Ohio Alternative Response Evaluation Extension

Interim Report

to the
Ohio Supreme Court

Prepared by
L. Anthony Loman, PhD and Gary L. Siegel, PhD
Institute of Applied Research
St. Louis, Missouri

January 2012

Copyright © 2012 by the Institute of Applied Research
Institute of Applied Research
103 W. Lockwood, Suite 200
St. Louis, MO 63119
(314) 968-9625
www.iarstl.org

Executive Summary

This is the interim report of the extended evaluation of the Ohio Alternative Response (AR) demonstration. The evaluation is being conducted by the Institute of Applied Research and is funded under a Ohio State Supreme Court contract with the American Humane Association. The evaluation is ongoing through late 2013 and includes three tasks: a continuation of the original analysis of the impacts of AR, a continuation of the cost analysis and an updated analysis of worker responses to AR. The present report considers interim findings on the first of these, using follow-up data from Ohio SACWIS through September 2011.

The experimental and control groups were reanalyzed using the latest SACWIS data confirming that county assignments were correct and that the groups were equivalent for target report allegations, race and other family demographics, various background family characteristics, history in child welfare and other variables. This analysis reconfirmed the validity of the field experiment.

After determining families that were appropriate for AR, each county utilized a web-based random assignment program. Families assigned to the experimental group were provided with a family assessment while those assigned to the control group received a traditional investigation.

Findings

1. Analysis of subsequent reports revealed a significant reduction in the proportions of any new screened-in reports of child maltreatment among experimental families that had no previous reports before being assigned to the study. This included more than half the families in the study. Proportional hazards analysis confirmed this finding.

2. A second analysis of subsequent reports focused on cumulative reports, counting all new reports and comparing the amounts per 100 experimental and control families. This showed that control families had 68/100 during the follow-up period compared to 60/100 for experimental families, a statistically significant difference, including reductions in both unsubstantiated and substantiated investigations of experimental families.

    a. A sub-analysis showed that reassignment to AR family assessments upon receiving a screened-in report varied significantly between counties, with Franklin County, which accounted for three in every ten study families, offering AR three times as often to experimental families.

3. A further analysis of cumulative screened-in reports revealed that control families received significantly more reports but that wide differences existed among counties in the levels of later reports on families. Because the differences were generally consistent

for both experimental and control families *within* each county, the best explanation lay in different levels of risk of study families arising from pathway assignment differences among the counties, as discussed in the April 2010 AR evaluation report.

4. It was found that AR appeared to be particularly effective among African American families in Ohio. An explanation was sought looking for differences in the level of services to families, the socio-economic status of families and previous contacts of families with child protection. African American families in this study did not differ substantially from Caucasian families on any of these variables and they were set aside as explanations. The most likely source of the difference was thought to lie in a combination of increased basic anti-poverty services coupled with the family-friendly and supportive approach that characterized family assessments.

5. The types of allegations of later reports were also examined. Reductions were observed in physical abuse, sexual abuse and emotional maltreatment allegations. However, only the first of these included enough reports and a large enough difference between experimental and control families to be statistically significant. Thus, it was concluded that AR leads to reduced reporting, and by implication reduced instances, of physical abuse in families.

6. The original difference in subsequent child removals that was reported among the April 2010 findings continued to be present but only at a lower level of one percent, which was not large enough to be statistically significant. A surprising finding was revealed in a sub-analysis, however. About 10 percent of experimental and control families had experienced a child removal and placement prior to entering the study during the 2008 to 2009 period. In most cases reunification had occurred. Among this smaller group of families AR appeared to produce large differences, with an estimated reduction of children being removed and placed of about 50 percent.

1 Introduction

This is the interim report of the extended evaluation of the Ohio Alternative Response (AR) demonstration. The study is being conducted by the Institute of Applied Research (IAR) and is funded under a Ohio State Supreme Court contract with the American Humane Association. The extended evaluation has three general goals. First, the impact analysis that was begun in the original report is being extended for an additional three years. Second, the cost analysis is to be updated, and third, the analysis of worker responses to AR will be extended and updated. The present report is focused on the first of these tasks.

The demonstration officially began in ten Ohio counties on July 1, 2008 after several months of planning and with the assistance of the American Humane Association. IAR conducted the independent evaluation of the demonstration. Under AR, each county did a second assessment of screened-in reports of child maltreatment using a *pathway assignment tool* to determine to which of two paths the report should be assigned: an investigation or a family assessment. During the period from July 2008 through September 2009, counties submitted each case that had been determined to be appropriate for a family assessment to a web-based program designed by IAR that randomly assigned families to either an experimental or control group. Experimental families received a family assessment while control families were provided with a traditional forensic investigation (which in Ohio has generally also been termed an *assessment*).

The time period for the evaluation was relatively short in that it was necessary to conduct data analysis during the period from December 2009 through March 2010. This permitted a maximum follow-up of less than 16 months on the first families that entered, which diminished to less than two months for families that entered during the last months of random assignment. The hypotheses of the evaluation regarding long-term outcomes were that experimental families would experience reduced new accepted reports of child maltreatment and fewer subsequent removals and placements of children. In order to allow sufficient follow-up time for outcomes to appear the analysis was limited to two groups of families: those that entered the study during the first three quarters (July 2008-March 2009) and those that entered during the first full year (July 2008-June 2009). Nonetheless, the results of the final analysis were positive and promising, as documented in the April 2010 final evaluation report. For this reason the state determined that a follow-up study was needed that would include all experimental and control families and a longer follow-up period extending through 2013.

SACWIS data extractions were redesigned and improved to permit a fuller analysis of data. The design was finalized in November 2010 and the first new data set was received in early January 2011. The data files were cumulative and included all data from the time the AR demonstration project began in July 2008 to the present as well as information that had been imported from earlier pre-SACWIS data systems prior to that date. Additional information was included from new tables and fields designed to accommodate AR procedures, personnel records that had not been part of the original extractions and tables that added detail on intakes, persons and out-of-home placements. The present analysis extends the follow-up through September 2011.

2 The Experimental and Control Groups

The complete sample will be referred to as the *full sample*. This consisted of 2,285 experimental and 2,244 control families. Random assignment occurred separately in each county, and as is evident in Table 1, approximately equal proportions of families fell into the experimental and control group in each county. Generally, the smaller the county total, the more variation between experimental and control. However, the statewide proportions were very nearly equivalent.

Table 1. Percent of families in final experimental and control groups in each Ohio county in the AR evaluation

County	Control	Experimental	Total
Clark	49.4%	50.6%	520
Fairfield	45.0%	55.0%	360
Franklin	50.3%	49.7%	1374
Greene	55.1%	44.9%	374
Guernsey	48.5%	51.5%	169
Licking	49.0%	51.0%	259
Lucas	51.2%	48.8%	609
Ross	48.0%	52.0%	179
Trumbull	47.6%	52.4%	538
Tuscarawas	44.2%	55.8%	147
Total	49.5%	50.5%	4529

As is evident in the rightmost column in Table 1, however, the contribution of individual counties to the statewide total varied significantly. This is shown graphically in Figure 1.

