Using Meta-Analysis of Many Studies to Summarize Evidence about Program Effectiveness

Mark W. Lipsey

Peabody Research Institute, Vanderbilt

Blueprints Conference April 2010

What is evidence based practice?

There's more than one way of defining *practice*

and

corresponding differences in the *evidence* that supports it

1. The specific operating procedure (SOP) of a particular program

Example: Tumbleweed Family Counseling program in Tucson, AZ

Evidence base:

- Evaluation study of the effects of that program delivered by that provider (usually no more than one study).

2. ‘Brand name’ protocol programs

Examples: Functional Family Therapy, Multisystemic Therapy, Multidimensional Treatment Foster Care

Evidence base:

- Evaluation studies of implementations of that protocol in different places (usually only a few studies)

EBP is currently defined around 'brand name' programs

Lists of "model" programs, e.g.:

- Blueprints for Violence Prevention
- National Registry of Evidence-based Programs and Practices (NREPP)
- OJJDP Model Programs Guide

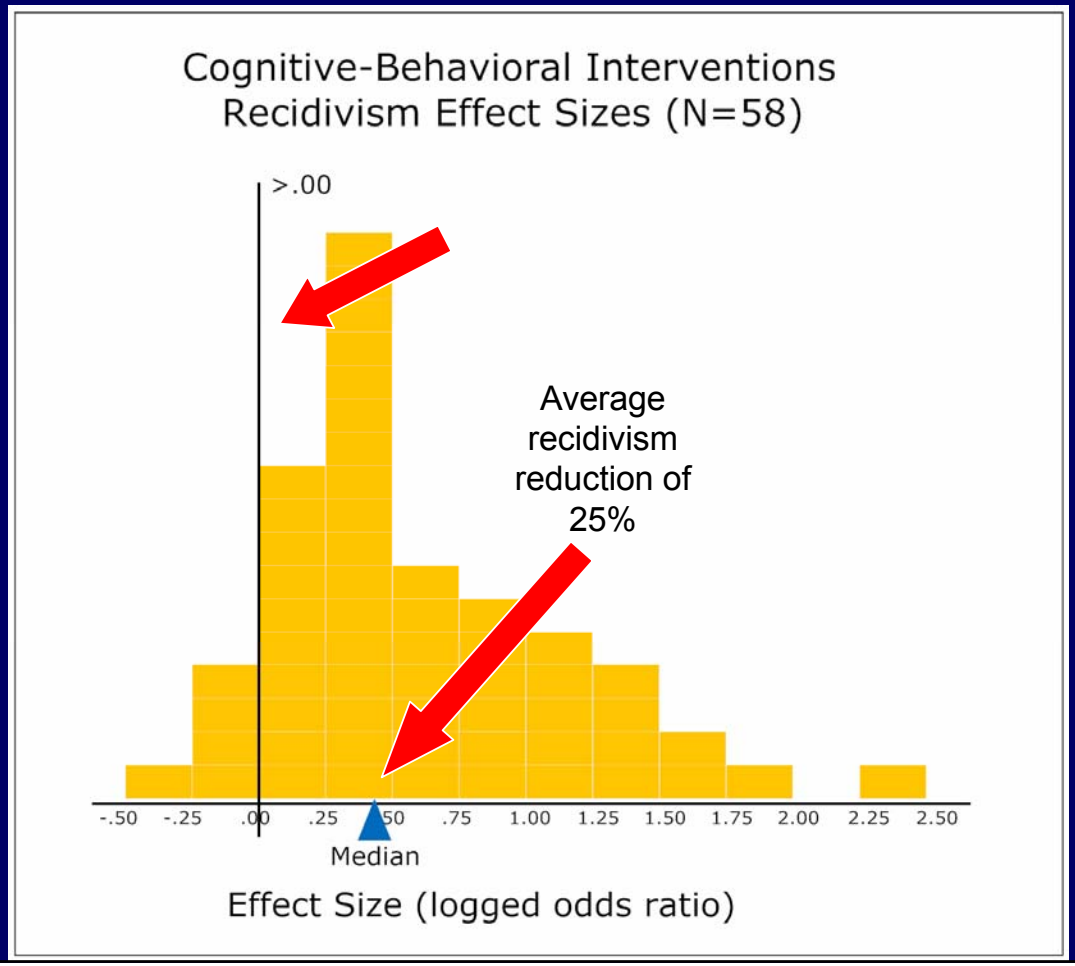
3. Generic intervention types

Examples: Interpersonal skills training, family therapy, group counseling, cognitive behavioral therapy

