



How To Use Rigorous Evidence About “What Works” To Improve Program Effectiveness:

Practical, Cost-Effective Strategies

Jon Baron

Coalition for Evidence-Based Policy

March 18, 2008 – Blueprints Conference

Coalition for Evidence-Based Policy

Board of Advisors

Robert Boruch - Co-Chair,
Campbell Collaboration

David Ellwood - Dean, Harvard's
JFK School

Judith Gueron – fmr President,
MDRC

Ron Haskins – Sr Advisor to
President for Welfare Policy

Robert Hoyt – Founder, Jennison
Associates

Blair Hull – Founder, Hull Trading
Co

David Kessler – fmr FDA
Commissioner

Jerry Lee – President, Jerry Lee
Foundation

Dan Levy – Researcher,
Mathematica

Diane Ravitch – fmr Asst
Secretary of Education

Laurie Robinson – fmr Asst
Attorney General, DOJ

Howard Rolston – fmr Director of
Research, HHS/ACF

Isabel Sawhill – fmr Associate
Director of OMB

Martin Seligman – fmr President,
American Psychological Assn

Robert Solow – Nobel Laureate in
Economics, MIT

Nicholas Zill – Vice-President,
Westat



1. Brief Overview of Evidence-Based Policy

(10 minutes, + 5 minutes Q&A)

The Problem: Little progress in many areas of policy

- ◆ U.S. has made very limited progress in raising K-12 achievement over past 30 years.
- ◆ U.S. poverty rate today is higher than in 1973.
- ◆ U.S. has made no significant progress versus drug/alcohol abuse since 1990.



Rigorous evaluations have identified interventions that are ineffective/harmful:

◆ **Vouchers for disadvantaged workers, to subsidize their employment**

Well-designed randomized trial found large negative effects on employment.

◆ **Scared Straight**

Shown in randomized trials to cause a small *increase* in subsequent criminal activity by participating youth.

◆ **Drug Abuse Resistance Education (DARE)**

Ineffective in preventing substance use, according to randomized trials (is now being redesigned).

Rigorous evaluations have identified a few highly-effective interventions:

◆ Nurse-Family Partnership

By age 15, results in 40-70% reductions in child abuse/neglect, and criminal arrests of children.

◆ Carrera Teen Pregnancy Prevention Program – a youth program w/ academic assistance, job club, & sex ed (reduced females' pregnancies/births by 40-50% at 3-year followup; increased high school completion & college enrollment by 30-40% at 7 year followup, vs. controls).

Evidence-based policy seeks to incorporate two main reforms into social programs:

1. Increased funding for rigorous evaluations, to build the number of research-proven interventions.
2. Strong incentives and assistance for program grantees to adopt the research-proven interventions.

(We suggest not focusing on evaluating whole federal programs.)

Congressional trend toward randomized controlled trials where feasible

- ◆ House and Senate Committee reports accompanying the FY08 Appropriations Act (Public Law 110-161):

"The Committee strongly supports the [Education] Department's efforts to carry out Congressionally-authorized evaluations . . . using rigorous methodologies, particularly random assignment, that are capable of producing scientifically-valid knowledge regarding which program activities are effective."

Congressional trend, continued

- ◆ FY 08 Appropriations Act creates an evidence-based home visitation program at HHS, directing HHS to –

“ensure that States use the funds to support models that have been shown, in well-designed randomized controlled trials, to produce sizeable, sustained effects on important child outcomes such as abuse and neglect . . . [and] not to incorporate any additional initiatives that have not met these high evidentiary standards.”

Congressional trend, continued

- ◆ Second Chance Act (enacted last week) to facilitate the re-entry of prisoners into the community contains a 2% set-aside for evaluations that –

"include, to the maximum extent feasible, random assignment . . . And generate evidence on which re-entry approaches and strategies are most effective."

U.S. Office of Management and Budget (OMB) Guidance

- ◆ Identifies the randomized controlled trial, where feasible, as the “highest quality, unbiased design” for measuring a program’s effectiveness, and well-matched comparison-group studies as a second-best alternative when a randomized controlled trial is not feasible.