Franklin County (Columbus) was responsible for over three of every ten families in the study. For this reason, outcomes in Franklin County had a potent effect on outcomes statewide. On the other hand, very small counties (such as Guernsey or Ross) may have done better or worse than Franklin or the other larger counties, such as Lucas (Toledo), but the effects may have been concealed in the statewide totals. Because of the variation in county sample sizes and the small sizes of many, analysis of individual counties and comparison of counties is less reliable.

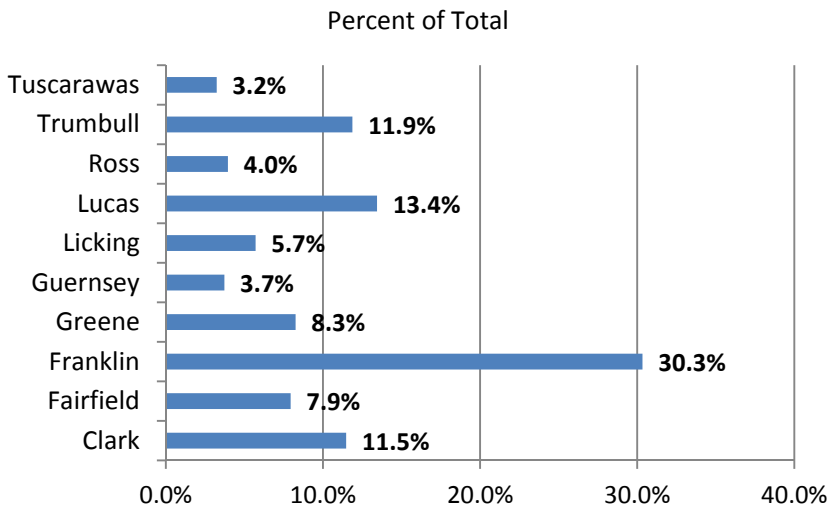


Figure 1. Percent of families from each demonstration county in the full sample (n = 4,529)

Full-sample experimental-control comparison. Random assignment assures that the experimental and control groups have similar characteristics. The goal is to produce two groups that are similar on all important variables, both those that are known and others that are unknown. No statistically significant difference was detected in the slight variations shown in Table 1, which indicates success in the assignment process. In Table 2, demographic and case background variables are compared for experimental and control families. The close similarity and lack of statistically significant differences in each of the categories confirms the essential similarity of experimental and control groups *as collections of families*. The major categories of report allegations were virtually identical. Racial identifications were comparable. Various background variables, some of which are associated with risk of future maltreatment, were also similar. Counts of report dispositions of various types were essentially the same for each group. Finally, various demographic variables were similar across the two groups. None of the differences shown were statistically significant, nor were differences in other variable that were not included. Each of these supports the basic validity of the AR field experiment.

These are important findings that strengthen any findings of *subsequent* outcome differences between the two groups. Outcome differences detected in experiments can be regarded as being caused either by pre-existing differences between groups or by the experimental treatment (and/or measuring and sampling error). By succeeding in random assignment, we have essentially eliminated (or made unlikely) the former of these two sets of causes.

Table 2. Comparative Statistics for the Experimental and Control Groups

Variable	Control	Experi-mental
Allegations of target report		
Neglect	53.7%	53.0%
Physical Abuse	43.9%	44.5%
Emotional Abuse	4.8%	5.3%
Race		
Caucasian	63.5%	62.0%
African-American	24.7%	24.8%
Unknown	11.8%	12.8%
Hispanic Designation in SACWIS	3.1%	2.6%
Background		
Criminal history indicator	7.6%	6.5%
Drug activity indicator	8.8%	8.8%
Mental health hazard (adults)	3.3%	3.6%
Domestic violence	8.2%	7.4%
Substance abuse known in family	4.6%	3.8%
Child in family with behavior problem	8.6%	7.0%
Before Assignment to Study (means)		
Reports screened out or in	1.72	1.82
Reports screened in	1.45	1.53
Investigations unsubstantiated	.83	.90
Investigations substantiated	.32	.31
Investigations indicated	.18	.19
Demographics (means)		
Number of children	2.03	2.01
Infants	.15	.14
Ages 1-3 years	.41	.40
Ages 4-5 years	.25	.24
Ages 6-11 years	.63	.63
Ages 12-17 years	.41	.40
Male children	1.03	1.01
Female children	.99	.98
Number of adults	1.72	1.72

The original 2008-10 evaluation included several methods of data collection, including monthly uploads of SACWIS data, general workers surveys, worker case-specific (case review) surveys, multiple field interviews of workers and supervisors, community surveys and interviews, family interviews, and family questionnaire surveys. Families that completed the family questionnaire make up a subsample of the full sample in this analysis and will be considered further in this report.

The *family-survey subsample* is composed of families that could be contacted and that chose to respond. It consisted of 419 control and 337 experimental families. Comparing the full sample to family-survey sample, no significant differences (with probability less than .05) were found for criminal history, drug activities or mental health problems, domestic violence or children with behavior problems. Allegations of the target report were less often for neglect (full: 53.4%; family-survey: 49.7%) and more often for physical abuse (full: 44.2%; family-survey: 47.6%). Sample cases had higher proportions of past accepted reports (full: 1.49 per family; family-survey: 1.70 per family). No differences were found in the number of adults and children in families or the number of children in various age groups or genders. On this basis it was assumed that the full sample and family-survey subsample were comparable.

3 Follow-up Impact Analysis

The analysis of subsequent reports and child removals in the original report focused on the *first* such report received. The statistical method used was *survival analysis* (Cox proportional hazards analysis). Survival analysis is a method that is used extensively in medical studies and the name conjures up images of severely ill patients some of whom survive after being treated while others die. However, the survival event can be anything that is measurable and occurs on a particular day after a treatment process. A subsequent accepted report of child maltreatment or a new removal and placement of a child after a family assessment or investigation both fit this definition. Survival analysis takes into account the number of individuals who survive with no instances of the problem event and the length of time until the problem event occurs. Thus, it considers not only the occurrence of the event but differences in the *delay* until it occurs. Survival analysis also adjusts for cases that survive through the end of data (*censored* cases), since it is possible that the event (a new report or placement) may occur after the end of data collection. This method is appropriate for studies in which subjects enter at different times with variable periods of follow-up, both of which are characteristics of the AR evaluation design.

