

To: Colorado Department of Human Services (CDHS)

From: Elysia Clemens, Deputy Director, Colorado Evaluation and Action Lab

**Date:** February 17, 2020

Subject: Colorado FFPSA Technical Review Submission for Colorado Community Response

(CCR)

• Independent reviewers Sara Bayless and Maggie Schultz Patel, independent contractors working on behalf of the Colorado Lab, found *low support of causal evidence* for the Community Response program. This means that the program does not currently meet the criteria to request transitional payments.

- After conducting a comprehensive literature review, reviewers identified two eligible studies: (1) a study of the Wisconsin Community Response program conducted by researchers from Ohio State University and University of Wisconsin-Stevens Point, and (2) a study of the Colorado Community Response program conducted by Colorado State University and the Kempe Center.
- The Wisconsin study¹ received a rating of low causal evidence based on a lack of available pretest data, pre-test alternatives, and demographic data (race/ethnicity or socioeconomic status). Because of this, baseline equivalence could not be established. This can be a challenge for evaluation of prevention programs that use screen-outs as a comparison group.
- The Colorado Community Response (CCR) study<sup>2</sup> also received a rating of low causal evidence as a result of not meeting statistical model standards. Reviewers concluded that the statistical models used were invalid in their current iteration. The major issues arose with respect to (1) the outcome analyses, which used McNemar's tests, and (2) the propensity score matching procedures (PSM); these issues were interrelated. McNemar's is a paired samples test that was, in this case, applied to independent samples. Although Austin (2010) suggests that PSM samples should be treated as paired samples, that assertion is based on the idea that "treated and untreated subjects matched on the propensity score will have observed baseline covariates that come from the same multivariate distribution" (p. 1293). However, given that PSM procedures were in this case based on the somewhat unconventional categorical treatment of continuous data (e.g., age of youngest child as 1 year old or less v. 2 years or older) and differences between the intervention and control groups were observed for some baseline variables in the effect size ranges of .11-.19,

<sup>&</sup>lt;sup>1</sup> Marathon County Community Response: Voluntary Services for Families Screened out of Child Protective Services. Maguire-Jack, K., & Bowers, J.

https://www.researchgate.net/publication/313373504 Marathon County Community Response Voluntary Services for Families Screened out of Child Protective Services

<sup>&</sup>lt;sup>2</sup> Colorado Department of Human Services Colorado Community Response Final Evaluation Report 2014-2018. Allen, H., Currie, D., Drury, I., Merkel-Holguin, L., & Fluke, J. <a href="http://earlychildhoodframework.org/wp-content/uploads/2018/10/CCR-Evaluation-Report-FINAL-DRAFT-2014-2018.pdf">http://earlychildhoodframework.org/wp-content/uploads/2018/10/CCR-Evaluation-Report-FINAL-DRAFT-2014-2018.pdf</a>

this suggests that the subjects do not come from the same multivariate distribution. Ultimately, this mismatch leaves reviewers with concerns about achieving the statistical model standards.

- Reviewers noted that there may be an opportunity to address the concerns regarding statistical
  model standards if additional analyses could be conducted on the dataset (e.g., with continuous
  data demonstrating baseline equivalence and/or outcome analyses using independent samples
  tests and accounting for baseline differences). The Colorado Lab recommends discussing this
  possibility with the authors of this CCR study directly.
- Please consider the findings from the technical review in the context that the Prevention Services Clearinghouse Standards were not published at the time the evaluations were conducted.
- Moving forward, the Colorado Lab recommends that CDHS require evaluators to pre-register
  their analytic plans for causal evaluation (e.g., Open Science Framework) and that these plans are
  proactively reviewed—prior to beginning the study—for alignment with the Clearinghouse
  Standards.

# Attachment B: Checklist for Program or Service Designation for HHS Consideration

### Instructions:

**Section I:** The state must complete Section I (Table 1) once to summarize all of the programs and services that the state reviewed and submitted and the designations for HHS consideration.