2. How To Assess Whether An Intervention Is Supported By Rigorous Evidence: Key Principles

(20 minutes, + 10 min Q&A)

There is strong evidence to support:

- ◆ Well-designed randomized controlled trials as the highest quality evaluation to determine program impact.
- ◆ Well-matched comparison-group studies as a second-best alternative when randomized controlled trials are not feasible.

(Comparison-group studies are sometimes called “quasi-experimental” studies.)

Less rigorous study designs:

- ◆ Less rigorous study designs include:
 - Unmatched or partially-matched comparison-group studies; and
 - Pre-post studies
- ◆ Such designs can be very useful in generating hypotheses about what works, but often produce erroneous conclusions.

Job Training Partnership Act: Impact on Earnings of Male Youth (Non-arrestees)

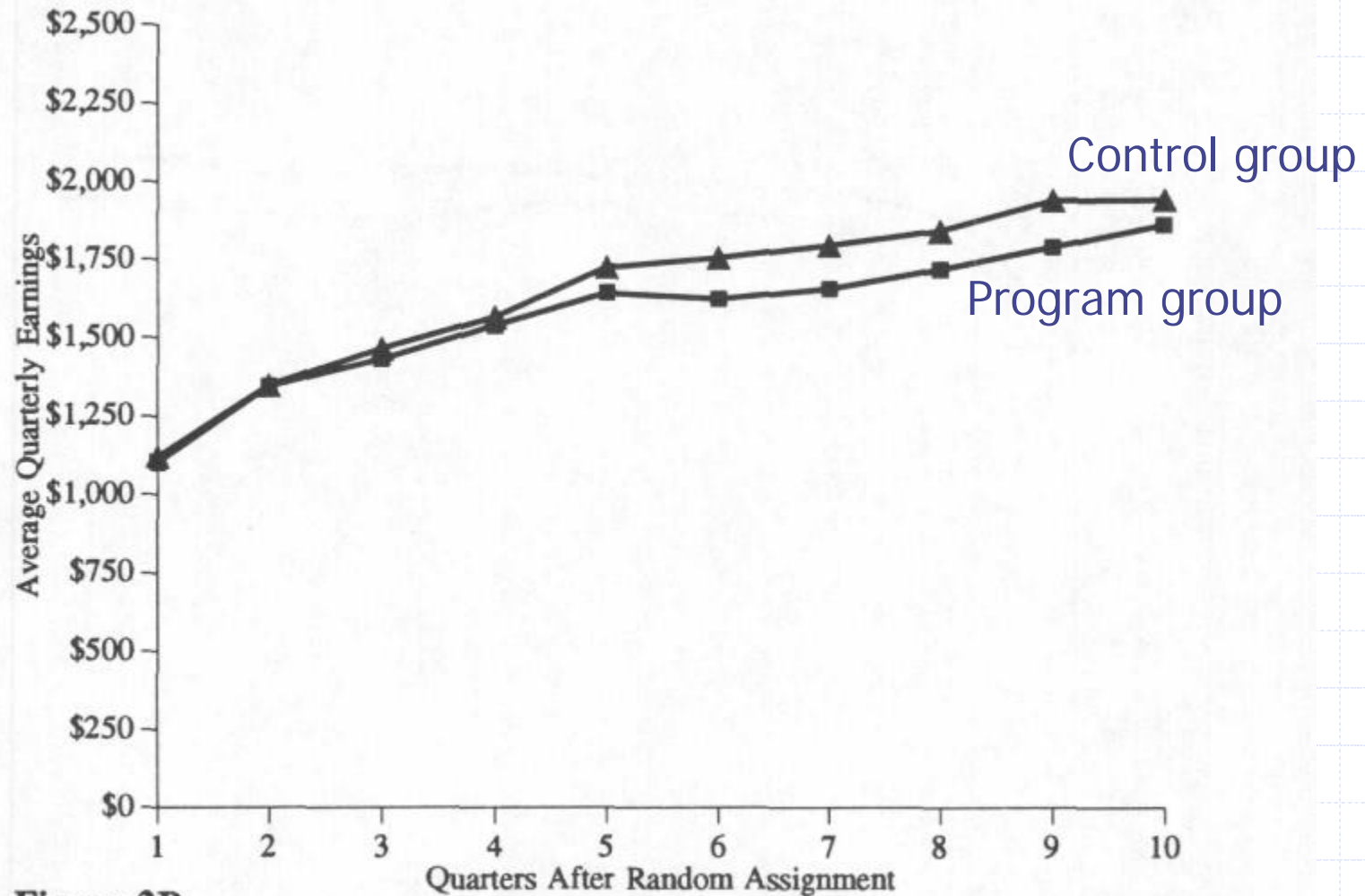
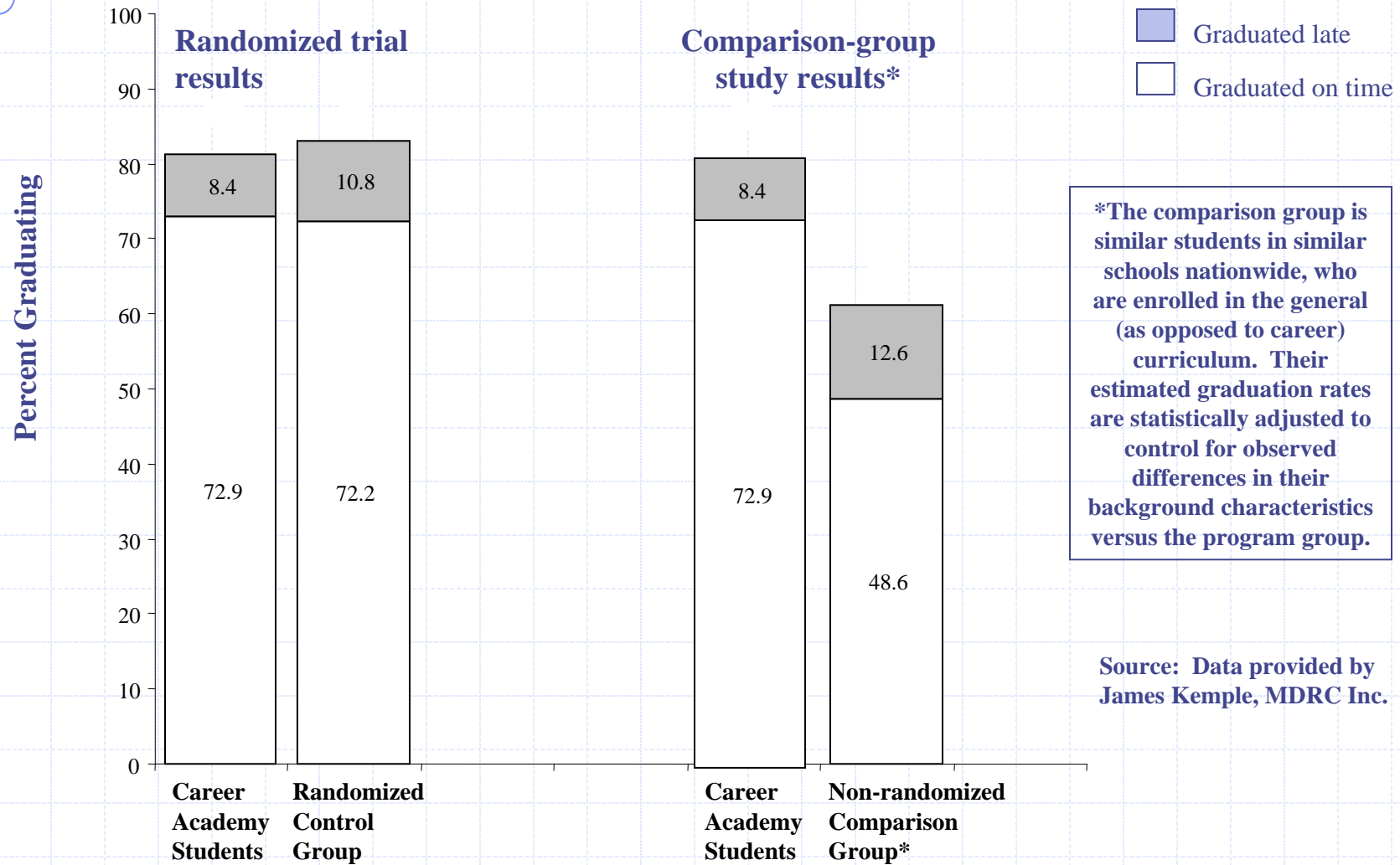