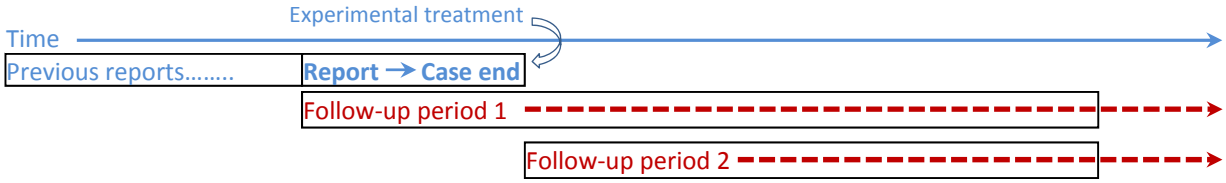
There is a sense, however, in which survival analysis is not appropriate for child welfare studies. New reports of child maltreatment are indeed point-in-time events but they are not one-time events, like dying or getting a disease. They may recur and accumulate for families, that is, they may be counted and summed. The present analysis also considers *quantity of new reports and placements per 100 families*, which represents an extension of the original approach to the evaluation. We will refer to this as a *cumulative analysis*. The criticism of the cumulative analysis is that it does not take into account differences in follow-up time for families, which vary at this point in the follow-up (September 30, 2011) from 24 to 39 months. However, since random assignment occurred on a daily basis throughout the study period, the *average* follow-up periods for experimental and control families were virtually identical.

Thus, the first hypothesis of the analysis—addressed by the survival analysis—is that experimental families will have fewer instances of *any new reports/placements.* The second hypothesis—addressed by the cumulative analysis—is that experimental families will have *fewer cumulative subsequent reports* than control families. The latter would imply that AR was effective in *reducing the total number* of returns to the child welfare or child protection system, a finding that implies improved child and family welfare but also has implications for the cost of operating the child welfare system.

3.1 Subsequent Reports: Survival Analysis

The tracking of subsequent reports began after the target case was closed. The target case was measured from the day of the report that brought the family into the system until the end of the assessment/investigation or, when a service case was opened, the end of the service case. If new screened-in reports were received on a family the date of first new screened-in report was noted and the length of time since the target case closing was recorded. This information was necessary for the survival analysis. To insure that two reports on the same incident were not recorded; reports were counted only when 14 days had elapsed after the immediately previous report.

There are two ways to track the occurrence of new reports: 1) counting from the date of the original report that led the family into the study and 2) counting from the date the target case ended. The two follow-up periods are represented in the following diagram. Both will be considered in this analysis, but with an emphasis on the second, which is the more appropriate experimental measure. Follow-up period 1 counted families with a new report after the date of the target report/ follow-up period 2 counted only new reports after the target case had ended.



First, we will consider simple percentages of new reports among experimental and control families. During follow-up period 1, 41.6% of control families experience at least one screened-in report compared to 39.2% of experimental families. This was statistically significant (p = .05). During follow-up period 2, which did not include new reports during the target case, 38.9% of control families had at least one new screened-in report compared to 37.4% of experimental families. This difference was not statistically significant (p = .17).

One the important findings of a recent long-term follow-up analysis of Minnesota families from the original AR evaluation (2001 to 2005) utilizing up to nine years of follow-up on families was that lasting effects of AR were stronger among families that had *no previous contact with Child Protection Services* when they entered the evaluation study. The same finding is shown for Ohio in Table 3, again as simple percentages. The proportions of families with previous contacts (via screened-in child maltreatment reports) was high in the Ohio sample: 1,075 of 2,244 control families (47.9%) and 1,104 of 2,285 experimental families (48.3%). This value of slightly less than half of families with previous experience in CPS was higher than was observed in Minnesota, where the proportion was slightly more than one-fifth. This may be attributable to differences in initial screening of reports or to the availability of fuller historical data in the Ohio SACWIS.

It appears that AR was effective (in reducing new reports of child maltreatment) *only for families not previously encountered by child welfare*. The difference in new reports was in the range of four percent or slightly less, regardless of the follow-up period. Families with no previous reports are <u>on average</u> lower-risk families on a variety of other variables.[1] On the other hand, families that had previous encounters showed no effects from the intervention on this variable. Previous encounters are one indicator of higher risk and such families may have been in need of more intensive services than is typically provided in the short-term intervention characteristic of most AR family assessments. However, we hasten to point out that

[1] The term "on average" is introduced to emphasize that some first encounter families are nonetheless moderate to high risk on other variables, which in turn contribute to recurring reports of child maltreatment.

subsequent maltreatment reports is only one measure of success, albeit an important measure, and it is possible that such families were influenced and affected in other ways by AR.

Table 3. Percent of families with any new screened-in reports separated into groups with no screened-in reports before the target report versus one or more reports during the two follow-up periods

Follow-up period 1			
Previous reports	Control	Experimental	Probability*
None	33.4%	29.7%	**.029**
One or more	50.5%	49.4%	.311
Follow-up period 2			
Previous reports	Control	Experimental	
None	31.1%	27.1%	**.017**
One or more	47.3%	48.5%	.302

* Significant values shown in bold

Turning to the survival analysis, for simplicity's sake we will show only the analysis for follow-up period 2, after the target case was closed. (The analysis for period 1 was analogous.) Table 4 shows the results of three separate Proportional Hazards Analyses (survival analysis) of experimental and control families. The top portion of the analysis shows that at this point in the follow-up no significant difference appears for experimental and control families regarding the variable *any new screened-in reports of child maltreatment*. The bottom portion shows a separate analysis of families with no previous screened-in reports when they entered the study versus those that had one or more previous screened in reports. The analysis indicates that a significant continuing (through 9/30/11) difference in reduced new reports ($p = .025$) was found among families that had no earlier screened-in reports before entering the AR demonstration. No such difference ($p = .775$) was found for families that had a history with CPS. Because previous reports are associated with other risk factors, this suggests, as noted above, that the most significant effects of AR are found among lower-risk families and may also suggest that families with a previous history in CPS (so called *frequently encountered families*) need more extensive help than is typically available through AR.

The relative hazard statistic indicates that for families with no previous reports about 84 experimental families received at least one new report for every 100 control families. This is a statistical estimate based on the follow-up evidence to date. If correct it suggests that AR has important consequences among such families and also reduces the level of response required by local child welfare offices.

Table 4. Cox Proportional hazards analysis for families with no previous screened-in reports versus families with one or more previous screened-in reports

	B	SE	Wald	p	Relative Hazard	95.0% CI for R.H.	
						Lower	Upper
Entire Study Group							
Experimental vs. Control	-.075	.048	2.415	.120	.928	.844	1.020
No previous screened-in reports							
Experimental vs. Control	-.172	.077	5.045	**.025**	.842	.724	.978
One or more previous screened-in reports							
Experimental vs. Control	-.018	.062	.082	.775	.982	.870	1.109

The survival function graph in Figure 2 illustrates the difference for families with no previous history. The lower line for the control group in the graph reflects the lower "survival" of those families, that is, that the proportion of families with no new report decreased more rapidly among control families (thus the more sharply descending curve) and such reports occurred sooner for control families compared to their experimental counterparts.