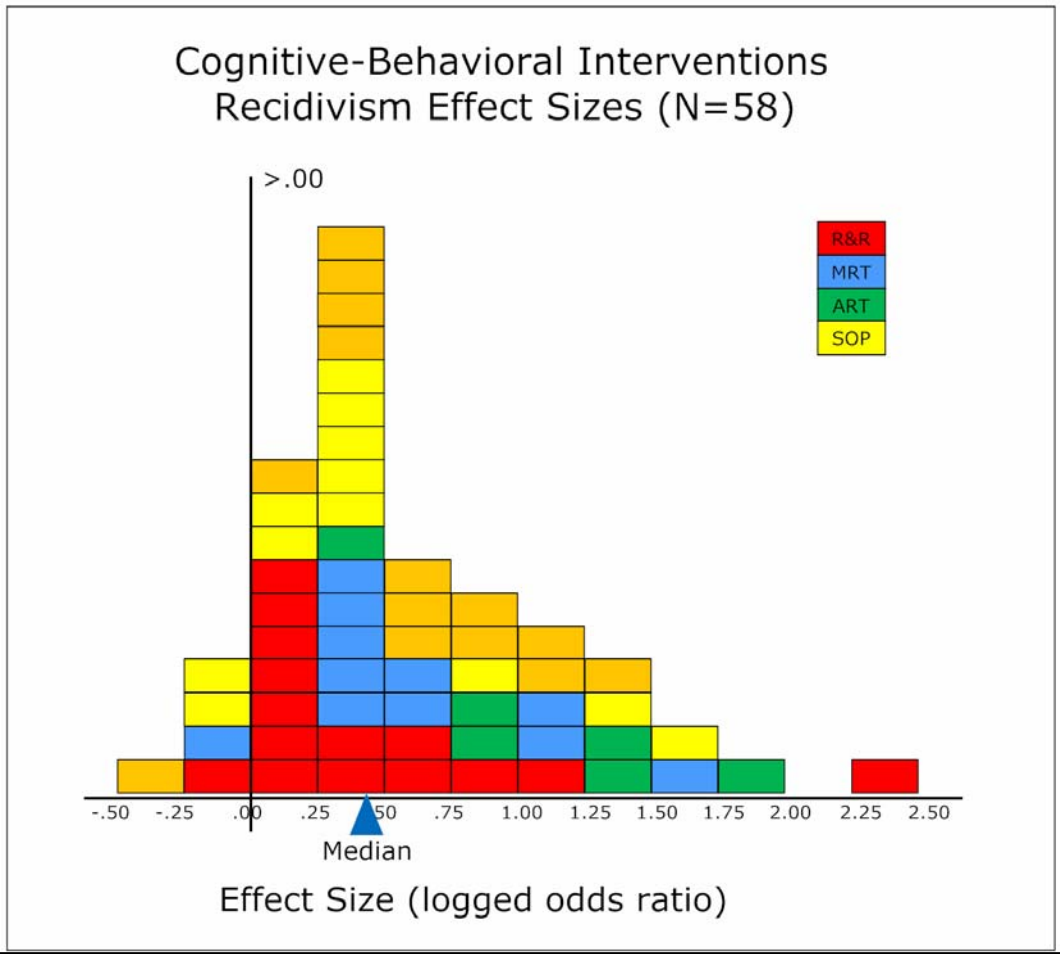
Evidence base:

- Evaluation studies of different programs of that type in different places (often many studies).