Section II: The state must complete Section II (Tables 2 and 3) once to describe the independent systematic review methodology used to determine a program or service (listed in Table 1) designation for HHS consideration. Section II outlines the criteria for an independent systematic review. To demonstrate that the state conducted an independent systematic review consistent with sections 471(e)(4)(C)(iii)(I), (iv)(I)(aa) and (v)(I)(aa) of the Act, the state must answer each question in the affirmative. If the independent systematic review used the Prevention Services Clearinghouse Handbook of Standards and Procedures, the relevant sections must be indicated in the "Handbook Section" column. If other systematic standards and procedures were used, states must submit documentation of the standards and procedures used to review programs and services. States should determine the standards and procedures to be used prior to beginning the independent systematic review process. If the state cannot answer each question in Table 2 and Table 3 in the affirmative, ACF will not make transition payments for the program or service reviewed by the state using those standards and procedures.

**Section III:** The state must complete Section III (Tables 4 and 5) for each program or service listed in Table 1, and provide all required documentation. Section III outlines the requirements for the review of the program or service. States should complete Table 4 prior to conducting an independent systematic review to determine if a program or service is eligible for review. For a program or service to be eligible for review, the answer to both questions in Table 4 must be affirmative and the state must provide the required documentation. If a program or service is eligible for review, the state must conduct the review and identify each study reviewed in Table 5, regardless of whether a study was determined to be eligible to be included in the review.

**Section IV:** The state must complete Section IV (Tables 6-10) for each program or service (listed in Table 1) reviewed and submitted and provide all required documentation. Section IV lists studies the state determined to be "well-designed" and "well-executed" and outlines characteristics of those studies. Do <u>not</u> include eligible studies that were <u>not</u> determined to be "well-designed" and "well-executed" in Tables 6 -10. States should complete Table 6 with a list of all eligible studies determined to be "well-designed" and "well-executed." States should complete Table 7 to describe the design and execution of each eligible "well-designed" and "well-executed" study. States should complete Table 8 to describe the practice setting and study sample. States must answer in the affirmative that the program or service included in each study was not substantially modified or adapted from the version under review. States must detail favorable effects on target outcomes present in eligible studies determined to be "well-designed" and "well-executed." States must detail unfavorable effects on target and non-target outcomes present in eligible studies determined to be "well-designed" and "well-executed."

**Section V:** The state must complete Section V (Table 11) for each program or service reviewed and submitted. Section V lists the program or service designation for HHS consideration and verification questions relevant to that designation. The state must answer the questions applicable to the relevant designation in the affirmative.

# Section I: Summary of Programs and Services Reviewed and their Designations for HHS Consideration

# **Section I. Summary of Programs and Services Reviewed**

# Table 1. Summary of Programs and Services Reviewed

To be considered for transitional payments, list programs and services reviewed and provide designations for HHS consideration.

Program or Service Name (if there are multiple versions, specify the specific version reviewed)	Proposed Designations for HHS consideration (Promising, Supported, or Well-Supported)
Community Response	Not Rated – Low Causal Evidence

# Section II: Standards and Procedures for an Independent Systematic Review

# Section II. Standards and Procedures for a Systematic Review

(Complete Table 2 and Table 3 to provide the requested information on the independent systematic review. The same standards and procedures should be used to review all programs and services.)

### Table 2. Systematic Review

Sections 471(e)(4)(C)(iii)(I), (iv)(I)(aa) and (v)(I)(aa) of the Act require that systematic standards and procedures must be used for all phases of the review process. In the table below, verify that systematic (i.e., explicit and reproducible) standards and procedures were used and submit documentation of reviewer qualifications. If the systematic review used the Prevention Services Clearinghouse Handbook of Standards and Procedures, indicate the relevant sections in the "Handbook Section" column. If other systematic standards and procedures were used, submit documentation of the standards and procedures.

