A Critical Review of Evidence Based Repositories (EBPRs) in Behavioral Health

Stephen Magura, Director
Miranda Lee, Project Manager
Stephanie N. Means, Project Manager

Evaluation 2015
The Annual Conference of the American Evaluation Association
Chicago, Illinois
November 13 2015

Introduction

Overview of the studies and the papers presented here

Presenters

Dr. Stephen Magura
stephen.magura@wmich.edu

Stephanie N. Means
stephanie.means@wmich.edu

Miranda J. Lee
miranda.lee@wmich.edu

Study 1 & 2 Additional Authors

Dr. Daniela Schröter
daniela.schroeter@wmich.edu

Dr. Chris Coryn
chris.coryn@wmich.edu

Expert Advisory Panel

Katrina Bledsoe
Robert Boruch
Tina Christie
Lois-Ellin Datta
Stuart Donaldson
Jennifer Greene
George Julnes
Rebecca Kilburn
Francis Lawrenz
Michael Quinn Patton
Mitch Prinstein
Paul Roman
Roger Weissberg
Brian Yates

Main Study

Stephanie N. Means

Register Structure and Paradigms of Evidence

Funded by grant# R21DA032151 from the National Institute of Drug Abuse

Study Premise

Evidence-based registers are not well understood in their structure or the manner in which they include/exclude or rank interventions/projects.