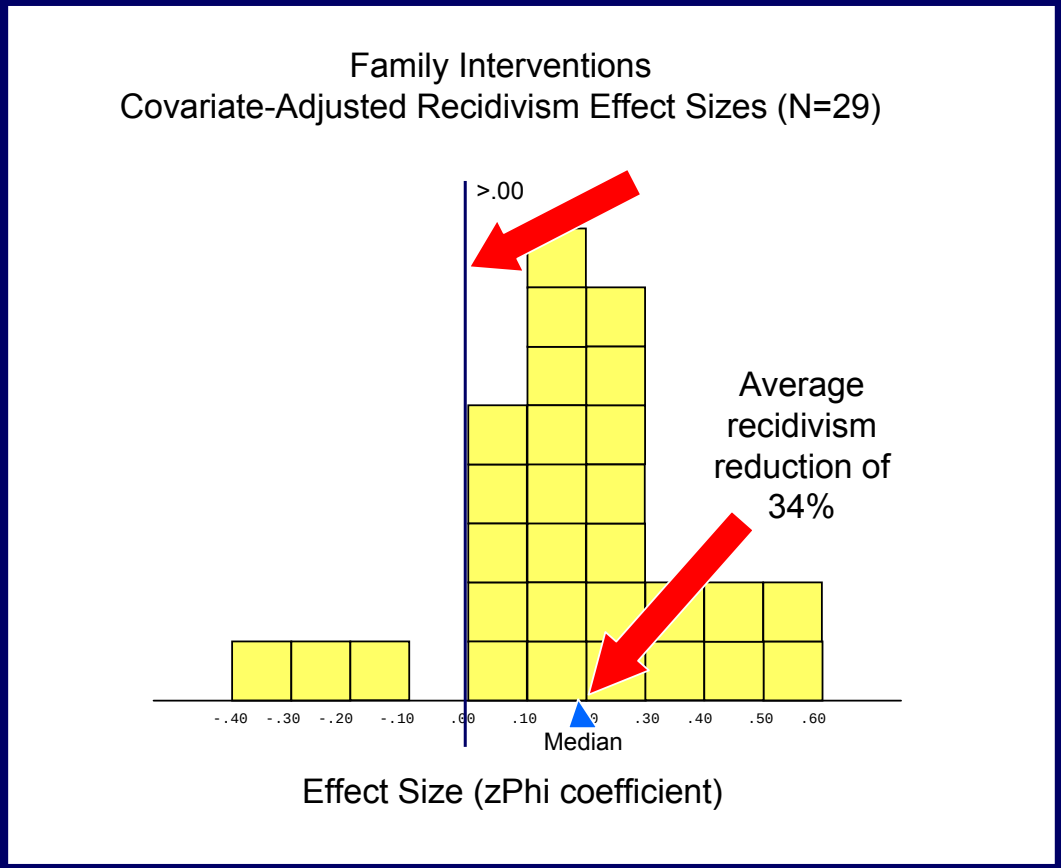
Effects for a generic intervention



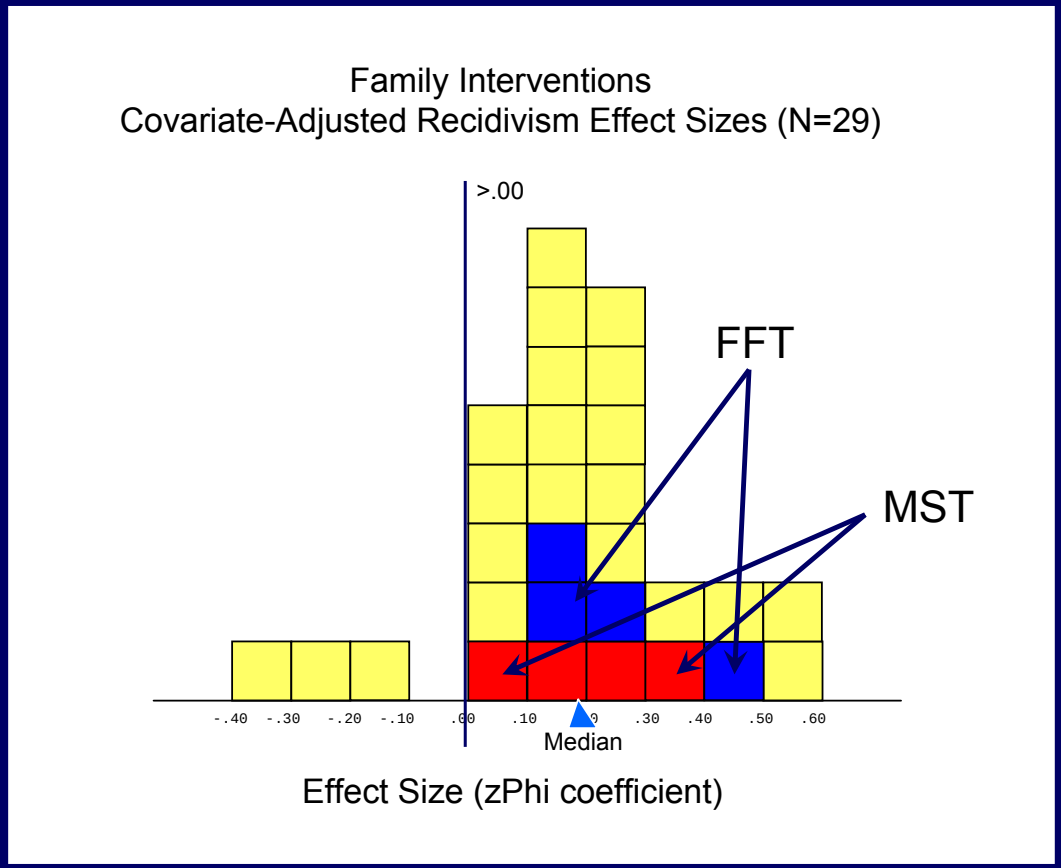
Mix of brand name & SOP programs



Programs with a Primary Family Emphasis



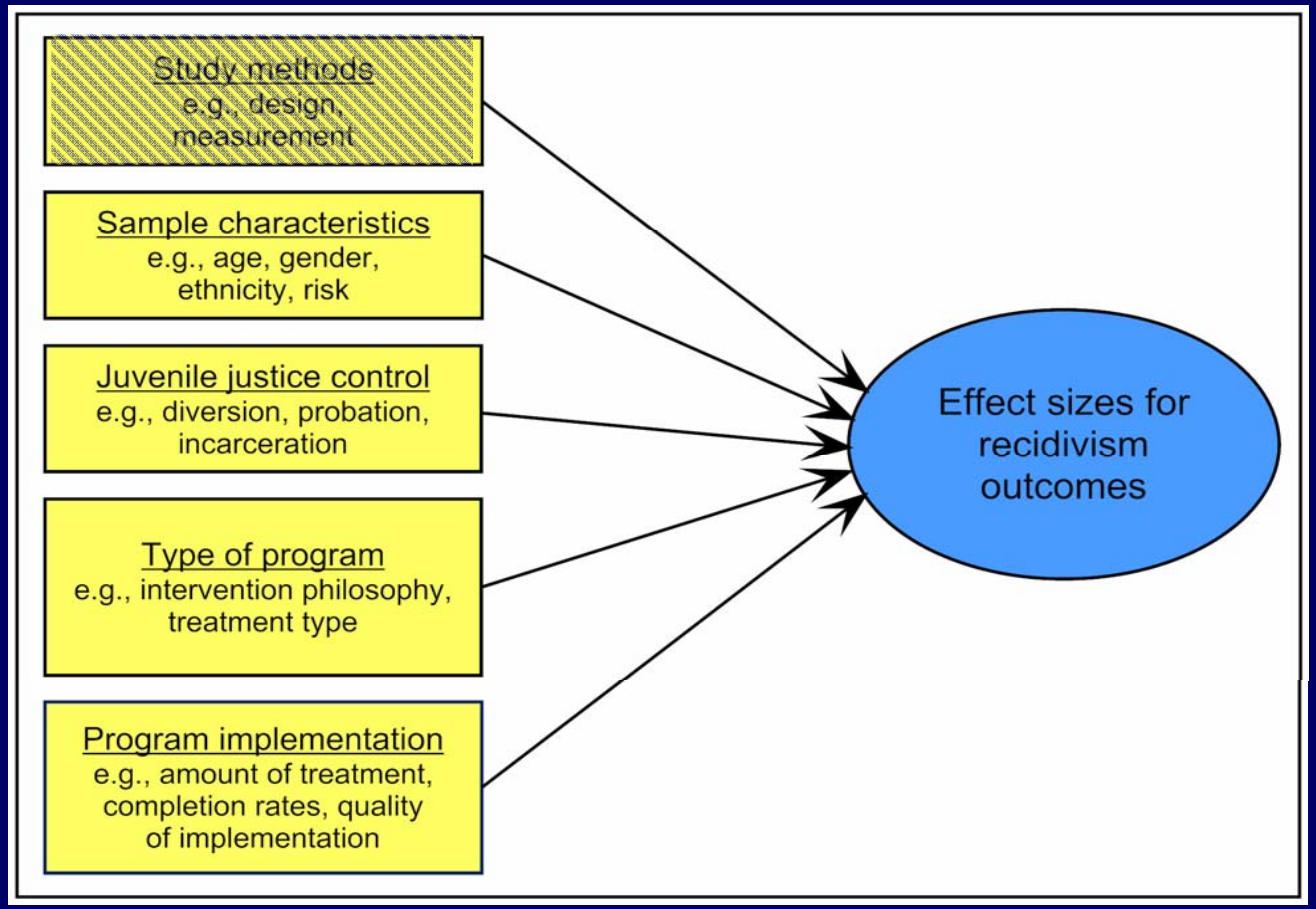
Blueprint Programs with a Primary Family Emphasis



Database of existing studies of interventions for juvenile offenders

- 548 independent study samples from 361 primary research reports
- Juveniles aged 12-21 received an intervention intended to have positive effects on their subsequent delinquency
- At least one delinquency outcome; assignment was random or, if not, pretreatment differences were reported or matched
- Conducted in English speaking countries between 1958 and 2002

Effect sizes analyzed as a function of study and program characteristics



Key Findings

Some characteristics of the juveniles matter

Effect size differences associated with:

- Delinquency risk (strong positive)
- Aggressive history (moderate negative)

Effect size differences not associated with:

- Mean age
- Gender mix
- Ethnicity

JJ supervision doesn't matter much (with risk, etc. controlled)