	☐ to Verify	Handbook Section
Were the same systematic standards and procedures used to review all programs and services?	X	
Were qualified reviewers trained on systematic standards and procedures used to review all programs and services?	X	
Were standards and procedures in accordance with section 471(e) of the Social Security Act?	X	
Were standards and procedures in accordance with the Initial Practice Criteria published in Attachment C of <u>ACYF-CB-PI-18-09</u> ?	X	
Program or Service Eligibility: Were systematic standards and procedures used to determine if programs or services were eligible for review? At a minimum, this includes standards and procedures to:	X	2.1
<ul> <li>Determine if a program or service is a mental health, substance abuse, in-home parent-skill based, or kinship navigator program; and</li> </ul>	X	2.1.1
<ul> <li>Determine if there was a book/manual or writing available that specifies the components of the practice protocol and describes how to administer the practice.</li> </ul>	X	2.1.2
Literature Review: Were systematic standards and procedures used to conduct a comprehensive literature review for studies of programs and services under review? At a minimum, this includes standards and procedures to:	X	3
Search bibliographic databases; and Search other sources of publicly available	X	3
<ul> <li>Studies (e.g., websites of federal, state, and local governments, foundations, or other organizations).</li> </ul>	X	3
Study Eligibility: Were systematic standards and procedures used to determine if studies found through the comprehensive literature review were eligible for review? At a minimum, this includes standards and procedures to:	X	4.1
<ul> <li>Determine if each study examined the program or service under review (as described in the book/manual or writing) or if it examined an adaptation;</li> </ul>	X	4.1 & 2.1.2
Determine if each study was published or prepared in or after 1990;	X	4.1.1 & 4.1.2
Determine if each study was publicly available in English;	X	4.1.3
<ul> <li>Determine if each study had an eligible design (i.e., randomized control trial or quasi- experimental design);</li> </ul>	X	4.1.4
Determine if each study had an intervention <i>and</i> appropriate comparison condition;	X	4.1.4
<ul> <li>Determine if each study examined impacts of program or service on at least one 'target' outcome that falls broadly under the domains of child safety, child permanency, child well-being, or adult (parent or kin-caregiver) well-being. Target</li> </ul>	X	4.1.5

outcomes for kinship navigator programs can instead or also include access to, referral		
to, and satisfaction with services; and		
Identify studies that meet the above criteria and are eligible for review.	X	4.1
Study Design and Execution: Were systematic standards and procedures used to determine if eligible studies were well-designed and well-executed? At a minimum, this includes standards and procedures to:	X	5
Assess overall and differential sample attrition;	X	5.6 – N/A (QED)
<ul> <li>Assess the equivalence of intervention and comparison groups at baseline and whether the study statistically controlled for baseline differences;</li> </ul>	X	5.7, 5.7.1-5.7.3 & WWC
Assess whether the study has design confounds;	X	5.9.3
<ul> <li>Assess, if applicable, whether the study accounted for clustering (e.g., assessed risk of joiner bias<sup>1</sup>);</li> </ul>	X	5.5 – N/A (QED)
Assess whether the study accounted for missing data; and	X	5.9.4
Determine if studies meet the above criteria and can be designated as well-designed and well-executed.	X	5
Defining Studies: Sometimes study results are reported in more than one document, or a single document reports results from multiple studies. Were systematic standards and procedures used to determine if eligible, well-designed and well-executed studies of a program and service have non-overlapping samples?	X	4.1
Study Effects: Were systematic standards and procedures used to examine favorable and unfavorable effects in eligible, well-designed and well-executed studies? At a minimum, this includes standards and procedures to:  • Determine if eligible, well-designed and well-executed studies found a favorable effect (using conventional standards of statistical significance) on each target outcome; and  • Determine if eligible, well-designed and well-executed studies found an unfavorable		
effect (using conventional standards of statistical significance) on each target or non- target outcome.		
Beyond the End of Treatment: Were systematic standards and procedures used to determine the length of sustained favorable effects beyond the end of treatment in eligible, well-defined and well-executed studies? At a minimum, this includes standards and procedures to:		
Identify (and if needed, define) the end of treatment; and		
Calculate the length of a favorable effect beyond the end of treatment.	N/A because	no eligible contrasts
Usual Care or Practice Setting: Were systematic standards and procedures used to determine if a study was conducted in a usual care or practice setting?	were deter	mined to be well- nd well-executed.
<i>Risk of Harm:</i> Were systematic standards and procedures used to determine if there is evidence of risk of harm?		
Designation: Were systematic standards and procedures used to designate programs and services for HHS consideration (as promising, supported, well-supported, or does not currently meet the criteria)? At a minimum, this includes standards and procedures to:		
<ul> <li>Determine if a program or service has one eligible, well-designed and well-executed study that demonstrates a favorable effect on a target outcome and should be considered for a designation of promising;</li> </ul>		
<ul> <li>Determine if a program or service has at least one eligible, well-designed and well- executed study carried out in a usual care or practice setting that demonstrates a favorable effect on a target outcome at least 6 months beyond the end oftreatment and should be considered for a designation of supported; and</li> </ul>		
Determine if a program or service has at least two eligible, well-designed and well- executed studies with non-overlapping samples carried out in usual care or practice		