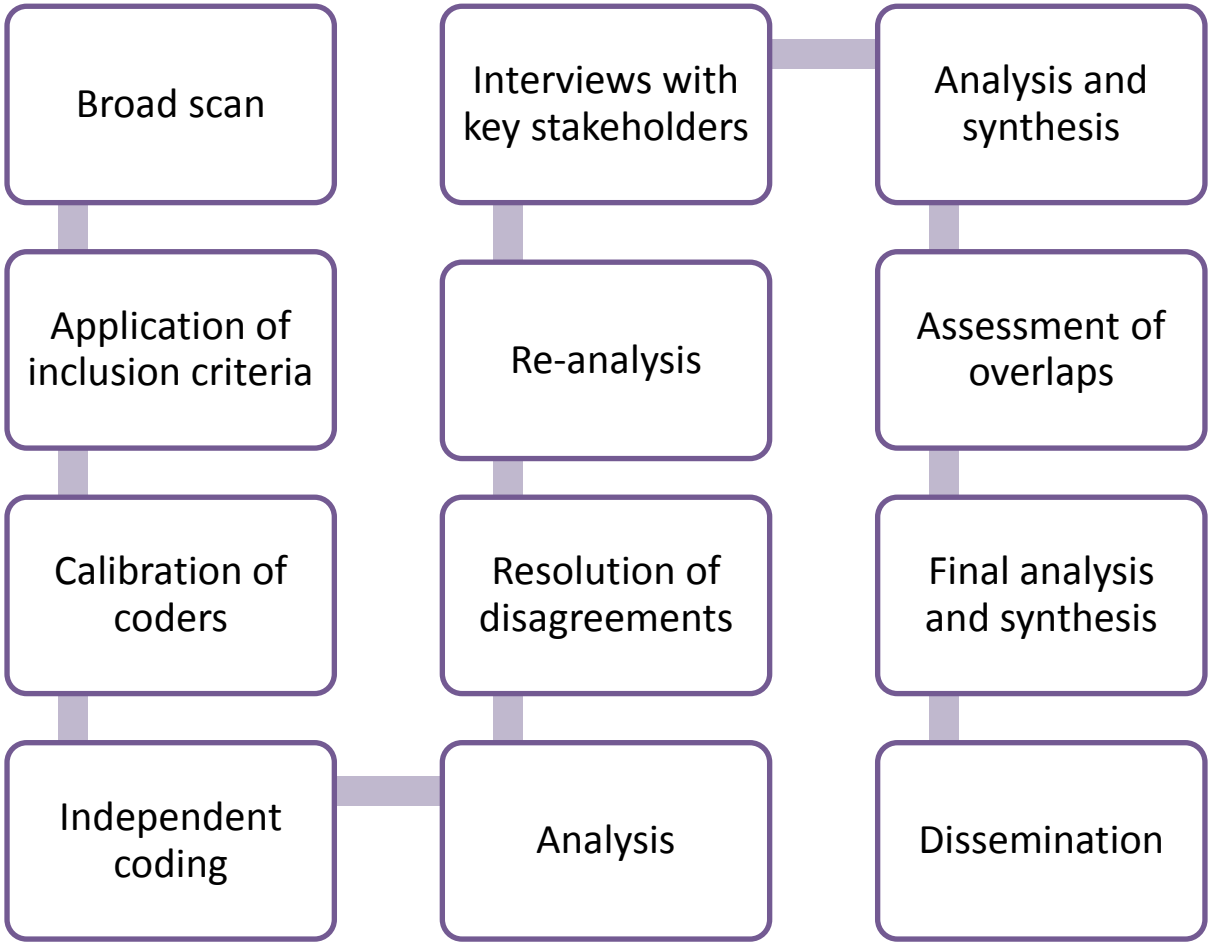
EBPR Definition

EBPRs are interactive, web-based databases of behavioral health interventions that have been evaluated against some set of criteria and standards with the purpose of supporting policy implementation or program selection

Study Aims

1. *Compile a comprehensive list of evidence-based registers for behavioral health-related interventions*
2. *Classify the registers according to their purposes, methodologies, acceptable types and standards of evidence, and other factors that are used to include and certify "effective" interventions*
3. *Determine the practical consequences of using different types and standards of evidence for including and certifying interventions as effective.*

Methods


Broad scan
Application of inclusion criteria
Calibration of coders
Independent coding
Interviews with key stakeholders
Re-analysis
Resolution of disagreements
Analysis
Analysis and synthesis
Assessment of overlaps
Final analysis and synthesis
Dissemination

Sample of registers

Select registers that appear in our study

Blueprints for Healthy Youth Development (Programs)

California Evidence-Based Clearinghouse for Child Welfare (Programs)

Cochrane Collaboration (Modalities)

CrimeSolutions.gov (Mixed)

Effective Child Therapy (Modalities)

NREPP - National Registry of Evidence-based Programs and Practices (Programs)

Social Programs that Work (Programs)

What Works Clearinghouse (Programs)

Findings

Structure of the registers

Reporting Schemas

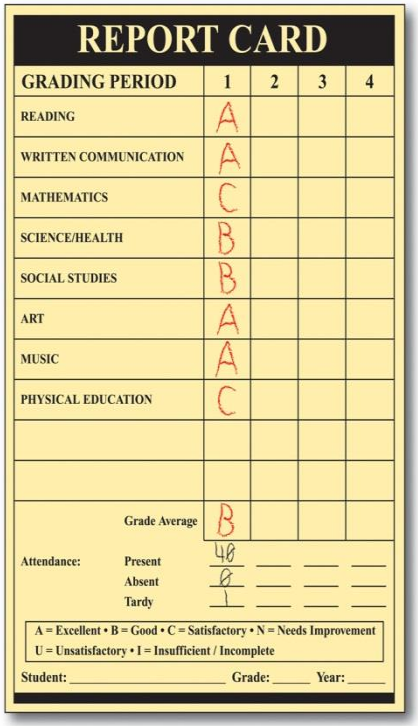
Ranking(10)

Best Practice

Promising Practice

Not Evaluated

Scoring(1)



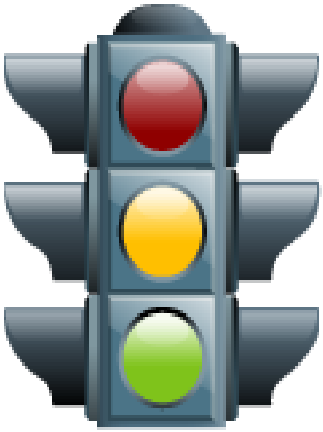
REPORT CARD

GRADING PERIOD	1	2	3	4
READING	A			
WRITTEN COMMUNICATION	A			
MATHEMATICS	C			
SCIENCE/HEALTH	B			
SOCIAL STUDIES	B			
ART	A			
MUSIC	A			
PHYSICAL EDUCATION	C			
Grade Average	B			
Attendance: Present	48			
Absent	0			
Tardy	1			

A = Excellent • B = Good • C = Satisfactory • N = Needs Improvement
U = Unsatisfactory • I = Insufficient / Incomplete

