



Northern California Training Academy Research Evidence Use Practicum

SESSION 2: Developing a hypothesis

February 27, 2017

Jennifer Haight and Lily Alpert

The Center for State Child Welfare Data
Chapin Hall at the University of Chicago

1

Today's agenda:

9:00 – 9:30	Welcome back
9:30 – 10:30	A closer look at few core concepts
10:30 – 10:45	Stretch break
10:45 – 12:00	Recap and review group your work to date: "I observe that...I think it's because..."
12:00 – 12:45	Lunch break
12:45 – 2:00	"So I plan to...which I think will result in..." – identifying an intervention that is rooted in theory, if not evidence.
2:00 – 2:15	Stretch break
2:15 – 3:15	Group work: crafting a plan
3:15 – 3:45	Assign homework & video call with Renee Boothworthy
3:45 – 4:00	Conclude

2

Last time...

I observe [**some outcome that I want to improve**]. ✓

I think it's because of [**this reason**].

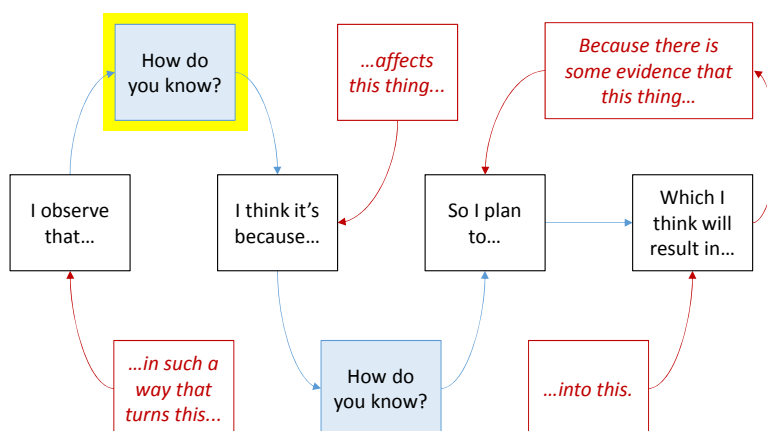
So I plan to [**implement this intervention**],

which I think will result in [**an improved outcome**].

- You worked with FCDA administrative data analyses to make an **observation** about **reentry** in your county and identify a problem that needs solving.

3

Theory of change



© The Center for State Child Welfare Data

4

Some important concepts

5

1.

**Case review:
Who should you **compare** to whom?**

6

Example...

I observe that children placed as **infants** are **more likely** to reenter care **than** children placed at **older ages**.

I think it's because my county does not have access to very good **age-appropriate parenting programs** for parents of infants.

Some of you said... I bet if I **compared infants who reentered to infants who didn't**, I would find that the parents of those who reentered did not receive age-appropriate parenting support while their children were in care.

Among...	I expect to see that...
Infants who reentered within 12 months	...parents did not have age appropriate parenting support
Infants who did not reenter within 12 months	...parents did have age appropriate parenting support

But I just said that I actually think that this is pretty rare...

In order build an argument that a lack of age-appropriate parenting support is the thing that makes **infants more likely than older children** to reenter (which is your observation), you'll need to compare children based on age.

Among children placed as...	I expect to see that...	
	Reentered within 12 months	Did not reenter within 12 months
Infants	...parents did not have age appropriate parenting support	...parents did have age appropriate parenting support.
1 and older	...parents did have age appropriate parenting support.	...parents did have age appropriate parenting support.

Therefore... the **question** for the case review is:

Among the cases reviewed, how many had parents who received age-appropriate parenting support while the child was in care?

2.

Stratified random sampling

9

Among the cases reviewed, how many had parents who received age-appropriate parenting support while the child was in care?

How do I select a case review sample to answer this question?