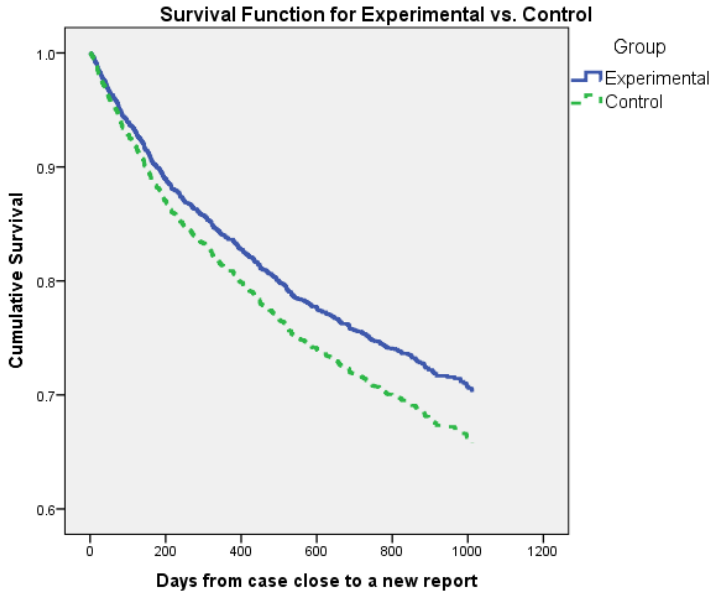


Figure 2. Survival function for experimental and control families with no previous screened in reports

Another difference was discovered that is not shown here. It was evident that African American families fared better on this outcome variable than did Caucasian families. This topic will be pursued in the next section where we attempt to discover why this might be true.

3.2 Subsequent Reports: Cumulative Analysis

Moving to the second measure of *cumulative subsequent reports*, Table 5 shows the counts of reports (and dispositions) per 100 families. This analysis is also limited to follow-up period 2. As can be seen, control families received 68 new screened-in reports per 100 control families compared to 60 per 100 experimental families. This difference was statistically significant (p = .011). Moving down in the table, it can be seen that experimental families experienced significantly fewer unsubstantiated and substantiated reports as well. (Note: for approximately 14 percent of screened-in reports the disposition was that the family could not be located or had moved from the county, or no disposition was indicated.)

Table 5. Child maltreatment reports per 100 families after target case closure

Number of reports per 100 families	Control	Experimental	Probability*
Screened-In	68	60	**0.011**
Unsubstantiated	30	21	**0.000**
Substantiated	10	8	**0.002**
Indicated	5	4	0.088
AR family assessment	16	18	0.248

Probability < .05 shown in bold

As noted previously, this finding has cost implications for the state. Assuming the control group represents CPS before the AR reform, it implies nearly a 12 percent reduction in new reports among families determined to be appropriate for AR. Each family that does not come back with a new report presumably represents a situation that is safer and more secure for the children. And since it is obvious that some families return more than one time, it also represents a potential savings in worker time and administrative expenses for local agencies.

Offering families AR a second time. An interesting finding was that no (significant) difference appeared between the experimental and control group regarding the number of *subsequent AR family assessments* (last row in Table 5). Experimental families, who had received a family assessment during the original target case, received 16 new family assessments per 100 families. Control families who had originally received an investigation received 18 per 100. This difference, however, was not consistent across the ten study counties, as is evident in Figure 3. Franklin County, which as noted accounted for three in every ten study families, appeared to

steer experimental families three times as often to new family assessments than investigations (23.1/100 experimental versus 8.5/100 for control). Why these differences occurred cannot be known without a further investigation. In some case they may be random variations, but the extreme differences in some counties for two groups of families that were highly similar in other ways may reflect differences in pathway assignment practice among counties. Pathway assignment policies will be investigated in subsequent data collection.

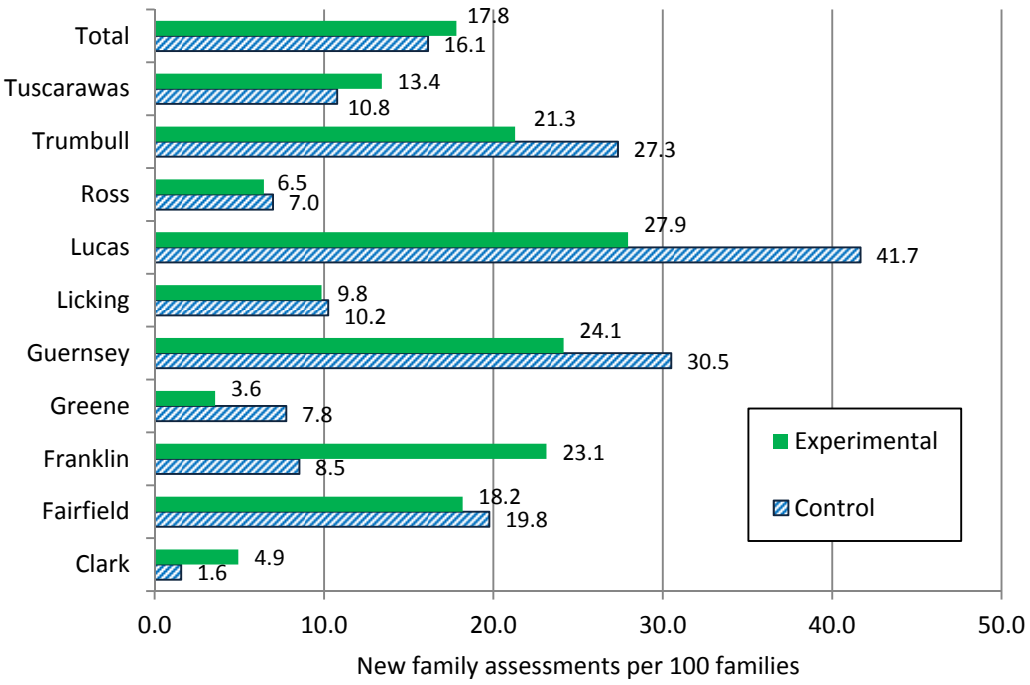


Figure 3. Number of new family assessments per 100 experimental and control families for each study county

Considering multiple factors simultaneously. The proper analysis for this variable is a factorial analysis of variance. In the following we consider counties, groups (experimental or control) and the interaction of counties with groups. The dependent variable is counts of screened-in reports after the target case was closed, that is, during follow-up period 2. The results of the analysis are shown in Table 6. As can be seen, in this combined analysis the experimental and control groups were significantly different ($p = .012$) but also the overall difference among counties (of combined experimental or control outcomes) was statistically significant ($p < .001$). The variation among counties and between groups can best be illustrated by showing the estimated marginal means for the interaction effect (County * Group in Table 6). This is illustrated Figure 4.

