Northern California Training Academy
Fundamentals in Evidence-based Decision-Making

SESSION 2: Developing a hypothesis

July 31, 2018

Jennifer Haight, Daniel Webster
and Lily Alpert

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

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 your group 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
3:30 – 3:45	Assign homework & video call with Renee Boothroyd
3:45—4:00	Conclude

Last time…

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

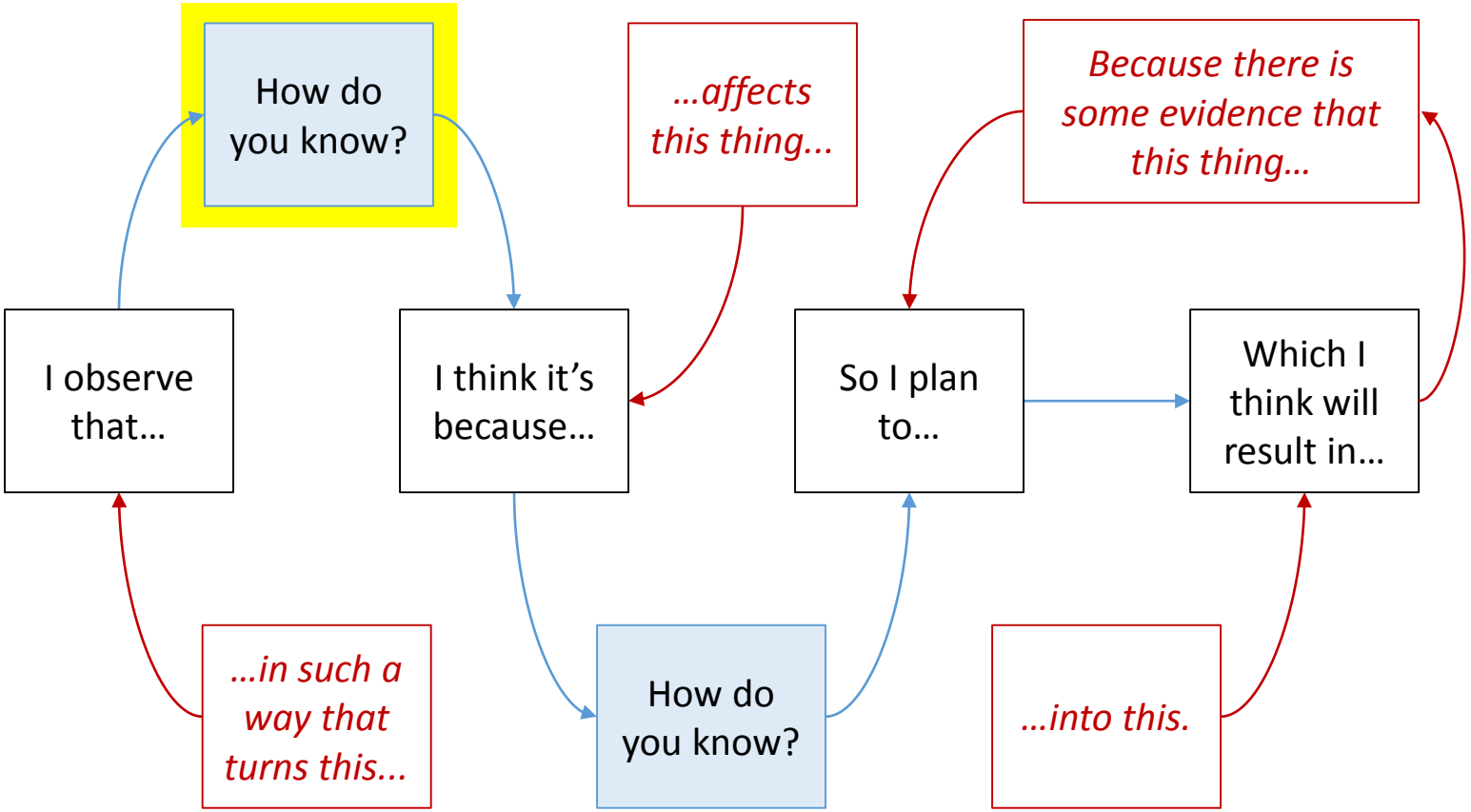
I think it's because of [**this reason**].

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 **permanency within 12 months of entry** in your county and to identify a problem that needs solving.

Theory of change



Some important concepts

1.

Case review:
Who should you compare to whom?