Let's say 1,000 children reunified in 2014 and 15% of them reentered within 12 months:

	Rntr < 12m	No rntr < 12 m	Total
Number			
Under 1	80	220	300
1 +	70	630	700
Total	150	850	1,000
Percent of grand total			
Under 1	8%	22%	30%
1 +	7%	63%	70%
Total	15%	85%	100%

If we draw a **random sample** of 40 cases, we're probably going to wind up with this:

	Rntr < 12m	No rntr < 12 m	Total
Under 1	3	9	12
1 +	3	25	28
Total	6	34	40

Representative of the population...
but unequal and tiny cell sizes

To ensure equal cell sizes, we could draw a **stratified random sample**.

- Create the strata first and then draw a random sample from each.

Let's say 1,000 children reunified in 2014 and 15% of them reentered within 12 months:

	Rntr < 12m	No rntr < 12 m	Total
Number			
Under 1	80	220	300
1 +	70	630	700
Total	150	850	1,000
Percent of <u>grand total</u>			
Under 1	8%	22%	30%
1 +	7%	63%	70%
Total	15%	85%	100%

If we draw a **stratified random** sample of 40 cases:

	Rntr < 12m	No rntr < 12 m	Total
Under 1	10	10	20
1 +	10	10	20
Total	20	20	40

Representative **within each stratum**
Equal cell sizes that are as large as manageable

A word on sample size...

We conduct case review in order to acquire rich, qualitative data on the process and quality of care.

We select a sample because a review of all cases in the population would be prohibitive.

Random sampling gives us the best chance of reviewing a **representative** sample.

That said, we should still extrapolate from case review with a **caveat** about **generalizability**.

	Rntr < 12m	No rntr < 12 m	Total
Under 1	10	10	20
1 +	10	10	20
Total	20	20	40

3.

The difference between a fraction that describes the **characteristics of population** and one that describes the **likelihood of an outcome**

13

What do we mean when we talk about **likelihood**?

First off, try to reserve the word **likelihood** for referring to the **likelihood of an outcome**.

- The likelihood of entering foster care.
- The likelihood of exiting to permanency.
- The likelihood of reunifying within six months.
- The likelihood of experiencing 2 or more placement moves.
- The likelihood of spending time in a congregate care setting.
- The likelihood of aging out of foster care.
- The likelihood of re-entering care within 12 months of reunification.

What do we mean when we talk about **likelihood**?

Second, recognize the synonyms.

When we ask “What is the **likelihood** that...?”

We’re asking:

- What are the **odds**?
- What are the **chances**?
- What is the **probability**?

What do we mean when we talk about **likelihood**?

Probability is usually expressed as a rate (i.e., a percentage). $\frac{\text{Numerator}}{\text{Denominator}}$

1,000 reunified in Acme County in 2014 and 150 of those children reentered care within 12 months.

- **Rate of reentry within 12 months = 150 / 1000 = 15%**

Make a statement about what **actually** happened:



“Of all children who reunified in Acme County in 2014, 15% reentered within 12 months.”

Use that fact to predict what is **likely** to happen in the future:



“Children reunifying in Acme County have about a 15% *likelihood* of reentering within 12 months.”

How am I going to know whether or not I'm right about that? I am going to look at past cohorts and see what actually happened.

I'm going to calculate a fraction that tells me the **actual rates** of reunification and aging out for those past cohorts.

I think teens are more **likely** to age out than they are to reunify.

Huh. Looks like teens are actually more **likely** to reunify than they are to age out.

$$\frac{\# \text{ aged out}}{\text{all admissions}} = \frac{32}{100} = 32\%$$

$$\frac{\# \text{ reunified}}{\text{all admissions}} = \frac{56}{100} = 56\%$$

Beep boop beep...

Variables that predict the likelihood of an outcome

Different factors may increase or decrease **the likelihood of reentry**.

These factors are known as predictors or independent variables.

- e.g., child's age at placement

The outcome is the dependent variable (it "depends" on the independent variable)

Independent variable	Reentered within 12 months? (Dependent variable)					
	Number			Percent		
	Reentered < 12 m	Did not reenter < 12 m	Total	Reentered < 12 m	Did not reenter < 12 m	Total
Total reunified	150	850	1,000	15%	85%	100%
<i>Age at placement</i>						
Under 1 y.o.	80	220	300	27%	73%	100%
1 y.o. and older	70	630	700	10%	90%	100%

Measures of likelihood...