Table 6. Factorial ANOVA of number of new accepted reports after case is closed by count and group (Full evaluation sample)

Source	Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	89.976	19	4.736	4.254	**.000**
Intercept	1114.696	1	1114.696	1001.290	**.000**
County	71.858	9	7.984	7.172	**.000**
Group (Exp. vs. Control)	6.983	1	6.983	6.273	**.012**
County * Group	9.852	9	1.095	.983	.451
Error	5019.690	4509	1.113		
Total	6991.000	4529			
Corrected Total	5109.666	4528			

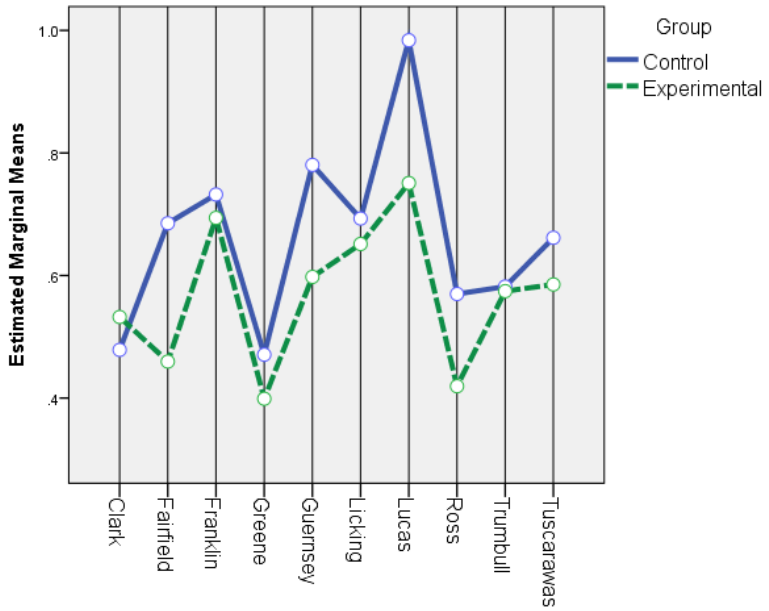


Figure 4. Estimated marginal means of new reports (per family) for experimental and control groups by county (full sample)

The means shown in Figure 4 illustrate the variation in levels of new reports among counties. The two lines are relatively parallel with some exceptions. The lower levels of experimental group means in most counties illustrate the overall statewide effects of AR in reducing the levels of new reports of child maltreatment. The variation among counties, without regard to experimental-control differences, shows how families, whether experimental or control, differ across counties in the level of return to the system. This may, of course, be due to county practice but more likely arises from initial study group differences among counties due to variations in pathway assignment. In other words, the April 2010 evaluation report showed that there was a great deal of variation among counties in the types of families

assigned to the AR pathway. This means that some counties assigned higher risk families that had a history in the system while others were more reluctant to assign such families to the AR pathway. The combined experimental and control groups in each county reflect this pathways assignment variation. Counties assigning more higher-risk families would have more re-reporting of both experimental and control families than counties who assigned fewer higher-risk families. Nonetheless, the chart shows improvements among experimental families in most counties whether the experimental and control means were low or high. This is a desirable outcome and illustrates a finding consistent with earlier evaluations—that AR produces effects in offices that select only the lowest risk families as well as in offices in which moderate and high risk families are assigned to the AR pathway.

3.3 Effectiveness of AR among Minority Families in Ohio

No strong evidence of a disproportional negative response to minority families was found in IAR's previous evaluations of differential response in Missouri and Minnesota. This term has generally referred to more substantiated investigations and more child removals for African American, American Indian and Latino families attributable to DR (AR). Under AR, there has been no evidence to date that minority families are assigned to the AR pathway less frequently, or when they are assigned, that they receive services less often. This was true of both the Missouri and Minnesota evaluation studies. However, a difference in outcomes among minority families appears to have occurred in Ohio. As the April 2010 evaluation report pointed out, AR appeared to be more effective among African American families. Building on the quantitative analysis illustrated in Table 5, families were split into Caucasian and African American, the two dominant racial categories in the Ohio study (Table 7). While an experimental control difference for screened in report is still seen among Caucasian families (control: 72/100 versus experimental: 68/100), the difference was not statistically significant. A substantial difference is shown between experimental African American families compared to control (control: 75/100 versus experimental: 61/100). This difference was statistically significant. For new substantiated and unsubstantiated investigations the difference was significant for both Caucasian and African American.

The large majority of African American families were from Franklin (65.4%) and Lucas (18.0%). Thus, the difference may reflect the approach to AR practice in Franklin or in both these counties. On the other hand, we thought that it was possible that the differences were due to the situation or background of minority families. The following analyses are an attempt to confirm or disconfirm this supposition.

Table 7. Child maltreatment reports per 100 families after target case closure by race

Number of reports per 100 families		Control	Experimen-tal	Probability*
Screened-In	Caucasian	72	68	.303
	African-American	75	61	**.031**
Unsubstantiated	Caucasian	31	25	**.008**
	African-American	35	19	**.0001**
Substantiated	Caucasian	11	08	**.054**
	African-American	12	8	**.018**
Indicated	Caucasian	6	5	.465
	African-American	5	3	.166
AR family assessment	Caucasian	18	19	.615
	African-American	16	20	.156

Probability <= .05 shown in bold

Are there differences in services to African American Families and are such families more financially disadvantaged? It was clearly established in the Minnesota study[2] and in recent follow-up analysis of Minnesota data[3] that the AR approach to families leads to increased provision of material (anti-poverty) services to the most impoverished portion of the CPS population. The question is whether this might be a partial explanation of the apparent effectiveness of AR with African American families. To demonstrate the change in types of services delivered under AR, the graphics that were based on responses of families in Minnesota and Ohio are shown in Figure 5 and 6.

[2] Siegel, G.L. and L.A. Loman. (2006) Extended follow-up study of Minnesota's Family Assessment Response. St. Louis: Institute of Applied Research. Retrievable at http://www.iarstl.org/papers/FinalMNFARReport.pdf.
[3] Loman, L.A. and G. L. Siegel. (2012), *Anti-poverty services under the Differential Response Approach to Children Welfare.* Submitted for journal publication.