Figure 2B
Mean Earnings, by Quarter: Male Youth Non-arrestees

Impact of Career Academies on High School Graduation Rates



Key elements of a “well-designed” randomized controlled trial

1. The study should clearly describe the intervention.
2. If appropriate, the study randomly assigns groups (e.g., classrooms), not just individuals within those groups (e.g., students).
3. The study has an adequate sample size – one large enough to detect meaningful effects of the intervention.

Key elements of a “well-designed” randomized controlled trial

4. The study should show that the intervention & control groups were similar in key characteristics prior to the intervention.
5. Few or no control group members participated in the intervention, or otherwise benefited from it.
6. The study obtained outcome data for a high proportion of the sample members originally randomized (i.e., there is “low attrition”).
7. The study report outcome data even for those in the intervention group who don't complete the intervention (i.e., “intention-to-treat” approach).

Key elements of a “well-designed” randomized controlled trial

8. The study uses outcome measures that are valid – i.e., are highly correlated with the true outcomes the intervention is designed to affect.
 - E.g., it uses educational or psychological tests whose validity is well-established;
 - Self-reported outcomes (e.g., about criminal activity) are preferably corroborated by independent and/or objective measures (e.g, arrest rates).

Key elements of a “well-designed” randomized controlled trial

9. Where appropriate, evaluators are kept unaware of who is in the intervention vs control group (i.e., are “blinded”).
10. The study measures important policy or practical outcomes that the intervention seeks to affect – not just surrogate outcomes.
11. The study preferably obtains data on long-term outcomes of the intervention (e.g., a year after intervention ends, preferably longer).

Key elements of a “well-designed” randomized controlled trial

12. If the study claims that the intervention is effective, it should report:

- the size of effect; and
- statistical tests showing the effect is unlikely to be due to chance (including “hierarchical” tests if groups were randomized).

Key elements of a “well-designed” randomized controlled trial

13. A study's claim that the intervention's effect on a subgroup (e.g., Hispanics) differs from the effect on overall population should be treated with caution.
14. The study reports the intervention's effect on all the outcomes the study measured.
15. Preferably, the study evaluated the intervention in the real-world community settings & conditions where it would normally be implemented.

Quantity of evidence needed to establish that an intervention is effective:

- ◆ Preferably, well-designed randomized trials showing effectiveness in 2 or more typical community settings (can be 2 separate trials, or one multi-site trial).

Characteristics of a “well-matched comparison-group study” (= 2nd best alternative if a randomized trial is not possible):

- ◆ Intervention and comparison group members are closely matched on characteristics that may predict their outcomes (e.g., for drug-abuse prevention program, match on prior drug use*, age, sex).
- ◆ Intervention group should not be comprised of volunteers, and comparison group is non-volunteers.
- ◆ The intervention & comparison groups are chosen “prospectively” – i.e. before intervention is administered.
- ◆ The study should meeting guidelines above for a well-designed randomized trial (low attrition etc).

Example of a well-matched comparison-group study

- ◆ Study at Univ of South Florida that evaluated 2 ways of teaching general chemistry – large lecture vs. large lecture w/ small-group instruction taught by students.