Categorize children by the PREDICTOR...

...to explain variation in the OUTCOME
"How many [experienced this]?"

The risk set
("of all children who reunified...")

Independent variable	Reentered within 12 months? (Dependent variable)					
	Number			Percent		
	Reentered < 12 m	Did not reenter < 12 m	Total	Reentered < 12 m	Did not reenter < 12 m	Total
Total reunified	150	850	1,000	15%	85%	100%
Age at placement						
Under 1 y.o.	80	220	300	27%	73%	100%
1 y.o. and older	70	630	700	10%	90%	100%

Row percent

"Among the children who reunified, children placed as infants **were more likely** (27%) **to reenter** within 12 months than children placed at age 1 or older (10%)."

But sometimes, fractions...

...And describe the CHARACTERISTICS of each group.

Categorize children by the OUTCOME...

	Independent variable	Reentered within 12 months? (Dependent variable)		
		Reentered < 12 m	Did not reenter < 12 m	Total
Number	Total reunified	150	850	1,000
	Age at placement			
	Under 1 y.o.	80	220	300
	1 y.o. and older	70	630	700
Percent	Total reunified	100%	100%	100%
	Age at placement			
	Under 1 y.o.	53%	26%	30%
	1 y.o. and older	47%	74%	70%

Not the risk set!
("of all children who reentered within 12 months...")

"Selecting on the dependent variable"

Column percent

"Of children who reentered within 12 months, 53% were placed as babies and 47% were placed at age 1 or older."

20

"Of children who reentered within 12 months, 53% were placed as babies and 47% were placed at age 1 or older."

	Independent variable	Reentered within 12 months? (Dependent variable)		
		Reentered < 12 m	Did not reenter < 12 m	Total
Number	Total reunified	150	850	1,000
	Age at placement			
	Under 1 y.o.	80	220	300
	1 y.o. and older	70	630	700
Percent	Total reunified	100%	100%	100%
	Age at placement			
	Under 1 y.o.	53%	26%	30%
	1 y.o. and older	47%	74%	70%

*This is aligned with my hypothesis that being placed as an infant **makes a child more likely to reenter.***

Indeed, the previous slide shows us that is true: "Among the children who reunified, children placed as infants were more likely (27%) to reenter within 12 months than children placed at age 1 or older (10%)."

21

WHAT IS THE POINT FOR CRYING OUT LOUD.

It's important to be precise!

Of children who reunified who were placed as infants, what proportion reentered within 12 months of exit?

$$80/300 = 27\%$$

Of children who reentered within 12 months of reunification, what proportion were placed as babies?

$$80/150 = 53\%$$

Which of these percentages are we trying to change?

*Creating your fraction too hastily could lead you to make an incorrect statement about the **likelihood** of reentering for children placed as infants...*

WHAT IS THE POINT FOR CRYING OUT LOUD.

We are trying to predict the **likelihood of an outcome** – namely, reentry – so that we can learn something that helps us know how to prevent it.

Not all fractions describe the likelihood of an *outcome*... check your denominator!

If your denominator reflects the risk set, your fraction describes a likelihood of that outcome. If it doesn't, that fraction means something else.

Fine, but what does this have to do with case review?

When generating evidence from administrative data, statements about likelihood require you to appreciate the risk set. ***The same is true for statements resulting from case record review.***

What questions are we asking in our case review and what do we learn from them?

Are we learning whether certain factors affect the likelihood of an outcome?

- (Hint: The sample is selected based on the risk set – children who had the potential to reenter)

Or are we learning about the characteristics of children who fall into one outcome category or another?