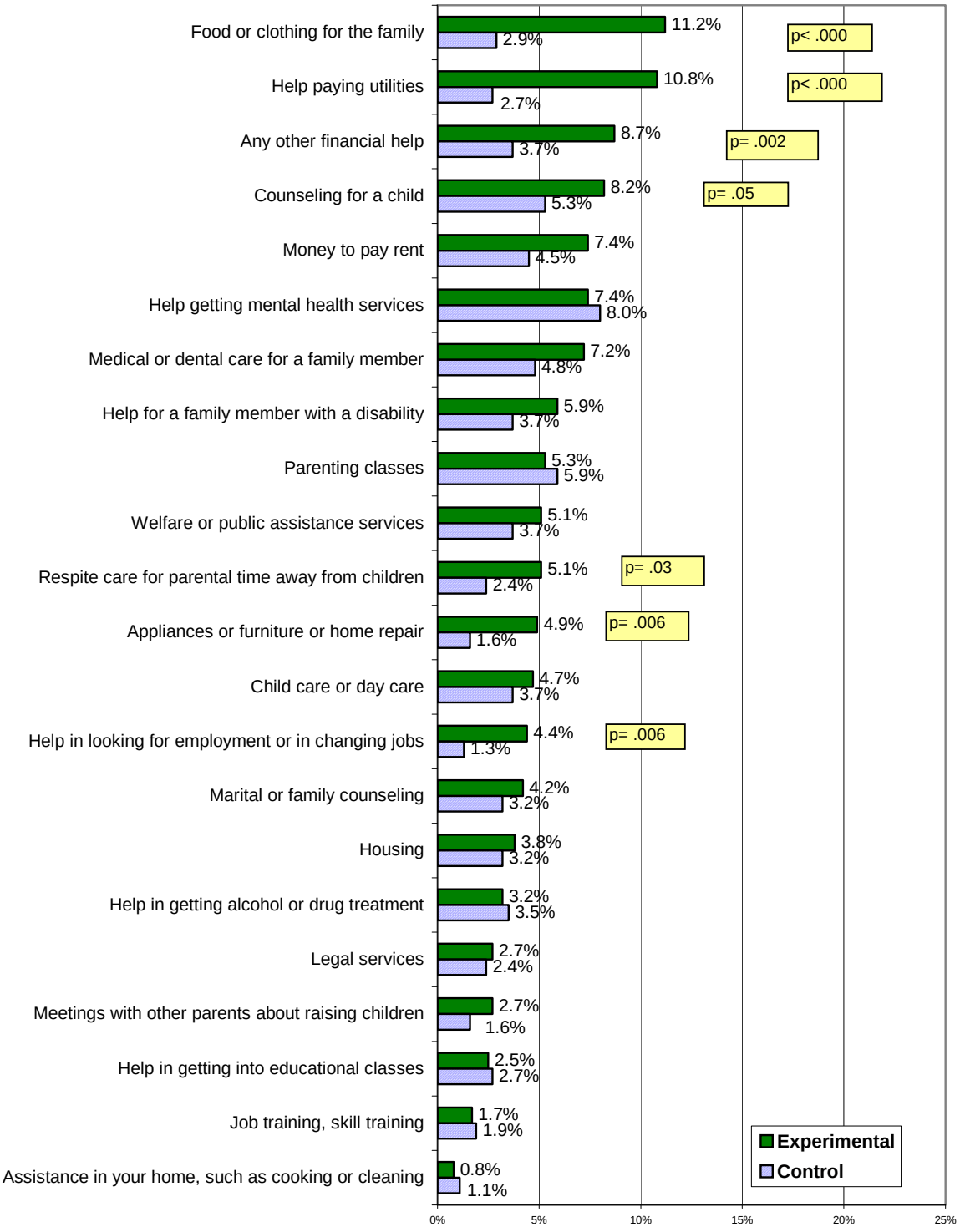


Figure 5. **Minnesota:** Percent of experimental and control families who reported receiving specific services.

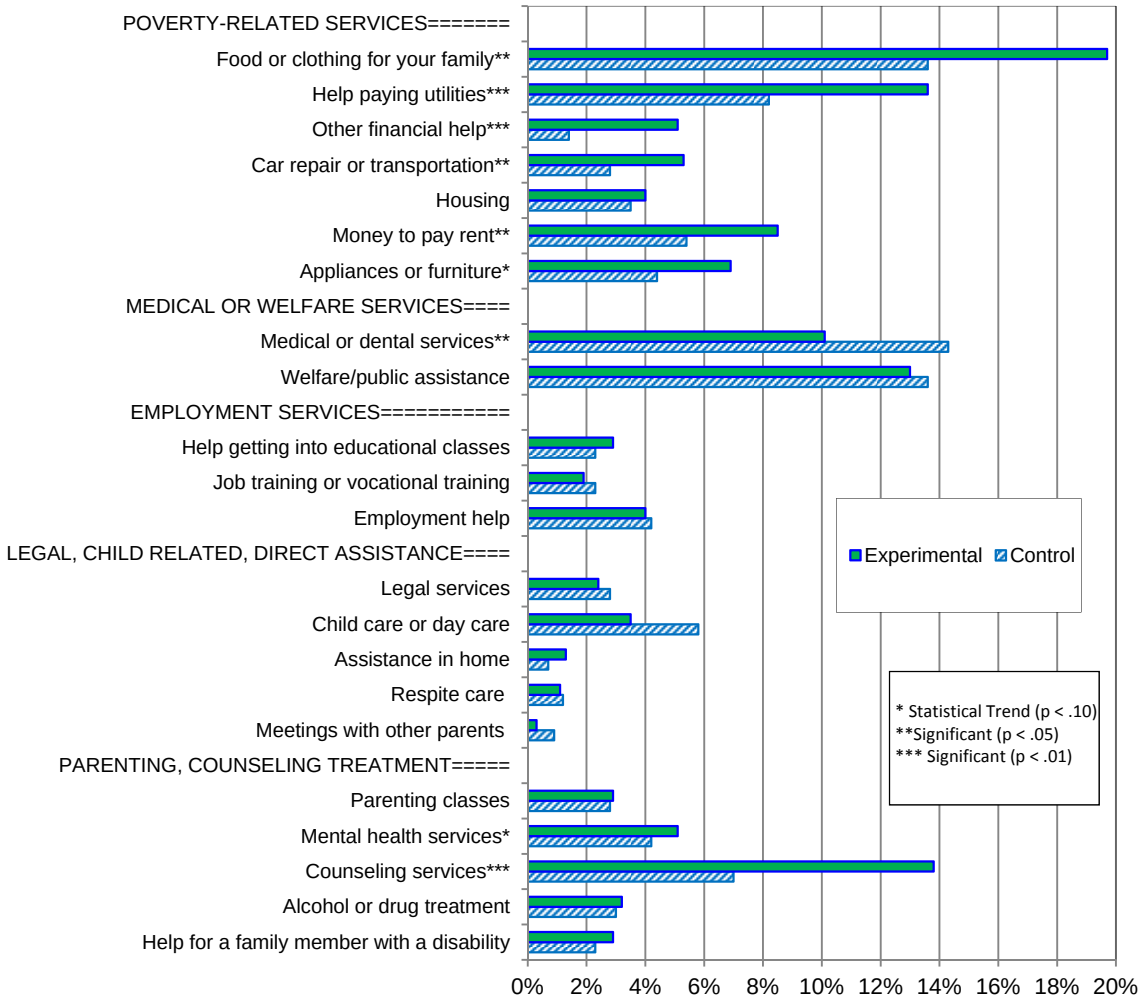


Figure 6. **Ohio:** Percent of experimental and control families who reported receiving specific services.

The two figures were arranged and presented slightly differently in the two final reports, but they show essentially the same immediate outcomes. In each study, experimental families, all of whom received family assessments, reported significant and sometimes substantial increases in material services (food, clothing, rent, utilities, transportation, etc.) when compared to control families, all of whom received traditional investigations.

Turning to Ohio alone and <u>focusing on the family-survey sample only</u>, a measure of material services was created that combined food, clothing, rent assistance, appliances, furniture, home repair, help with utilities, and other financial help. Results showed the 35.9% of experimental families received at least one of these services compared to 21.0% of control families ($p < .001$).