Student: _____ Grade: _____ Year: _____

Simple Inclusion/ Exclusion(7)



Tiers of Evidence

Top 2 tiers considered as “evidence-based”

Typical issues considered in the tiers

- *Design*
- *Sample size*
- *Outcomes*
- *Measures*
- *Analytical methods*

Research Design

First and Second Tiers

Top Tier

8 required at least one RCT

2 "functioning" random assignment

1 "at least" 2 RCTs and/or 1 RCT and 1 QED

1 "convincing" comparison group

1 "at least" one QED

Top Tier

5 required "at least" one RCT

2 "formal" QED

4 designs that have the essential features of a QED, but do not require formal QED

2 "minimally biased assignment"

Sample Size

First and Second Tiers

Top Tier

2 stated required minimum sample size

1 "adequate sample size/group equivalence to detect effects

1 clear statement of demographic characteristics

Second Tier

2 required minimum sample size

1 required an "adequate sample size to detect effects"

Outcomes

First and Second Tiers

There was only 1 register that specified outcome requirements

Top tier: change of 20% or more

2nd tier: change of 1% of more

Measures

First and Second Tiers

1 considered “the standard measure of effect size, the standardized mean difference” implying a cutoff for its categories

Most required that the outcomes measured be specially related to the domain of interest in the register

Analytic Methods

First and Second Tiers

Top Tier

4 placed constraints an analytic methods

2 mentioned power analysis

Second Tier

4 stated requirements for analytic methods

1 "appropriate statistical analysis based on an intent to treat model"

2 analysis of participants based on intervention exposure, based on two sided t-test

Other major differences

Reporting Bias

Sustainability of effects

Other methodological quality criteria

Additional requirements

Findings

Comparing Rating Paradigms for EBPRs

Sample

Random sample of 100 out of 355 programs contained in the registers

We assessed agreement and disagreements on program ratings between the registers

Program Rating

	Agreement		Disagreement		
	Top Tier	Lower than top	Tier placement	Effect vs. no effect	n
Two Registers	15 (33%)	6 (13%)	20 (44%)	4 (9%)	45
Three Registers	8 (33%)	0 (0%)	12 (50%)	4 (17%)	24
Four Registers	13 (68%)*	1 (5%)*	3 (16%)	2 (11%)	19
Five or More	4 (33%)*	0 (0%)*	7 (58%)	1 (8%)	12

*Agreement = More than 75% of registers agree on program rating

Program Rating

Overall, 42% of the programs were classified as disagreeing on tier placement

There were an additional 11% of the programs where one register rated the program as having no effect (or insufficient evidence) while the other registers classified the program as effective at some tier

Thus substantial disagreement among registers was identified for 53% of the programs rated by more than one register.

Summary

Some overlap between registers

Differences in what constitutes quality evidence

Differences in users and uses

Difficult to ascertain which register to prefer when making a decision

Challenges

Design

Use of qualitative evidence

Register eligibility

Strength of evidence

Addressing program costs

Assessing intervention modalities

What kind of evidence do you need to make a decision, and how much?

Is study quality or strength of findings more important for decision making?

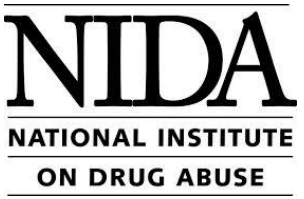
The development and use of evaluation criteria in evaluating scientific evidence

What are the barriers/facilitators to use?

What are the characteristics of evidence-based registers that facilitate use?

THANK YOU !

The Evaluation Center
Western Michigan University
Kalamazoo MI 49008-5237 USA
(269) 387-5895



Sub-study 1

Miranda J. Lee

Meta-Analysis as a Form of Evidence – Cochrane Reviews as an Example

Funded by grant# R21DA032151 from
the National Institute of Drug Abuse

Introduction

This study extends the research that was just described

Since meta-analysis is viewed as the highest form of evidence in traditional evidence hierarchies, we took a closer look at one of the most well-known databases of meta-analyses: The Cochrane Collaboration

What did we do?

We reviewed a purposive sample of 30 Cochrane Meta-Analysis reports, using several analytic dimensions

Analytic Dimensions

Types of Research Designs Included

Strength of Evidence Ratings

Number of Studies Used

Discussions of Clinical Significance

Study goals & Transportability

Other Factors

How did we pick the reports?