- (Hint: The sample is selected based on the outcome – children who did/didn't reenter)

Case review set you up to answer THIS: *Of children who did/didn't reenter, what percent had parents who received age-appropriate parenting support while the child was in care?*

		Number		
		Rntr <12	No Rntr <12	Total
Number	Infants	10	10	20
	AAPS	2	4	6
	No AAPS	8	6	14
	1 and older	10	10	20
	AAPS	7	9	16
	No AAPS	3	1	4
Percent	Infants	100%	100%	100%
	AAPS	20%	40%	30%
	No AAPS	80%	60%	70%
	1 and older	100%	100%	100%
	AAPS	70%	90%	80%
	No AAPS	30%	10%	20%

Sample is based on the outcome category:

Descriptive statistics (column percent)

- In general, infants received the service less frequently (30%) than older children (80%).
- Among all children, stable cases got the service more often (40%, 90%) than re-enterers (20%, 70%).
- Among children who reentered within 12 months, not getting the service was more common for infant cases (80%) than it was for older re-enterers (30%).

What do we know at this point?

- We already observed that infants are more likely to reenter than older children.
- Case review indicates that parents of infants don't receive age-appropriate parenting services as often as parents of older children do.
- Case review also indicates parents of re-enterers don't receive these services as often as parents of stable children.

All of this is supportive to our hypothesis that the lack of these services is behind reentry for infants.

BUT – to test whether receiving services, in fact, affects the likelihood of reentry, we have to ask:

- *Of children whose parents received the service, what percent reentered care within 12 months?*
- *Of children whose parents didn't receive the service, what percent reentered care within 12 months?*

We can get a sense of that (at least among the reviewed cases) by changing the denominator: *Of children whose parents received the service, what percent reentered within 12 months?*

Sample is based on the risk set (row percent)

	Number			Percent		
	Rntr <12	No Rntr <12	Total	Rntr <12	No Rntr <12	Total
Infants	10	10	20	50%	50%	100%
AAPS	2	4	6	33%	67%	100%
No AAPS	8	6	14	57%	43%	100%
1 and older	10	10	20	50%	50%	100%
AAPS	7	9	16	44%	56%	100%
No AAPS	3	1	4	75%	25%	100%

- Infants whose parents got the service were less likely to reenter within 12 months (33%) than those whose parents didn't (57%).
- The same was true for older children (44%, 75%). In fact, it seems like not receiving the service might have had an even stronger effect on the older children (75% reentered within 12 months) than younger children (44%).
- But as we saw a moment ago, it was relatively rare for an older child not to get the service ($4/20 = 20\%$) and much more likely for an infant ($14/20 = 70\%$).

**“I observe that... I think it’s because...”
...and the supporting **evidence**.**

What did you discover?

Your homework: Building the argument

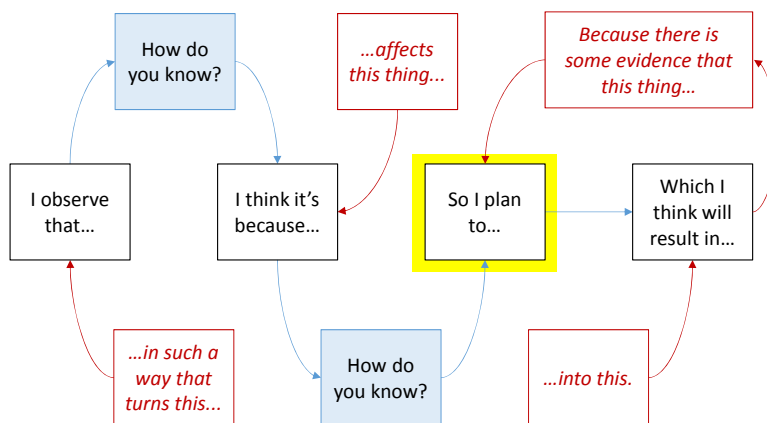
- Your “I observe that...” statement along with its supporting evidence
- Your “I think it’s because...” hypothesis
- An explanation of how you explored that hypothesis (i.e., an explanation of what questions you asked of whom)
- Your main findings from that exercise
- A statement about whether your hypothesis was supported

29

So I plan to...