We also asked whether African American families might be the more financially disadvantaged portion of the study sample. An index of socio-economic status (SES) was computed which divided the family-survey sample into two groups. Low SES was defined as either an income less than $15,000 during previous year, lack of a high school diploma or reception of assistance (food stamps, TANF, housing assistance **or** utilities assistance) during the previous year. Within the family-survey sample, 86.1% of families were in the low-SES segment. The family survey sample likely to have underrepresented financially disadvantage families.[4] It is likely, therefore, that the levels of poverty and financial distress were similar in the full sample. SES was balanced between the experimental and control groups with 85.4% of control families in the low-SES group compared to 86.9% of experimental families (p = .314). A similar analysis was conducted in Minnesota where roughly half of the sample fell into this group, highlighting again the differences between the Minnesota and Ohio families who entered the evaluation.

African American families were more disadvantaged, but only slightly more. Among the entire family survey sample without regard to experimental or control group membership, 86.1% were low SES as defined here. Among African American families, 92.6% were low SES compared to 86.7% of Caucasian families.

The analysis did not show disproportion—either negative or positive—regarding material services and race. Among African American families 28.9% reported receiving at least one of the material services listed as compared to 28.7% of Caucasian families. Nor did the slightly higher proportions of low SES families in the African American portion of the sample make a difference. While 100% of the low SES African American families received at least one material service, 94.8 percent of low-SES Caucasian families did as well.

These balances and the small proportion of families in the higher SES group restricted the possible analyses. However, it was possible to look at outcomes by race and by provision of services, remembering that many more material services were provided to low-SES families. This is shown in Figure 7, which represents the count per 100 families of new screened-in reports during the follow-up period. This figure is not definitive but only indicative, since the number of families in each subgroup, particularly African Americans, was small. Among the comparison bars only the difference on the far right was statistically significant (p = .05). The difference between these two bars indicates that **it was not** the differential effect of services

[4] Low SES families are slightly **less** likely to respond to written surveys and if that is the case then the family survey sample would underrepresent the financial disadvantaged population in the full sample.

that accounts for better outcomes for African American families. If it were, one would expect the height of the bars to be more or less equal.

Another possible explanation was the greater effectiveness of AR for lower-risk families, that is, families with no previous contacts with CPS. We asked whether African American families might have had fewer previous reports in comparison to Caucasian families. Among African American families, 45.3% had had no previous screened in reports compared to 47.2% of Caucasian families, a difference that was not statistically significant (p = .146). Thus, this explanation was rejected, as well.

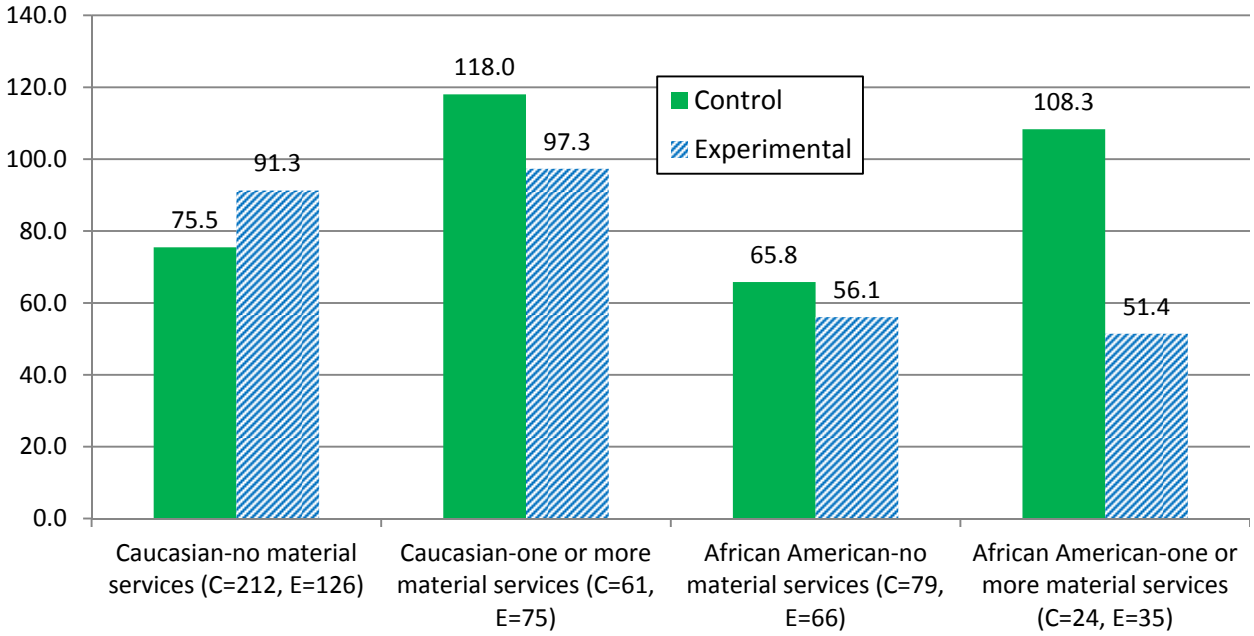


Figure 7. Number of new screened in reports per 100 families by race and reception of material services (family-survey sample, numbers of experimental and control families in the column titles)

To summarize, we asked whether the increase in material anti-poverty services that characterizes AR, difference in SES, or differences in previous reports before entering the study might explain the differences in outcomes between African American and Caucasian families. None of these alone showed a difference. However the large differences between African American families that did receive material services (far right set of bars in Figure 7), which is also reflected, at a lower level, among Caucasian families (second set of bars from left in Figure 7), may point to an explanation. The fact that experimental families appear to do better when such services are made available when compared to control families suggests that it is not simply material services that were effective but services combined with the family friendly,

supportive approach of AR. We are left with the original notion that the variation probably will be found in the AR practice toward minority families, and perhaps the approach in counties where most African American families resided. This will be investigated further in the third year of study.

3.5 Variations in Types of Later Reports

It may be informative to study the nature of the allegations of subsequent reports of child maltreatment among experimental and control families. We have shown that subsequent reports are reduced among experimental families when compared to their control counterparts. The question for this section is whether the reductions are uniform across the different types of allegations of child maltreatment or whether particular types of allegations are reported less often. Each subsequent report concerns allegations for at least one child but more often several children in the family. Allegations of a single kind of child maltreatment or several kinds may be made for each child in the family. In most cases the allegations, whether single or multiple, are the same for each of the children associated with the report. In a minority of cases they are different.

The approach in this analysis was to count allegations in five general categories: child neglect, physical abuse, sexual abuse, emotional maltreatment and medical neglect. If an allegation was contained in a report it was counted only one time per report regardless of whether it was alleged for one or several children in the family. Thus, a report of neglect for three children was counted as one report of neglect. In addition, a small number of families had two or more reports with the same allegation(s). Allegations were counted only once per family for this analysis.