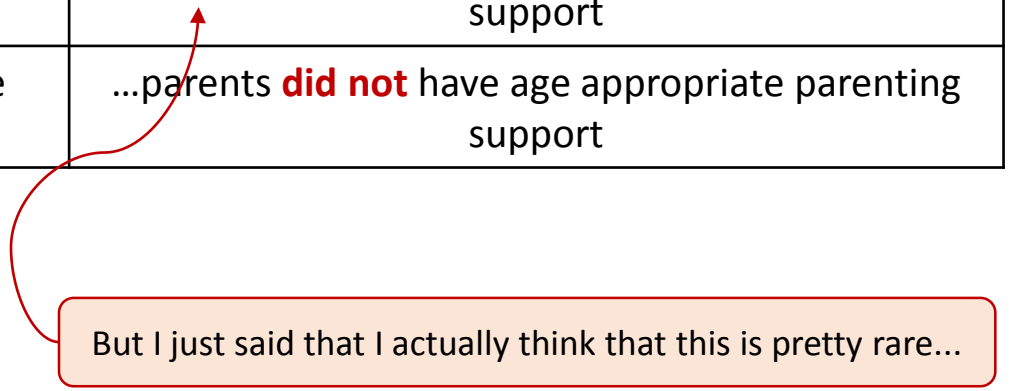
Example…

I observe that **younger** children are **less likely** to achieve permanency in 12 months compared to 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 younger children.

Some of you said… I bet if I **compared younger children who exited to permanency to those who didn't**, I would find that the parents of those who achieved permanency within 12 month received age-appropriate parenting support while their children were in care.

Among…	I expect to see that…
Younger children who exited to permanency within 12 months	…parents **did** have age appropriate parenting support
Younger children who did not achieve permanency within 12 months	…parents **did not** have age appropriate parenting support



In order build an argument that a lack of age-appropriate parenting support is the thing that makes **young children less likely than older children** to exit to permanency in 12 months (which is your observation), you'll need to compare children based on age.

Among children placed …	I expect to see that…	
	Exits to perm within 12 months	Did not exit to perm within 12 months
Under 6	…parents **did** have age appropriate parenting support	…parents **did not** have age appropriate parenting support.
6 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.

Random sampling and random samples from stratified selections

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 entered care in 2014 and 33% of them exited to permanency within 12 months:

	Perm < 12m	No perm < 12 m	Total
	Number		
Under 6	170	470	640
6 +	163	197	360
Total	333	667	1,000
	Percent of **grand total**		
Under 6	17%	47%	64%
6 +	16%	20%	36%
Total	33%	67%	100%

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

	Perm < 12m	No Perm < 12 m	Total
Under 6	7	18	25
6+	7	8	15
Total	14	26	40

Representative of the population with respect to age and outcome…but unequal and tiny cell sizes

If you want equal cell sizes, you could draw a random sample based on the combinations of the outcome of interest and the variable you think affects it.

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

Let's say 1,000 children entered in 2014 and 33% of them exited to permanency within 12 months:

	Perm < 12m	No Perm < 12 m	Total
	Number		
Under 6	170	470	640
6 +	163	197	360
Total	333	667	1,000
	Percent of **grand total**		
Under 6	17%	47%	64%
6 +	16%	20%	36%
Total	33%	67%	100%

If we draw a **random** sample of 40 cases in this way:

	Perm < 12m	Perm < 12 m	Total
Under 6	10	10	20
6 +	10	10	20
Total	20	20	40

Representative ***within each combination*** *but not representative of the population as a whole.*

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**.

	Perm < 12m	No Perm < 12 m	Total
Under 6	10	10	20
6+	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

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 entered care in Acme County in 2014 and 333 of those children achieved permanency within 12 months.

- **Rate of permanency within 12 months = 333/ 1000 = 33%**

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

"Of all children who entered care in Acme County in 2014, 33% exited to permanency within 12 months."

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

"Children entering care in Acme County have about a 33% *likelihood* of exiting to permanency 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.

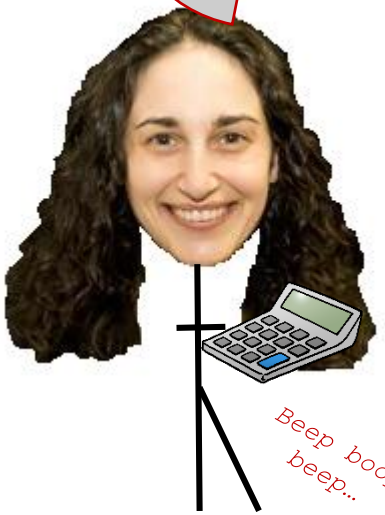
I'm going to calculate a fraction that tells me the **actual rates** of reunification and aging out for those past cohorts.