3. How to Effectively Replicate and/or Scale-up Research- Proven Interventions

(10 minutes + 5 minutes Q&A)

Example of Successful Replication -- Welfare-to-Work

- ◆ **Riverside GAIN Program:** to move welfare recipients quickly into workforce through short-term job search & training (at 5-year follow-up, increased single-parent employment & earnings by ~40% vs. control group).
- ◆ **L.A. Jobs-First GAIN:** scrapped their ineffective program, and replicated key features of Riverside program (results: at 2-year follow-up, increased employment & earnings by ~30% vs. control group).

Key Steps

- ◆ Select the intervention
- ◆ Select the implementation sites
- ◆ Implement a system to ensure close adherence to the intervention's key features

Step 1: Select the Intervention

◆ Key things to look for:

- The intervention has been shown in rigorous – preferably randomized -- evaluations to have sustained, sustained effects
- The intervention been evaluated in a site and with a population similar to the setting in which the intervention is to be delivered

Websites for identifying evidence-based social programs

- ◆ Social Programs that Work
www.evidencebasedprograms.org
- ◆ Blueprints for Violence Prevention
www.colorado.edu/cspv/blueprints/index.html
- ◆ What Works Clearinghouse (Dept of Education)
<http://www.whatworks.ed.gov>
- ◆ International Campbell Collaboration
www.campbellcollaboration.org

Step 2: Select Implementation sites

◆ Important site characteristics:

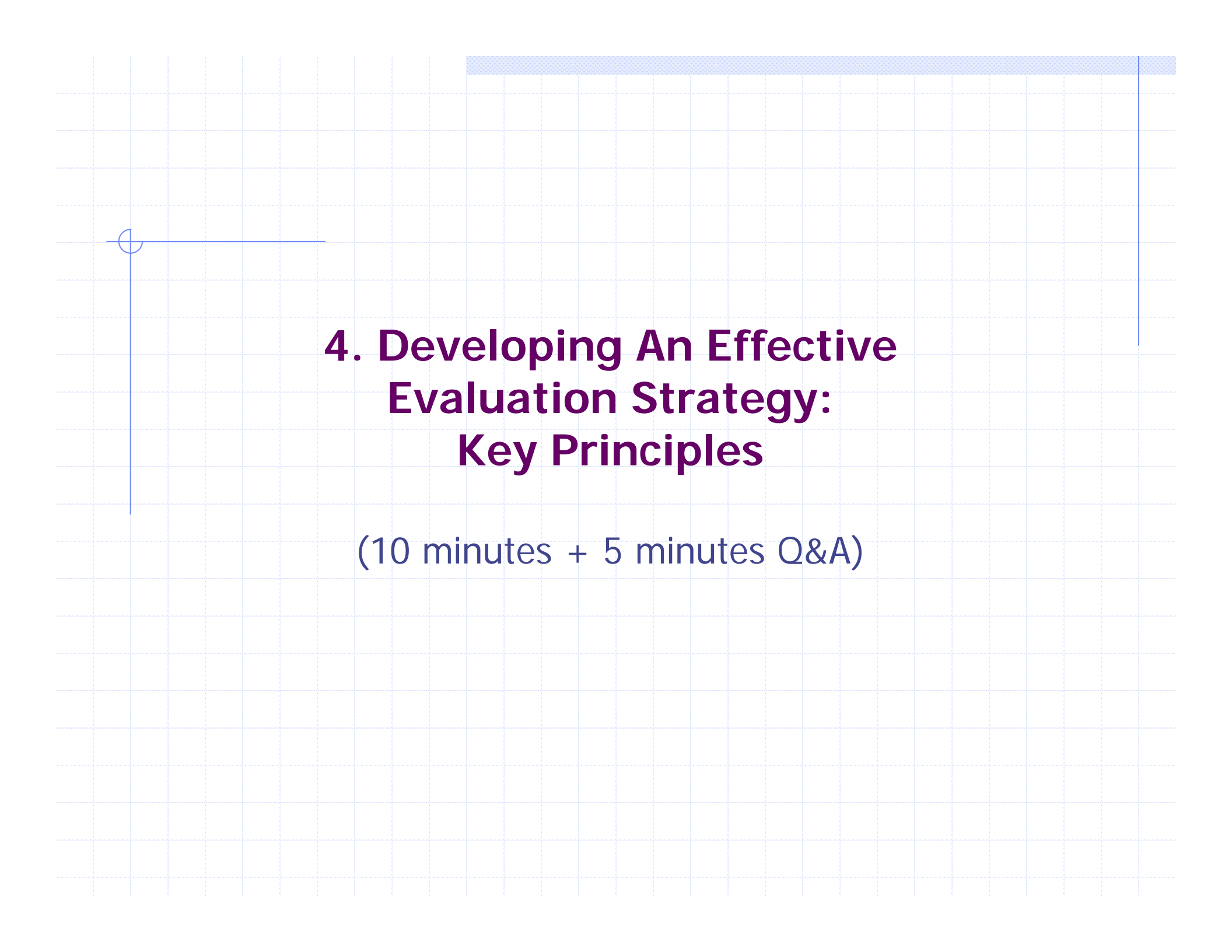
- A top official at the site who will be a capable, strong proponent of the intervention
- Support for the intervention among other administrators and program delivery staff
- The ability to commit sufficient financial and other resources to the intervention