In a previous analysis in another Differential Response evaluation, we found that child neglect reports were subsequently reduced among families that previously had received family assessments.[5] In Ohio, differences can be seen in each of the allegation categories (Table 8), except child neglect. Cumulatively, these add up to the differences shown earlier that led to a significant reduction of subsequent reports. However, among the single allegations, only the quantity and percent difference of physical abuse reports were large enough to register as statistically significant. The implication is that the AR family assessment intervention and

[5] Siegel, G.L., C.S. Filonow and L.A. Loman. (2010) Differential Response in Nevada: Final Evaluation Report. St. Louis: Institute of Applied Research. Retrievable at http://www.iarstl.org/papers/NevadaDRFinalReport.pdf.

change in services led some family caregivers later to avoid being reported for (and by implication avoid committing) physical abuse of their children. Reductions can also be observed in later sexual abuse or emotional maltreatment but the numbers and differences of reports of these kinds were individually too small to show up as significant.

Table 8. Percent of families by types of allegations in subsequent reports of child maltreatment after the target case (2,244 Control families; 2285 Experimental families)

Type of allegation	Control % (n)	Experimental % (n)	p
Child neglect	23.3% (523)	23.5% (537)	.452
Physical abuse	23.8% (534)	20.2% (461)	**.002**
Sexual abuse	8.4% (49)	7.6% (40)	.172
Emotional maltreatment	2.2% (49)	1.8% (40)	.173
Medical neglect	1.9% (42)	1.5% (34)	.187

* Significant values shown in bold, Chi Square – Exact Tests

Some find it surprising that sexual abuse allegations later arise in families that were determined to be appropriate for AR, as indeed all experimental and control families in this study were. However, those familiar with patterns of report allegations know that allegations for families that are reported more than one time *are more likely to be different than to be the same*. This is part of the *revolving iceberg metaphor* of child maltreatment. Families with multiple reports tend to have a variety of underlying problems and needs (the iceberg) that manifest themselves in different ways (revolving) over periods of months and years.[6]

3.6 Impact of AR on Removal and Placement of Children

Among the April 2010 evaluation findings, we reported that AR appeared to have led to a reduction in later removals and placements of children. The numbers of children removed from families was very small, and as noted above, the time for following families was short. While the smaller proportion of removals among experimental families has continued, at this point it is not large enough to register as statistically significant for the full sample of families. However, an interesting and surprising difference appeared as we examined the study groups in greater detail.

[6] For a detailed analysis of this phenomenon utilizing data from another state see: Loman, L.A., Family Frequently Encountered by Child Protection Services: A Report on Chronic Child Abuse and Neglect. St. Louis: Institute of Applied Research. Retrievable at: http://www.iarstl.org/papers/FEfamiliesChronicCAN.pdf

About one in every ten families that entered the study had at least one child removed from their family and put in out-of-home placement before assignment to the experimental or control group. In most cases children had been reunited with their families before assignment. As with other proportions shown earlier the percentages of children removed were virtually identical for the two groups (control: 10.3%; experimental: 10.1%, p = .438), again confirming the success of random assignment.

Looking at placements after the target case, the numbers remain low with 151 of 2,244 control families (6.7%) with a child removal compared to 133 of 2,285 experimental families (5.8%). This difference of about a percentage point was not large enough to specify it as statistically significant (p = .115). Thus, for the overall AR sample we cannot now say with confidence that child removals were reduced. However, another surprising difference did appear.

This difference is shown in Table 9. When the study groups are split into families with no previous child removals versus families with one or more child removals a different picture emerges. No difference is apparent between the large sets of families (2,013 and 2,054) in the first of these conditions. It is apparent that all the gains in avoiding child removals came within families that had a previous removal before being assigned to the study. Of the 231 control families in this condition 35 (15.2%) had a child removed later compared to 18 of 231 (7.8%) of experimental families. Even though the groups were relatively small, this difference was large enough to be statistically significant (p = .009).

Table 9. Percent of families with and without previous child removals that had a subsequent removal after the target case (2,244 Control families; 2285 Experimental families)

Type of allegation	Control % (n)	Experimental % (n)	p
No previous child removals			
Families with child removed	5.8% (116)	5.6% (115)	.437
Total	2013	2054	
A previous child removal			
Families with child removed	15.2% (35)	7.8% (18)	**.009**
Total	231	231	

* Significant values shown in bold, Chi Square – Exact Tests

Survival analysis reflects these findings as is shown in Table 10. The analysis suggests that virtually all the difference between the total population of experimental and control families is attributable to changes that occurred among families in which children were previously removed and placed. The implication is that AR was particularly effective in changing behavior among such families. The analysis suggests that about half as many families exposed to the family assessment process (relative hazard = .5) can be expected to have a child removed as their control family counterparts. This fits closely the actual percentages shown in Table 9 (7.8% versus 15.2%).

Table 10. Cox Proportional hazards analysis for families with no previous child removals versus families with one or more previous child removals

	B	SE	Wald	p	Relative Hazard	95.0% CI for R.H.	
						Lower	Upper
Entire Study Group							
Experimental vs. Control	-.144	.119	1.458	.227	.866	.686	1.094
No previous removals							
Experimental vs. Control	-.024	.132	.034	.854	.976	.754	1.263
One or more previous removals							
Experimental vs. Control	-.694	.290	5.727	**.017**	.500	.283	.882

We should point out that this finding stands in contrast to the earlier finding regarding subsequent reports of child maltreatment. We found that lower-risk families appear to be helped more for that variable. In the present analysis, the highest-risk families, those that were three times as likely to have a child removed, were assisted most. This finding will be tracked and further analyzed to determine whether it continues as more children are placed and, if so, to discover why it happens. One possible explanation that will be pursued is the overlap in Ohio between the AR and Title IV-E waiver demonstration. The latter is designed to find alternatives to foster placement that may dovetail with AR services and may have had positive consequences for AR families as well.

4 Conclusions

The finding of the analysis of racial differences should be emphasized because it suggests that the effectiveness of AR lies not simply in the increase in services—particularly material or anti-poverty services—but in the combination of broadened services with the family-friendly and participatory approach embodied in AR family assessments. This suggests

that the context of family-worker engagement in which services and other assistance are provided is critically important.

Another important finding concerns *cumulative reports*. The statistically significant and rather substantial difference in new reports per family suggests improvements in child and family welfare as a consequence of AR. It also has important cost implications for the system and will be examined from this standpoint in subsequent analysis.

Although this analysis of *any new reports* showed that AR was most effective among families that had not been encountered previously by Child Protection, the finding does not mean that investigations are better for previously encountered families. Nearly equivalent outcomes were achieved for the latter, with no new safety problems during the target case, as the original evaluation demonstrated. No support is found in these results for treating previously encountered families in a more forensic manner. Families with one or more past screened-in reports are higher-risk <u>on average</u> than families with one report, which means they have greater needs. The difference found may be an argument for more intensive services for such families. Indeed, as the analysis also showed, AR may reduce the need to remove children from families who, on the basis of past removals, would be classified as higher risk.