<sup>&</sup>lt;sup>1</sup>If a cluster randomized study permits individuals to join clusters after randomization, the estimate of the effect of the intervention on individual outcomes may be biased if individuals who join the intervention clusters are systematically different from those who join the comparison clusters.

settings that demonstrate favorable effects on a target outcome; at least one of the studies must demonstrate a sustained favorable effect of at least 12 months beyond the end of treatment on a target outcome; and should be considered for a designation of well-supported.

Reconciliation of Discrepancies: Were systematic standards and procedures used to reconcile discrepancies across reviewers? (applicable if more than one reviewer per study)

Author or Developer Queries: Were systematic standards and procedures used to query study

authors or program or service developers? (applicable if author or developer queries made)

## Table 3. Independent Review

The systematic review must be independent (i.e., objective and unbiased). In the table below, verify that an independent review was conducted using systematic standards and procedures by providing the names of each state agency and external partner that reviewed the program or service. States must answer all applicable questions in the affirmative. Submit MOUs, Conflict of Interest Policies, and other relevant documentation.

List all state agencies and external partners that reviewed programs and services.	
Colorado Evaluation and Action Lab:	
Sara Bayless	
Maggie Schultz Patel	
	<b>5</b>
	☐ to Verify
Was the review independent (conducted by reviewers without conflicts of interest including those that	X
authored studies, evaluated, or developed the program or service under review)?	
Was a Conflict of Interest Statement signed by reviewers attesting to their independence? If so, attach the	X
statement.	
Was a Memorandum of Understanding (MOU) signed by external partners (if applicable)? If so, attach MOU(s).	X

# Sections III-V: Describe and Document Findings from Each Program and Service Reviewed and Submitted

# Section III. Review of Programs and Services (Complete Tables 4-5 for each program or service reviewed.)

# Table 4. Determination of Program or Service Eligibility

Fill in the table below for each program or service reviewed.

	☐ to Verify
Does the program or service have a book, manual, or other available documentation specifying the	X
components of the practice protocol and describing how to administer the practice?	
Provide information about how the book/manual/other documentation can be accessed OR provide supporting availability of book/manual/other documentation.	e other information
Manuals for Colorado Community Response and Wisconsin Community Response are both public narrative for supporting information.	ly available. See
Is the program or service a mental health, substance abuse, in-home parent-skill based, or kinship navigator program or service?	X
Identify the program or service area(s).	
Community Response is a comprehensive case management child welfare prevention program.	

# Table 5. Determination of Study Eligibility

Fill in the table below for each study of the program or service reviewed. Provide a response in every column; N/A or unknown are not acceptable responses. The response in columns iii, v, vi, vii, and ix must be "yes" or "no." The response in column ix is "yes" only when the responses in columns iii, v, vi, and vii are "yes."

i. Study Title/Authors  Colorado	ii. Publicly Available Location	iii. Is the stud y in Engli sh? (Yes/ No)	iv. Design (RCT, QED, or other). If other, specify design.	v. Did the intervention condition receive the program or service under review in accordance with the book/manual/documentation? (Yes/No)	vi. Did the comparison condition receive no or minimal intervention or treatment as usual? (Yes/No)	vii. Did the study examine at least one target outcome? (Yes/No)	viii. Year Published	ix. Eligible for Review? (Yes/No)
Department of Human Services Colorado Community Response Final Evaluation Report 2014-2018. Allen, H., Currie, D., Drury, I., Merkel-Holguin, L., & Fluke, J.	http://earlychildhoodfram ework.org/wp- content/uploads/2018/10 /CCR-Evaluation-Report- FINAL-DRAFT-2014- 2018.pdf		QED	res	Yes	res	2018	1 es
Marathon County Community Response: Voluntary Services for Families Screened out of Child Protective Services. Maguire- Jack, K., & Bowers, J.	https://www.research gate.net/publication/ 313373504_Marathon County Community Response Voluntary Services for Families Screened out of Ch ild Protective Service S	Yes	QED	Yes	Yes	Yes	2013	Yes

# Section IV. Review of "Well-designed" and "Well-executed" Studies (Complete Tables 6-10 for each program or service reviewed.)