Effect size differences not associated with:

- No JJ supervision (prevention programs)
- Diversion
- Probation/parole
- Incarceration

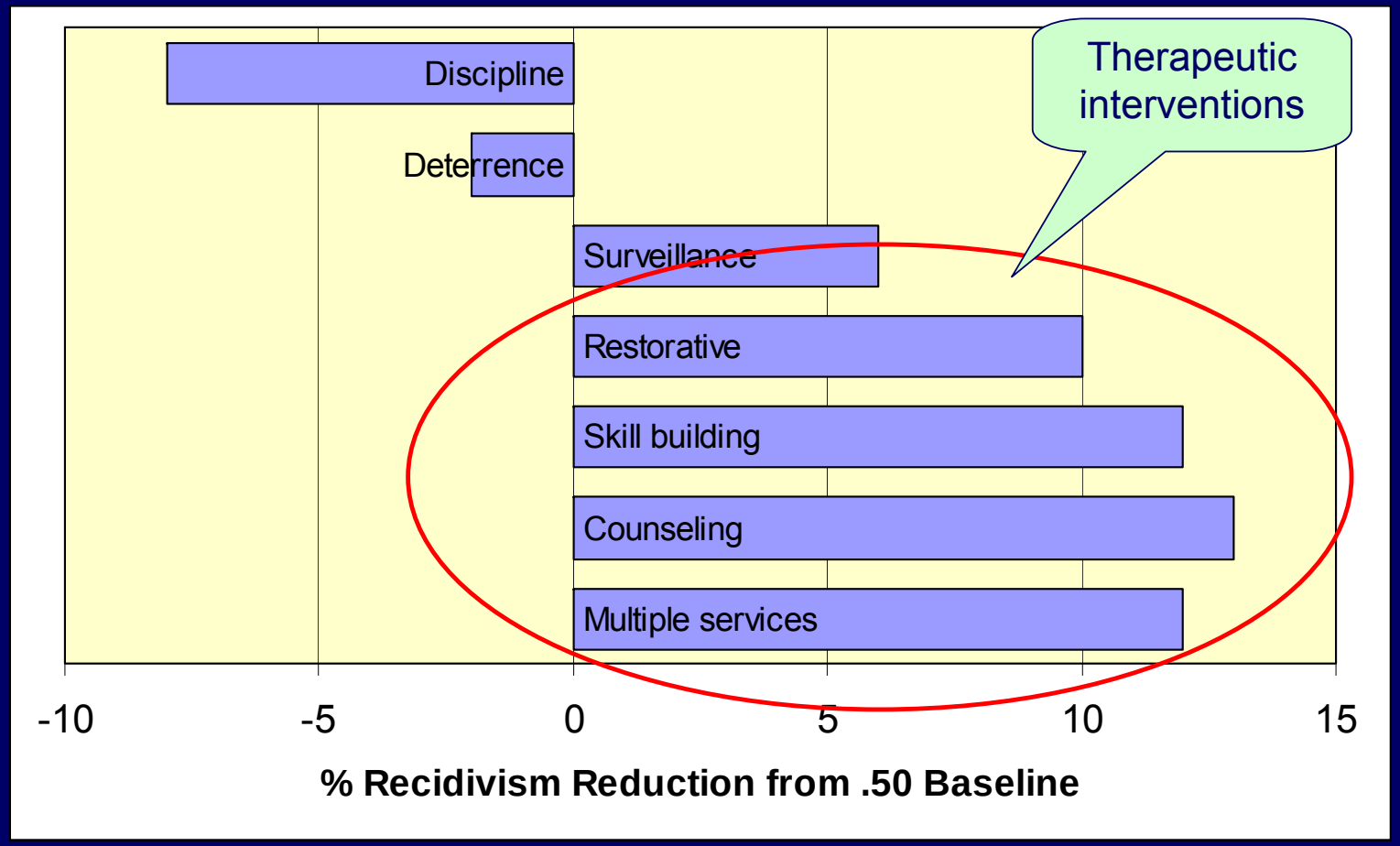
Type of program matters

Programs are identified first according to their broad approach or "philosophy:"