Step 3: Ensure close adherence to the intervention's key features

- ◆ *How* evidence-based interventions are implemented is extremely important, in that minor changes in implementation can often make a major difference in the size of the intervention's effects.

Implementing a system to ensure close adherence

- ◆ Work closely with program developer in delivering the program.
- ◆ Closely monitor whether program delivery is adhering closely to the model (e.g., via periodic observation, videotape, etc)
- ◆ Correct deviations (e.g., via booster training sessions, on-site coaching of program deliverers, etc)



4. Developing An Effective Evaluation Strategy: Key Principles

(10 minutes + 5 minutes Q&A)

Goal:

- ◆ To assist you in developing a research & evaluation strategy that is likely to succeed in identifying, in rigorous evaluations, at least a few interventions with meaningful positive effects.

Suggestions on Evaluation Strategy

1. To evaluate effectiveness (or “impact”) of an intervention, use a randomized controlled trial. If not feasible (only if), consider a well-matched comparison-group study.

Suggestions on Evaluation Strategy

2. Focus rigorous evaluations on most promising interventions (as opposed to evaluating everything).
3. Make sure an intervention is well-developed and well-implemented before rigorously evaluating its effectiveness.
4. Recognize that well-designed randomized trials can sometimes be done at modest cost (e.g., \$50,000).

Suggestions on Evaluation Strategy

5. Select a capable study team! (one with a demonstrated track record in well-designed, rigorous evaluations.)
6. Provide the study team with clear guidance and/or assistance to prevent common flaws in study design/implementation.
7. Where To Start? Start by trying to get a small number of rigorous evaluations underway (e.g., one or two).

To select a capable evaluator:

- ◆ Ask your candidate(s) to send you 2 randomized controlled trials they have conducted (at least one in program area similar to yours); and
- ◆ Have the trials briefly reviewed by 1-2 experts (using a checklist) to determine if well-design and implemented.

Grant/contract mechanisms that federal programs have used to advance rigorous evaluations

- ◆ Competitive priority for projects that include a rigorous (preferably randomized) evaluation.
- ◆ Absolute priority (i.e., requirement) for projects to include such an evaluation.
- ◆ The program sponsors the evaluation, and requires grantees to participate in the evaluation if asked.
- ◆ The program funds a sheltered competition to evaluate a specific model at several program sites.
- ◆ The agency “waives” law/regulation to allow demonstration projects, and requires rigorous evaluation.

Coalition for Evidence-Based Policy
www.excelgov.org/evidence

Evidence-Based Policy Help Desk for OMB/Agencies:
www.evidencebasedpolicy.org

Social Programs That Work
www.evidencebasedprograms.org

Jon Baron

jbaron@excelgov.org

202-530-3279