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

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

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

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



Beep boop beep...

Variables that predict the likelihood of an outcome

Different factors may increase or decrease **the likelihood of timely permanency.**

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	Permanency within 12 months? (Dependent variable)					
	Number			Percent		
	Perm Exit < 12 m	No Perm Ex < 12 m	Total	Perm Exit < 12 m	No Perm Ex < 12 m	Total
Total entered care	333	667	1,000	33%	67%	100%
Age at placement						
Under 6 y.o.	170	470	640	27%	73%	100%
6 y.o. and older	163	197	360	45%	55%	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 entered…")

Independent variable	Permanency within 12 months? (Dependent variable)					
	Number			Percent		
	Perm Exit < 12 m	No Perm < 12 m	Total	Perm Exit < 12 m	No Perm Exit < 12 m	Total
Total Entries	333	667	1,000	33%	67%	100%
Age at placement						
Under 6 y.o.	170	470	640	27%	73%	100%
6 y.o. and older	163	197	360	45%	55%	100%

Row percent

"Among the children who reunified, younger children (under 6) **were less likely** (27%) **to exit to permanency** within 12 months than children placed at age 6 or older (45%)."

But sometimes, fractions…

…And describe the CHARACTERISTICS of each group.

Categorize children by the OUTCOME…

	Independent variable	Permanency within 12 months? (Dependent variable)		
		Perm Exit < 12 m	No Perm Ex < 12 m	Total
Number	Total Entries	333	667	1,000
	Age at placement			
	Under 6 y.o.	170	470	640
	6 y.o. and older	163	197	360
Percent	Total reunified	100%	100%	100%
	Age at placement			
	Under 1 y.o.	51%	70%	64%
	1 y.o. and older	49%	30%	36%

Not the risk set! ("of all children who exited to perm within 12 months…")

"Selecting on the dependent variable"

Column percent

"Of children who exited to permanency within 12 months, 51% were 5 or younger at placement and 49% were placed at age 6 or older."

"Of children who exited to permanency within 12 months, 51% were under 6 at placement and 49% were placed at age 6 or older."

	Independent variable	Permanency within 12 months? (Dependent variable)		
		Perm Exit < 12 m	No Perm < 12 m	Total
Number	Total entries	333	667	1,000
	Age at placement			
	Under 6 y.o.	170	470	640
	6y.o. and older	163	197	360
Percent	Total Entries	100%	100%	100%
	Age at placement			
	Under 6 y.o.	51%	70%	64%
	6 y.o. and older	49%	30%	36%

This makes me think: If the group who achieve permanency about equally divided between older and younger children, then does that mean that all enterers are ***about equally likely to achieve timely permanency****?*

No, in fact the previous slide shows us that is not true: "Among the children who reunified, children placed as infants were less likely to exit to perm (27%) within 12 months than children placed at age 6 or older (45%)."

WHAT IS THE POINT FOR CRYING OUT LOUD.

It's important to be precise! The denominator is really important.

Of children who were under six at entry, what proportion exited to permanency within 12 months of entry?

170/640 = 27%

Of children who exited to permanency within 12 months of placement, what proportion were placed as babies?

170/333 = 51%

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 permanency children under 6 at placement…*

WHAT IS THE POINT FOR CRYING OUT LOUD.

We are trying to predict the **likelihood of an outcome** – namely, permanency in 12m months– so that we can learn something that helps us know how to promote 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 exit to permanency within 12 months)

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 exit to permanency within 12 months)

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

		Number		
		Perm <12	No Perm<12	Total
Number	Under 6	10	10	20
	AAPS	4	2	6
	No AAPS	6	8	14
	6 and older	10	10	20
	AAPS	9	7	16
	No AAPS	1	3	4
Percent	Under 6	100%	100%	100%
	AAPS	40%	20%	30%
	No AAPS	60%	80%	70%
	6 and older	100%	100%	100%
	AAPS	90%	70%	80%
	No AAPS	10%	30%	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, permanency cases got the service more often (40%, 90%) than non-perm (20%, 70%).
- Among children who did not exit to permanency within 12 months, not getting the service was more common for younger cases (80%) than it was for older enterers (30%).

What do we know at this point?

- We already observed that younger children are less likely to achieve permanency in 12 months than older children.
- Case review indicates that parents of younger children don't receive age-appropriate parenting services as often as parents of older children do.
- Case review also indicates parents of children who don't exit to permanency within 12 months also don't receive these services as often as parents of the children who do achieve timely permanency.