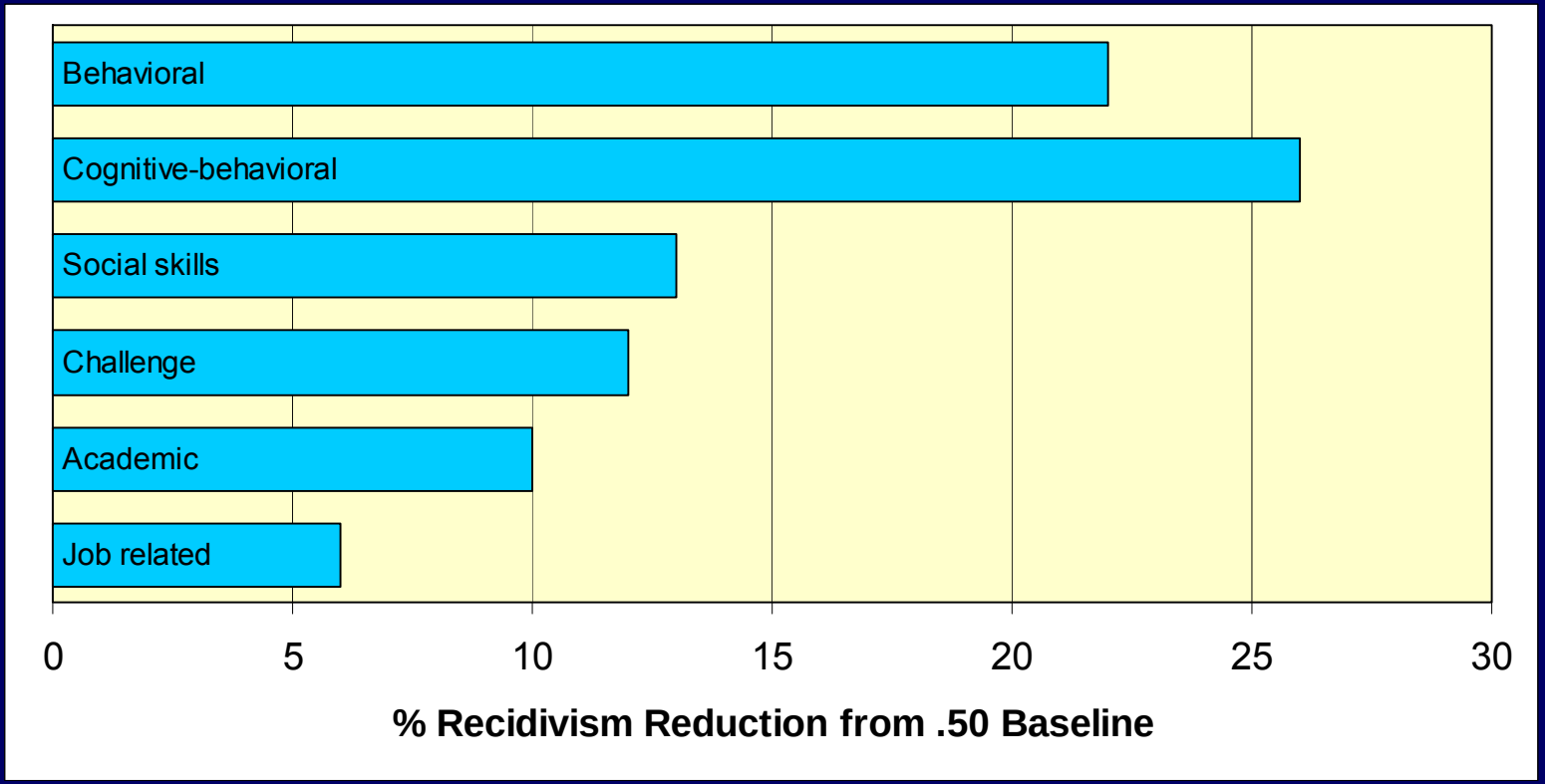
- Control philosophies
- Therapeutic philosophies

And, second, by their generic type, e.g., group counseling, interpersonal skills, cognitive behavioral therapy

Treatment “philosophy” matters



Type of intervention: Skill-building



Service amount and quality matters

Effect size differences associated with:

- Duration of service
- Total hours of service
- Quality of implementation

Summary of key findings

- Larger effects with high risk cases
- Effective interventions use a therapeutic approach
- Within a therapeutic category, some program types are more effective than others
- For a given program type, service must be delivered in adequate amounts and quality.

Operationalizing these practice guidelines

- A rating system for each program type within the therapeutic philosophies
- Applied to individual programs based on data about the services they actually provide
- Pilot projects with the juvenile justice systems of Arizona, North Carolina, & Tennessee

Points assigned proportionate to the contribution of each factor to recidivism reduction

Target values from the meta-analysis (generic) OR program manual (manualized)