Table 6. Studies that are "Well-Designed" and "Well-Executed"<sup>2</sup>

Provide an electronic copy of each of the studies determined to be eligible for review and determined to be "well-designed" and "well-executed."

List all eligible studies that are "well-designed" and "well-executed" (Study Title/Author)
Within the two studies eligible for review, none of the contrasts were determined to be well-designed and well-executed. For the study of Wisconsin Community Response, there were no pre-test data available (as indicated by authors in the report). For the study of Colorado Community Response, the study was assessed as not meeting statistical model standards (5.9.1; see the accompanying review narrative for more details). These determinations were made for both studies and thus affect all potentially eligible contrasts within the studies.

<sup>&</sup>lt;sup>2</sup> For reference, the Prevention Services Clearinghouse Handbook Chapter 5 defines "well-designed" and "well-executed" studies as those that meet design and execution standards for high or moderate support of causal evidence. Prevention Services Clearinghouse ratings apply to contrasts reported in a study. A single study may have multiple design and execution ratings corresponding to each of its reported contrasts.

# Table 7. Study Design and Execution - N/A

For each study eligible for review and determined to be "well-designed" and "well-executed," fill out the table below. Provide a response in every column; N/A or unknown are not acceptable responses for columns i, ii, iii, v, vi, and vii. The response in column ii must be "yes."

i. Study Title/Authors	ii. Verify the Absence of all Confounds? (Yes/No)	iii. List Measures that Achieved Baseline Equivalence	iv. List Measures that did NOT Achieve Baseline Equivalence but were Statistically Controlled for in Analyses	v. Overall Attrition <sup>3</sup> (for RCTs only)	vi. Differential Attrition <sup>4</sup> (for RCTs only)	vii. Does Study Meet Attrition Standards?	viii. Notes, as needed
Example Title. Smith,	Yes	-Center for Epidemiologic	-Income	2.0 percent	4.3 percentage points	Yes	N/A
A.B., Jones,		Studies		,	,		
C.D., and		Depression Scale					
Doe, E.F.		(CES-D)					
		-Child Behavior Checklist (CBCL)					

<sup>&</sup>lt;sup>3</sup> For reference, the Prevention Services Clearinghouse Handbook section 5.6 defines overall attrition as the number of individuals without post-test outcome data as a percentage of the total number of members in the sample at the time that they learned the condition to which they were randomly assigned.

<sup>&</sup>lt;sup>4</sup> For reference, the Prevention Services Clearinghouse Handbook section 5.6 defines *differential attrition* as the absolute value of the percentage point difference between the attrition rates for the intervention group and the comparison group.

# Table 8. Study Description - N/A

For each study eligible for review and determined to be "well-designed" and "well-executed," fill out the table below to describe the practice setting and study sample as well as affirm that the program or service evaluated was not substantially modified or adapted from the version under review. Provide a response in every column; N/A or unknown are not acceptable responses. The response in column v must be "yes."

i. Study Title/Autho rs	ii. Was the study conducted in a usual care or practice setting? (Yes/No)	iii. What is the study sample size?	iv. Describe the sample demographics and characteristics of the intervention group	v. Describe the sample demographics and characteristics of the comparison group	vi. Verify that the program or service evaluated in the study was <i>NOT</i> substantially modified or adapted from the manual or version of the program or service selected for review (Yes/No)
Example Title. Smith, A.B., Jones, C.D., and Doe, E.F.	Yes	N=354  Caregivers, N = 177  Child, N = 177	Caregivers – Average age = 37 years old (SD = 5 years); 95% female; 35% Black or African American, 25% White, 30% Latino or Hispanic, and 10% other; and 78% of households living 200% below the federal poverty level.  Children – Average age = 5 years old (SD=1.3 years); 47% female; 37% Black or African American, 27% White, 32% Latino or Hispanic, and 4% other.	Caregivers – Average age = 35 years old (SD = 5 years); 93% female; 33% Black or African American, 26% White, 31% Latino or Hispanic, and 10% other; and 76% of households living 200% below the federal poverty level.  Children – Average age = 5 years old (SD=1.4 years); 45% female; 34% Black or African American, 28% White, 33% Latino or Hispanic, and 4% other.	Yes