We identified major topic areas from the Cochrane Library Website

Then we picked the 7 most populated areas from mental health, and the three most populated areas from substance abuse

How did we pick the reports?

We identified major topic areas from the Cochrane Library Website

Then we picked the 7 most populated areas from mental health, and the three most populated areas from substance abuse

How did we code the reviews?

2 Raters each coded 15 of the studies

When one rater was uncertain about a rating, the second rater was asked to provide an opinion

What we found

Types of studies

23 of the studies were full meta-analyses, while 7 were changed to systematic reviews by the authors due to lack of poolable data.

What we found

Research designs

Type	Designs Allowed
Regular RCT	14
Cluster RCT	4
Controlled Clinical Trials	5
Quasi Experimental Designs	3
Other	3

What we found

Conclusions made by meta-analysis authors

12 studies stated definite conclusions about the effectiveness of the intervention

11 studies stated mixed results

7 studies showed no effect

What we found

Studies included in the meta-analyses

Range: 0 – 88

Mean number of studies: 16

Standard deviation: 19

What we found

Inclusion of studies in the meta-analyses

Percent of studies that made it into the final analysis

Range: <1% - 17%

Mean percent of studies: 4.4%

Standard deviation: 4.9%

What we found

Discussion of clinical significance

Virtually none of the reviews addressed practical importance.

One review converted its effect size in terms of reduced drinking days.

What we found

Quality of Evidence

Newer reviews address quality of evidence using the GRADE standards, developed by the GRADE Working Group*

Grading

Recommendations

Assessment

Development

Evaluation

*http://www.guidelinedevelopment.org/handbook/

What we found

Publication bias

While many of the reviews claimed to control for publication bias, only five (17%) studies included formal standard publication bias analysis (funnel plots, etc.)

What we found

Forest plots

For the reviews that presented Forest plots, many were presented by separate outcomes, which has advantages and disadvantages

What we found

Heterogeneity

High heterogeneity within the reviews, but also between the reviews!

I^2 is defined as being "the proportion of total variation in the estimates of the treatment effects that is due to heterogeneity rather than chance" (Cooper, Hedges, and Valentine, 2009)

25% is small heterogeneity, 50% is medium, and 75% is considered to be large.

What we found

Heterogeneity

I^2 Range = 0% – 91%

Mean I^2 = 43.2%

Standard Deviation = 35.6%

I^2 was not identifiable in 12 of the studies

Discussion

Heterogeneity

High heterogeneity was dealt with by creating subgroups.

However, the subgroups may only include one study. Is this really a viable strategy?

Discussion

Efficacy versus effectiveness

We can't assume that any individual intervention would be effective or not, solely based on the overall meta-analysis result

Discussion

Types of decisions

Lack of information about how to select one program over another –provider agencies may not simply implement a class of programming.

Discussion

Types of decisions

Funders may require provider organizations to implement programs within a particular class.

How can those funders use the Cochrane Database to inform program grant proposals?

Discussion

Policy implications

Meta-analyses are conducted post-hoc to the original research.

Do they have a role in setting policy agendas? If so, how responsive might they be?

Limitations

Small number of reviews sampled

Coding structure not validated on a new sample of reviews.

Agreement versus formal reliability study in the coding

Next steps

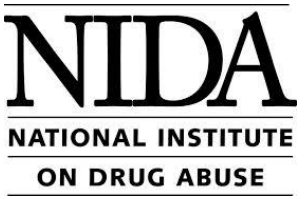
Validate coding structure

Independently code reviews and then assess reliability

Larger sample

THANK YOU !

The Evaluation Center
Western Michigan University
Kalamazoo MI 49008-5237 USA
(269) 387-5895



Studies of Evidence-Based Program Register Users

Presented by Dr. Stephen Magura

California Evidence-Based Clearinghouse for Child Welfare (CEBC)

Additional Collaborators

Cambria Rose Walsh
California Evidence-Based
Clearinghouse for Child Welfare

Jennifer Rolls-Reutz
California Evidence-Based
Clearinghouse for Child Welfare

Rady Children's Hospital
San Diego, California

Funding

Study 1: Funded by the California Evidence-Based Clearinghouse for Child Welfare

Study 2: Funded by grant 1R21DA032151 from the National Institute on Drug Abuse

Study 1

Survey of Child Welfare Professionals in California by California Evidence-Based Child Clearinghouse for Welfare (CEBC).