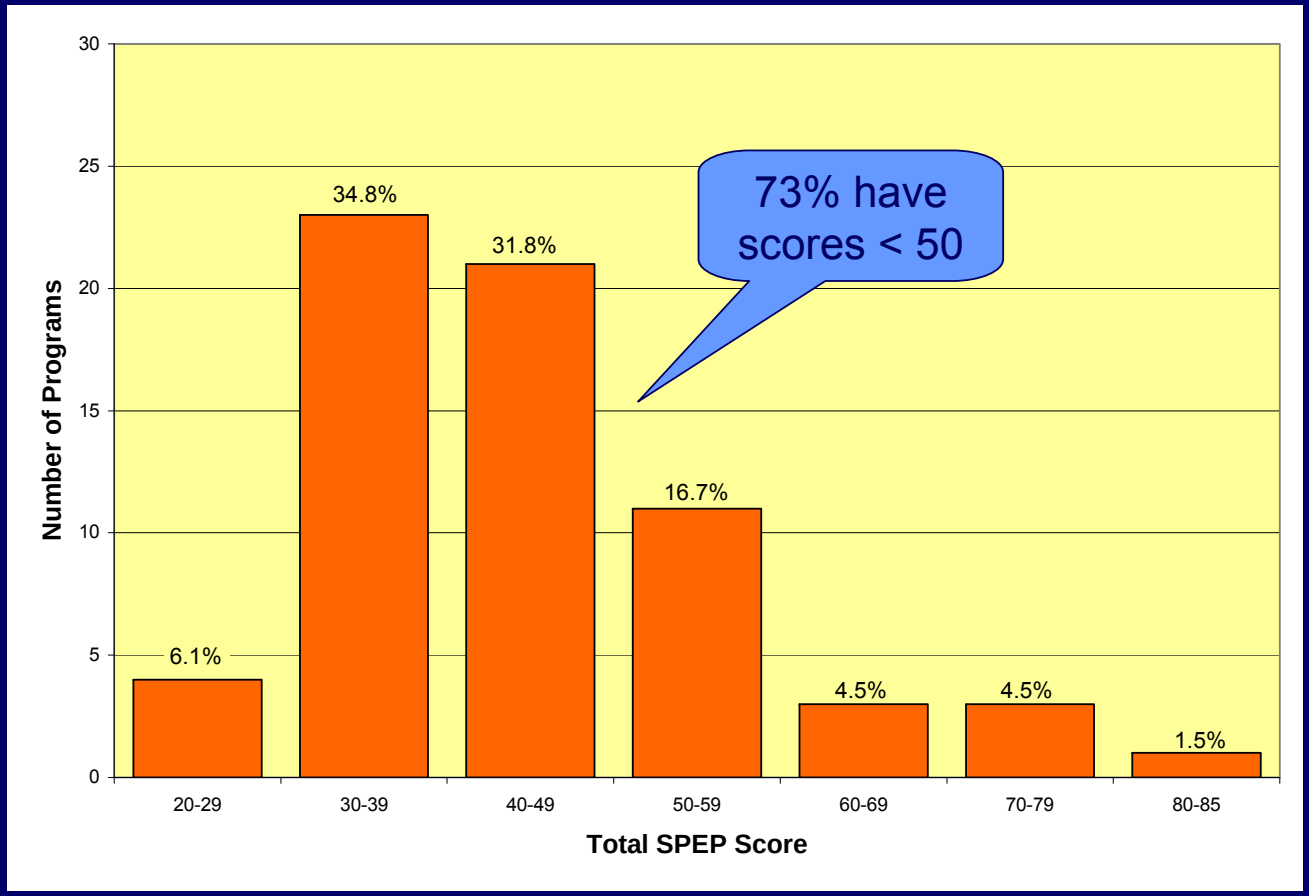
Standardized Program Evaluation Protocol (SPEP) for Services to Probation Youth

	Possible Points	Received Points
Primary Service: High average effect service (35 points) Moderate average effect service (25 points) Low average effect service (15 points)	35	
Supplemental Service: Qualifying supplemental service used (5 points)	5	
Treatment Amount: Duration: % of youth that received target number of weeks of service or more: 0% (0 points) 60% (6 points) 20% (2 points) 80% (8 points) 40% (4 points) 100% (10 points)	10	
Contact Hours: % of youth that received target hours of service or more: 0% (0 points) 60% (9 points) 20% (3 points) 80% (12 points) 40% (6 points) 100% (15 points)	15	
Treatment Quality: Rated quality of services delivered: Low (5 points) Medium (10 points) High (15 points)	15	
Youth Risk Level: % of youth with the target risk score or higher: 25% (5 points) 75% (15 points) 50% (10 points) 99% (20 points)	20	
Provider's Total SPEP Score:	100	[INSERT SCORE]

Validity study: Does it work?