# Table 9. Favorable Effects – N/A

For each study eligible for review and determined to be "well-designed" and "well-executed," fill out the table below listing only target outcomes with **favorable effects**. Provide a response in every column; N/A or unknown are **not acceptable** responses.

i. Study Title/Authors	ii. List the Target Outcome(s)	iii. List the Outcome Measures	iv. List the Reliability Coefficients for Each	v. Are Each of the Outcome Measures Valid?	vi. Are Each of the Outcome Measures Systematically Administered?	vii. List the P-Values for Each of the Outcome Measures	viii. List the Size of Effect for Each of the Outcome Measures	ix. Indicate the Length of Effect Beyond the End of Treatment (in months)
Example Title. Smith, A.B., Jones, C.D., and Doe, E.F.	Parent/Caregiver Mental Health (Depression)	CES-D	Cronbach's alpha coefficient = 0.91	Yes	Yes	p = 0.04	d = 0.13	8 mos
	Child Behavioral and Emotional Functioning (Externalizing Behaviors)	CBCL (Aggressive Behavior Scale)	Cronbach's alpha coefficient = 0.94	Yes	Yes	p = 0.03	d = 0.24	0 mos
	Child Behavioral and Emotional Functioning (Internalizing Behaviors)	CBCL (Anxious/Depressed Scale)	Cronbach's alpha coefficient = 0.84	Yes	Yes	p = 0.23 (non-sig)	N/A	0 mos

# Table 10. Unfavorable Effects – N/A

For each study eligible for review and determined to be "well-designed" and "well-executed," fill out the table below listing only target outcomes with unfavorable effects. Provide a response in every column; N/A or unknown are not acceptable responses.

i. Study Title/Authors	ii. List the Target or Non-Target Outcome(s)	iii. List the Outcome Measures	iv. List the Reliability Coefficients for Each	v. Are Each of the Outcome Measures Valid?	vi. Are Each of the Outcome Measures Systematically Administered?	vii. List the P-Values for Each of the Outcome Measures	viii. List the Size of Effect for Each of the Outcome Measures	ix. Indicate the Length of Effect Beyond the End of Treatment (in months)
Example Title. Smith, A.B., Jones, C.D., and Doe, E.F.	Adult Height	Inches	Cronbach's alpha coefficient = 0.99	Yes	Yes	p = 0.047	d = -0.05	0 mos

# Section V. Program or Service Designation for HHS Consideration

Table 11. Program or Service Designation for HHS Consideration - N/A

Fill out the table below for the program or service reviewed. Only select one designation. Answer questions relevant to the selected designation; relevant questions must be answered in the affirmative.

	☐ to Verify
There is <b>NOT</b> sufficient evidence of risk of harm such that the overall weight of evidence does not support the benefits of the program or service.	
	☐ the Designation and Provide a Response to the Questions Relevant to that Designation
Well-Supported	
<ul> <li>Does the program or service have at least two eligible, well-designed and well-executed studies with non-overlapping samples<sup>5</sup> that were carried out in a usual care or practice setting?</li> </ul>	
<ul> <li>Does one of the studies demonstrate a sustained favorable effect of at least 12 months beyond the end of treatment on at least one target outcome?</li> </ul>	
Supported	
<ul> <li>Does the program or service have at least one eligible, well-designed and well-executed study that was carried out in a usual care or practice setting and demonstrate a sustained favorable effect of at least 6 months beyond the end of treatment on at least one target outcome?</li> </ul>	
Promising	
<ul> <li>Does the program or service have at least one eligible, well-designed and well-executed study and demonstrate a favorable effect on at least one 'target outcome'?</li> </ul>	

<sup>&</sup>lt;sup>5</sup>Samples across multiple sources of a study are considered overlapping if the samples are the same or have a large degree of overlap. Findings from an eligible study determined to be "well-executed" and "well-designed" may be reported across multiple sources including peer-reviewed journal articles and publicly available government and foundation reports. In such instances, the multiple sources would have overlapping samples. The findings across multiple sources with these overlapping samples should be considered <u>one</u> study when designating a program or service as "well-supported," "supported," and "promising."