Background

There are virtually no systematic data on the users of evidence-based program registers in behavioral health care.

Two studies with different methodologies were conducted on users of the California Evidence-Based Clearinghouse for Child Welfare (CEBC).

Purpose

To assess knowledge and use of the CEBC website materials

Method

Targeted web-based survey of child welfare administrators and supervisors in the 58 counties in California

Objective - survey the director, 5 senior administrators and 5 front line supervisors in each county public child welfare (protective services) agency

Conducted during 2011 (unpublished)

Results

Survey response

67% of the counties were represented (n=39)

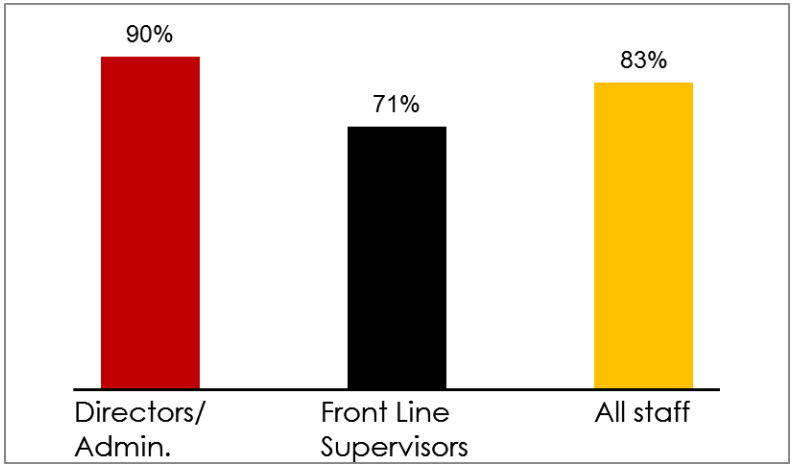
66% of the nominated staff members completed the surveys (n=160)

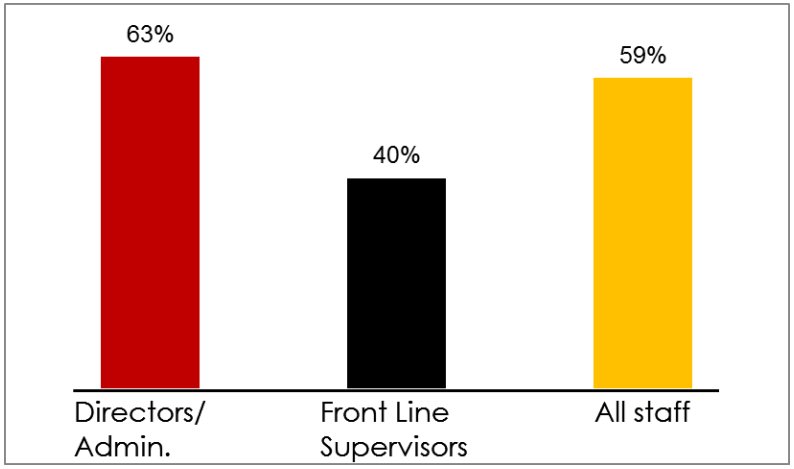
Within the range of the best reported results for on-line surveys of targeted samples (Monroe and Adams, 2012)