30

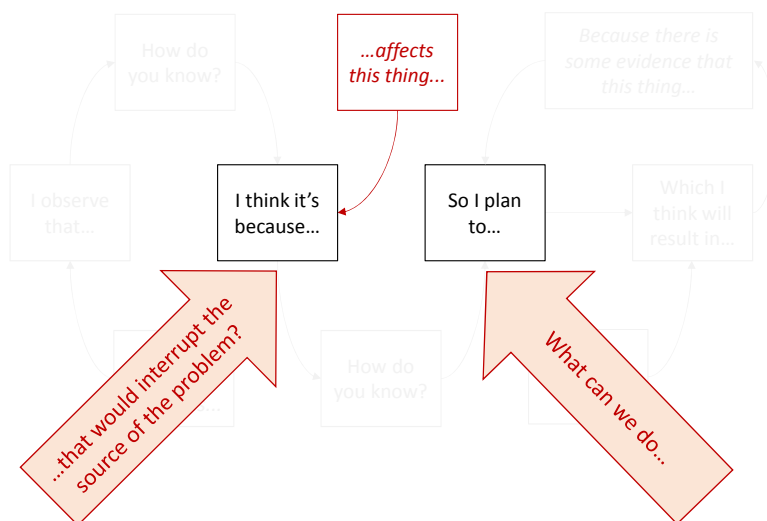
Theory of change



© The Center for State Child Welfare Data

31

Developing an intervention...



© The Center for State Child Welfare Data

32

Theory of change

You have observed the problem and have solidified your hypothesis as to its cause. Next you must make the argument for why your proposed solution can solve the problem.

- What are the components of the intervention? What will people do and when?
- Why and how is each component expected to produce the change you want to see?
- Is there any evidence to support the claim that this component will bring about the change you want to see? If yes, lay that out. If not, what, at least, is the theory?

33

REU practicum logic model

I observe that child welfare managers do not typically use evidence generated from administrative data to fuel their CQI decision-making.

I think it's because child welfare managers need to strengthen their REU knowledge and skills.

So I plan to implement an REU training program...

...which I think will result in participants using administrative data correctly and more frequently throughout the CQI process.

34

REU practicum logic model

What is the rationale for implementing this particular intervention? LOGIC MODEL

Component	How will this bring about the change I want to see?	Is there any evidence that this will work? If not, what, at least, is the theory?
Classroom based learning	Curriculum imparts new knowledge about best practices in measurement.	Pre/post tests from previous cohorts showing that knowledge improved.
Drills that involve practice with actual, available evidence sources	Repeated drills will make students more familiar with available evidence sources; skills will improve with practice.	Research literature on REU interventions noting the essential role of actual evidence sources. Research literature supporting the importance of repetition in learning new skills.
Group project in which students apply skills to a job relevant issue	Students will be more motivated to use evidence when they see the application to a real life scenario they are confronting.	Adult education literature; transfer of learning is optimized when project work links knowledge and skills to job-relevant tasks.
Coaching from experts	Coaches help students move up the learning curve and integrate new knowledge.	Coaching literature; transfer of learning is improved when experts facilitate learning process.

Developing an intervention...

Even if you have not gathered evidence to confirm your hypothesis, continue with this thought exercise...

Develop a theory of change: What can we do that would interrupt the source of the problem?

- What is the intervention? A change to process? Quality? Capacity?
- What makes you think that the intervention will bring about the change you want to see?

Think feasible...

- What does the intervention entail?
 - What needs to get done?
 - Who will do it and when?
- What resources does the intervention require? (time, money, supplies...)

Between now and next time

Start fleshing out your theory of change:
(Homework sheet on the Resource Barn)

1. Write a paragraph that summarizes your idea for your intervention. Make sure it starts like this:

I observe that _____. I think it's because _____. So I plan to...

2. Use the logic model template to start articulating your theory of change.

37