharp contrast to "clearinghouses" which usually do not screen for quality.