Results

Main findings

A majority of respondents have heard about the CEBC



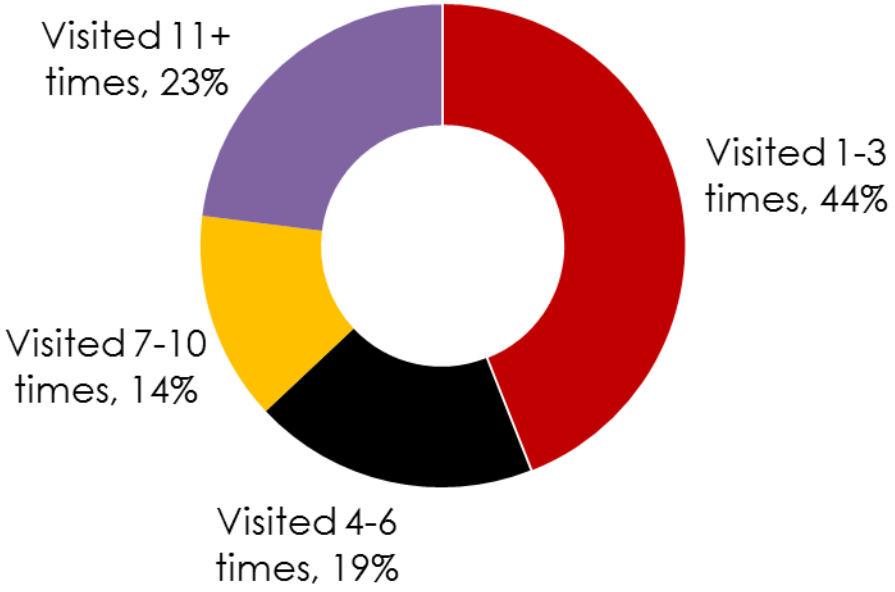


They also have visited the website at least once.