- Arizona Juvenile Justice Services Division
- Programs provided during 2005-06 to juvenile probationers in five pilot counties
- 1490 juveniles who received services from 66 SPEP rated programs
- 6-month recidivism data on all; 12-month recidivism for most

Distribution of SPEP scores across programs



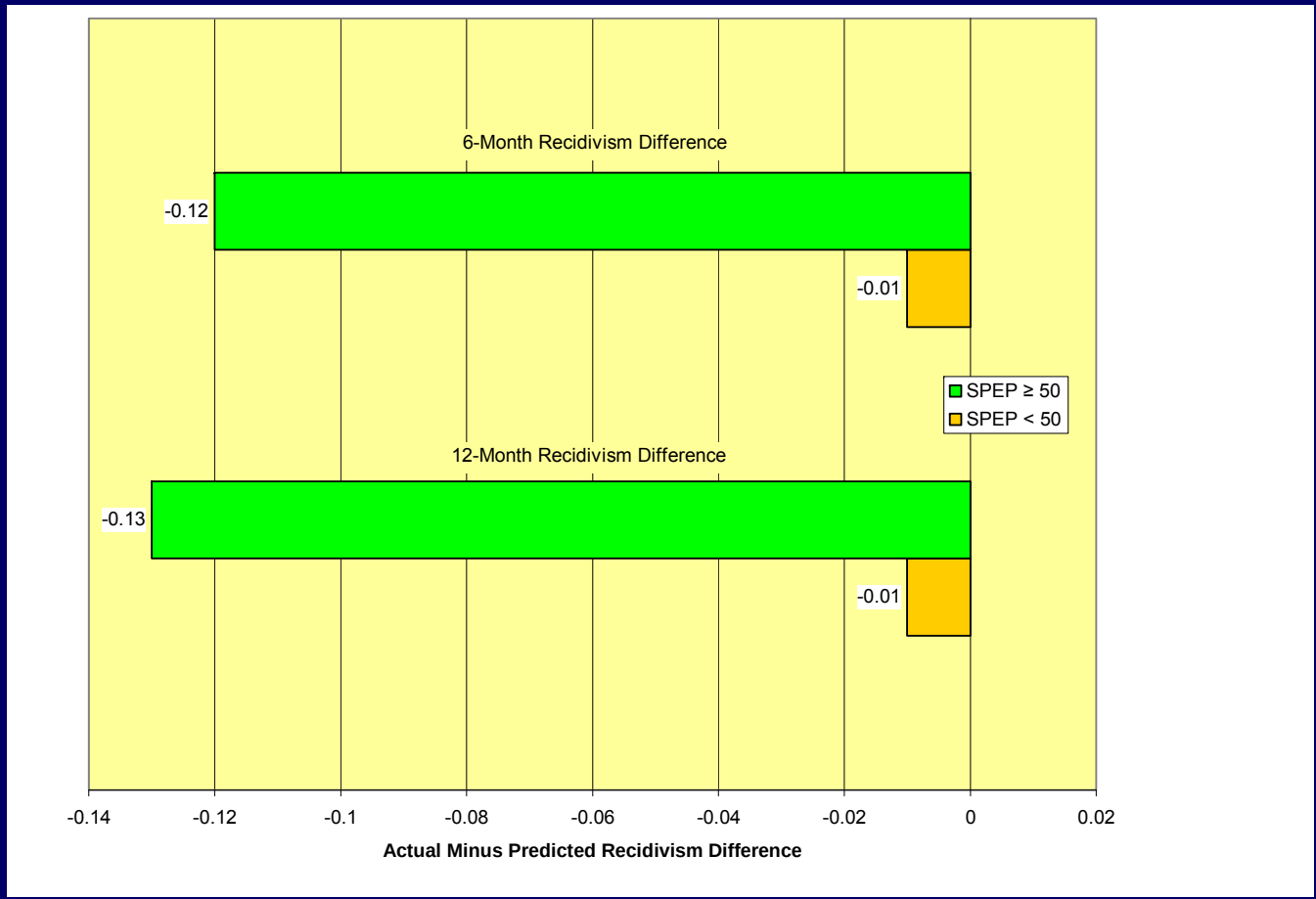
Estimate expected recidivism based on pre-existing risk factors

Recidivism predicted from archival data on:

- number and nature of prior offenses
- risk rating by probation officers
- age, sex, race, county
- number of prior service events

Actual recidivism: .27 at 6 mo, .44 at 12 mo

Actual vs. predicted recidivism for providers with scores ≥ 50 and < 50



Summary

- No one approach to EBP: There are different definitions of practice with correspondingly different bodies of evidence.
- Meta-analysis can be used to develop evidence-based practice profiles for generic interventions with wider applicability than other EBP approaches.
- Real world programs that better match these profiles do indeed show better outcomes.

Contact information:

mark.lipsey@vanderbilt.edu