Results

Main findings

The majority of people visited more than 3 times



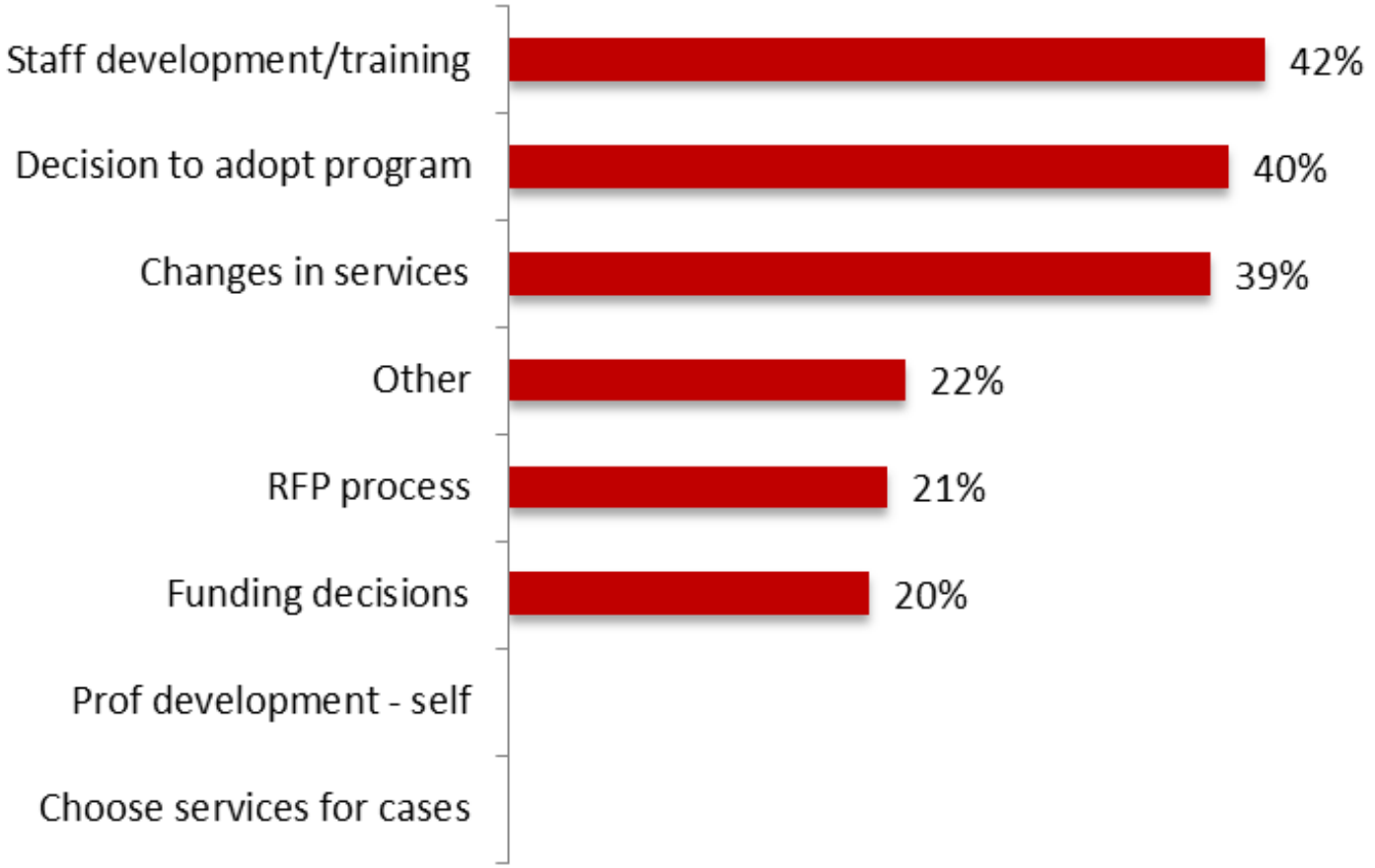
Results

Main findings

68% of Directors/Administrators and 53% of Front-Line Supervisors have used CEBC materials at work.

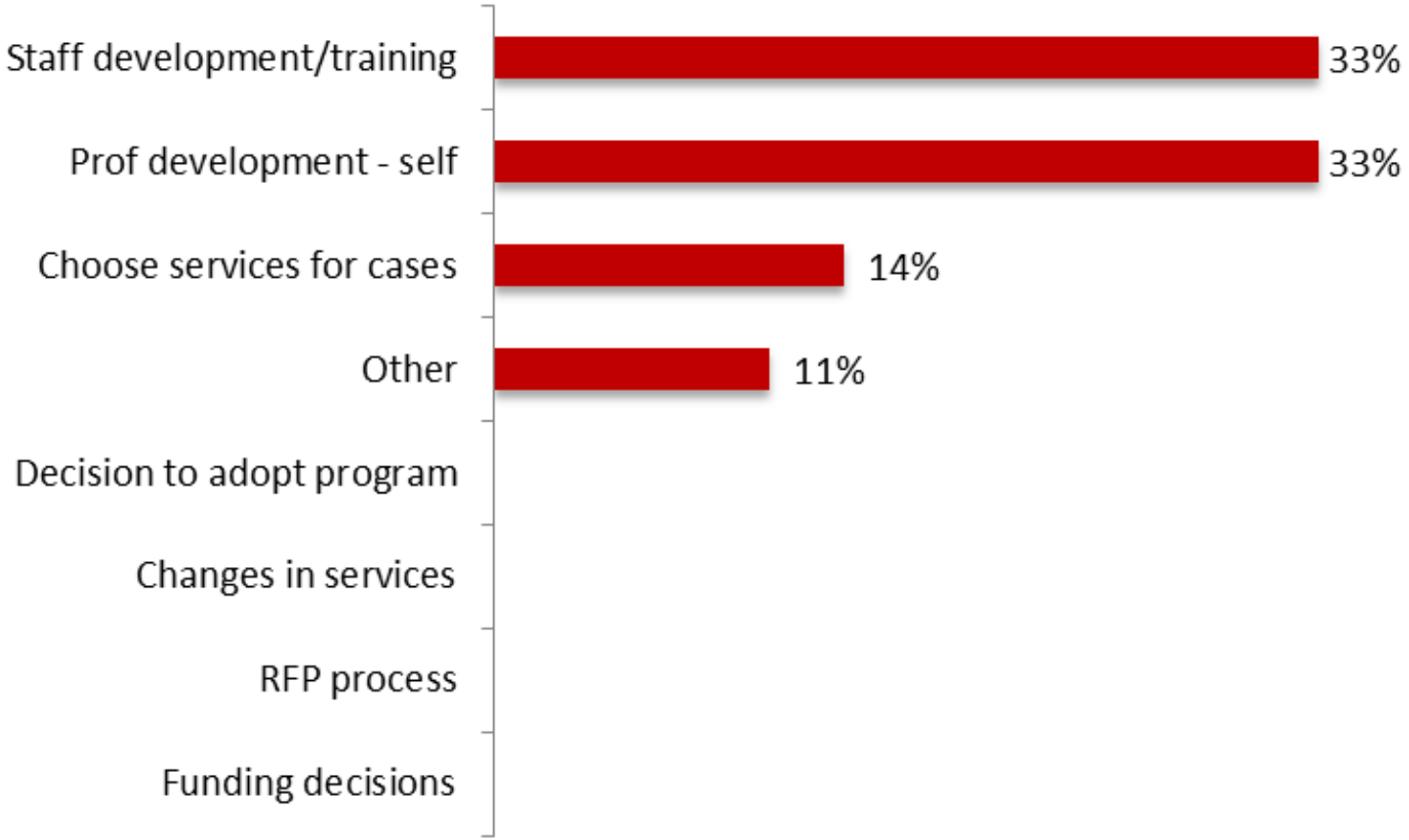
Results

Use by directors & administrators



Results

Use by front-line staff



Findings

Staff Comments

One pager on the CEBC would be useful for general distribution to families, advocates etc.

CEBC used to learn more about a program or intervention of interest (rather than finding a new program)

Movement is towards funding "evidence-based programs" (EBPs) but many definitions exist

CEBC definition of EBP is clearer than many others

CEBC includes not only info on programs, but also info on how to monitor program implementation, fidelity measures etc.

Discussion

There have been efforts to publicize the CEBC within the California child welfare system, including professional conferences and webinars sponsored by the CEBC; the survey results should be interpreted in that context.

Child Welfare administrators and front line supervisors were generally familiar with the CEBC and the majority had visited the web site, most several times.

A web-based EBPR can be a frequently-used source of information on evidence-based programs for professionals, given that the site is adequately publicized.

Discussion

However, details on the types and extent of changes that actually resulted for agencies or consumers could not be determined through this type of survey.

Since only child welfare professionals were queried, the full range of types of visitors to the site and their experiences was not determined.

Study 2

Pilot Study of Visitors of the California Evidence-Based Clearinghouse for Child Welfare (CEBC).

Purpose

To test a method of obtaining individualized feedback from visitors to the CEBC website.

Method

Pop-up window with invitation to participate on a 20 minute telephone interview – e-mail address requested to schedule

$20 incentive offered

Invitations extended during two one-week periods in Aug 2014 to those spending 3+ minutes on the site

Results

60 eligible visitors provided e-mails

Interviews were scheduled and completed with 25 visitors

Main Findings

General

Three main categories of users - direct service (50%), administrators (41%) and researchers (9%)

Organizations represented included public and private service providers, county and state level governments, universities, and private practices

Main Findings

Information needs

Most of respondents reported finding the information they needed.

Need of direct service providers to address particular problems using proven strategies

For those who help set policies, the typical need was to vet programs that were already in place for accountability purposes.

Main findings

Provider agency uses

Certify programs they employed as being evidence-based

Support grant writing activities for future funding

Selection of program models

Identify strategies for service delivery

Increase capacity for sound assessment practices.

Main findings

Useful Information

Basic information about the programs

Program ratings

Help on cultural competence issues

Criteria used in the rating process

Main findings

What users couldn't find

Negative evaluation findings for programs

Programs available in the local area ("resource list")

Transferability of program ratings to different types of clients

Information about several popular client assessments

Main findings

Concerns expressed

Reliability and validity of the program rating scales

Rationale for rating certain programs was not always clear, especially for programs listed as "not rated."

Several respondents did not understand the ratings or had not even viewed them

Generally, there was a mix of opinions about the usefulness of the program ratings

Main Findings

Impact of the CEBC

Did PRIOR VISITS to the site lead to any concrete changes in their organization?

Recommendations for organizational changes had been made

Some clients had changed their behavior or otherwise benefited as a result of CEBC-recommended interventions

Most of the respondents believed they would be able to use the information obtained on the site

Some said their use would be contingent on the project they were working on

Most also reported that they would visit the website again

Discussion

The telephone interview protocol was effective in eliciting relevant responses from respondents

The CEBC website is functional and generally useful to those visitors who were motivated to provide feedback

The user types were from a wide range of groups with a diverse set of needs, and as such, the CEBC is faced with the ongoing choice of how to clearly respond such diverse needs

Discussion

For some users, the scientific rating of programs is only one piece of desired information and more attention needs to be paid to feasibility and implementation

More could be done to address the needs of private citizens and community organizer-type users

This pilot study indicates the utility of a broader study of EBPRs with larger samples of site visitors, especially prospectively follow-up on the results of visits

THANK YOU !

The Evaluation Center
Western Michigan University
Kalamazoo MI 49008-5237 USA
(269) 387-5895