All of this is supportive to our hypothesis that the lack of these services is behind slower permanency for younger children

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

- *Of children whose parents received the service, what percent exited to permanency within 12 months.*
- *Of children whose parents didn't receive the service, what percent exited to permanency 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 achieved permanency within 12 months?* **Sample is based on the risk set (row percent)**

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

- Younger children whose parents got the service were more likely to achieve permanency within 12 months (67%) than those whose parents didn't (43%).
- The same was true for older children (56%, 25%). In fact, it seems like not receiving the service might have had an even stronger effect on the older children (75% did not achieve permanency within 12 months) than younger children (57%)
- 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 a younger child (14/20 = 70%).

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

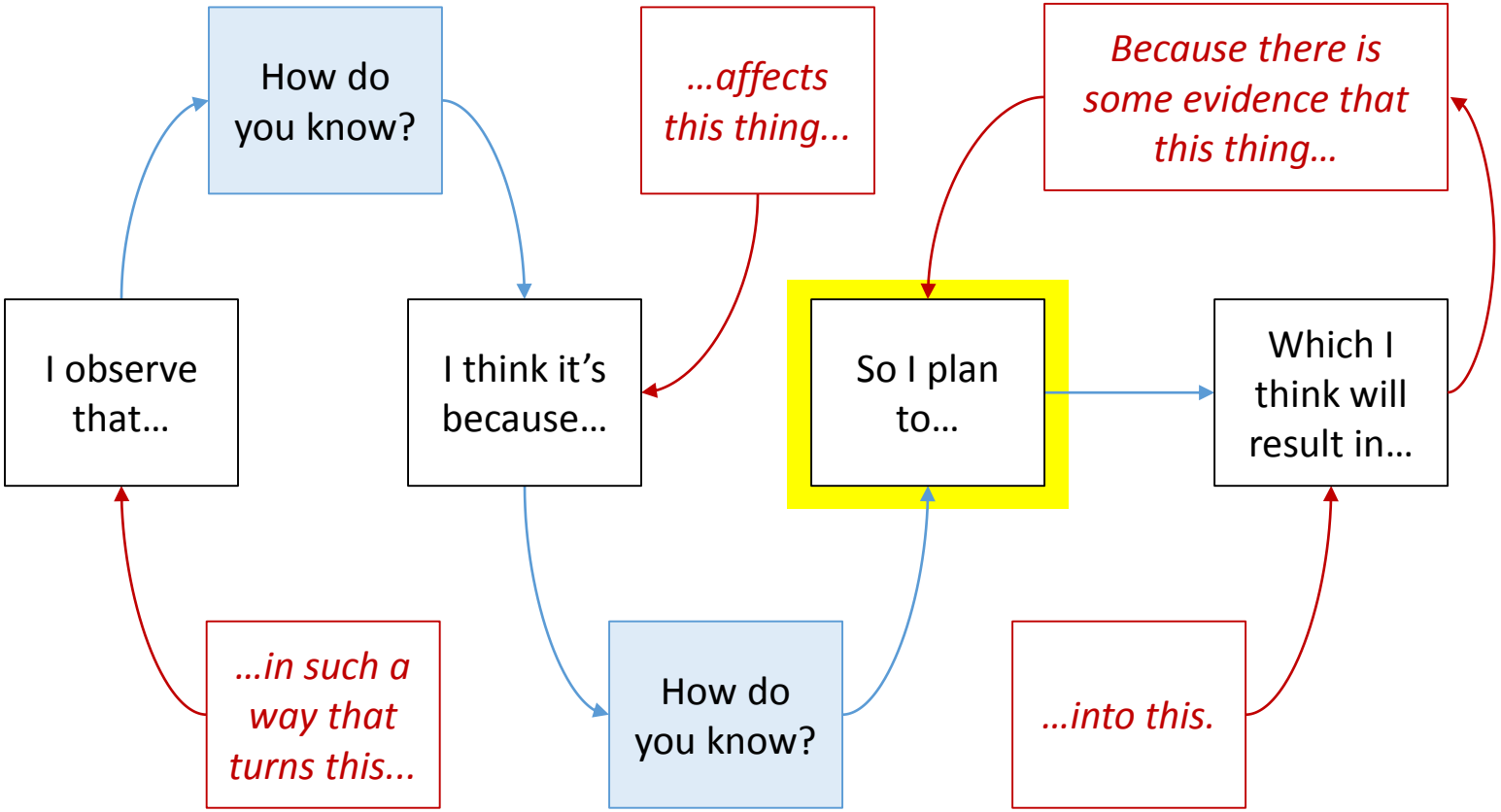
What did you discover?

Your homework: Building the argument

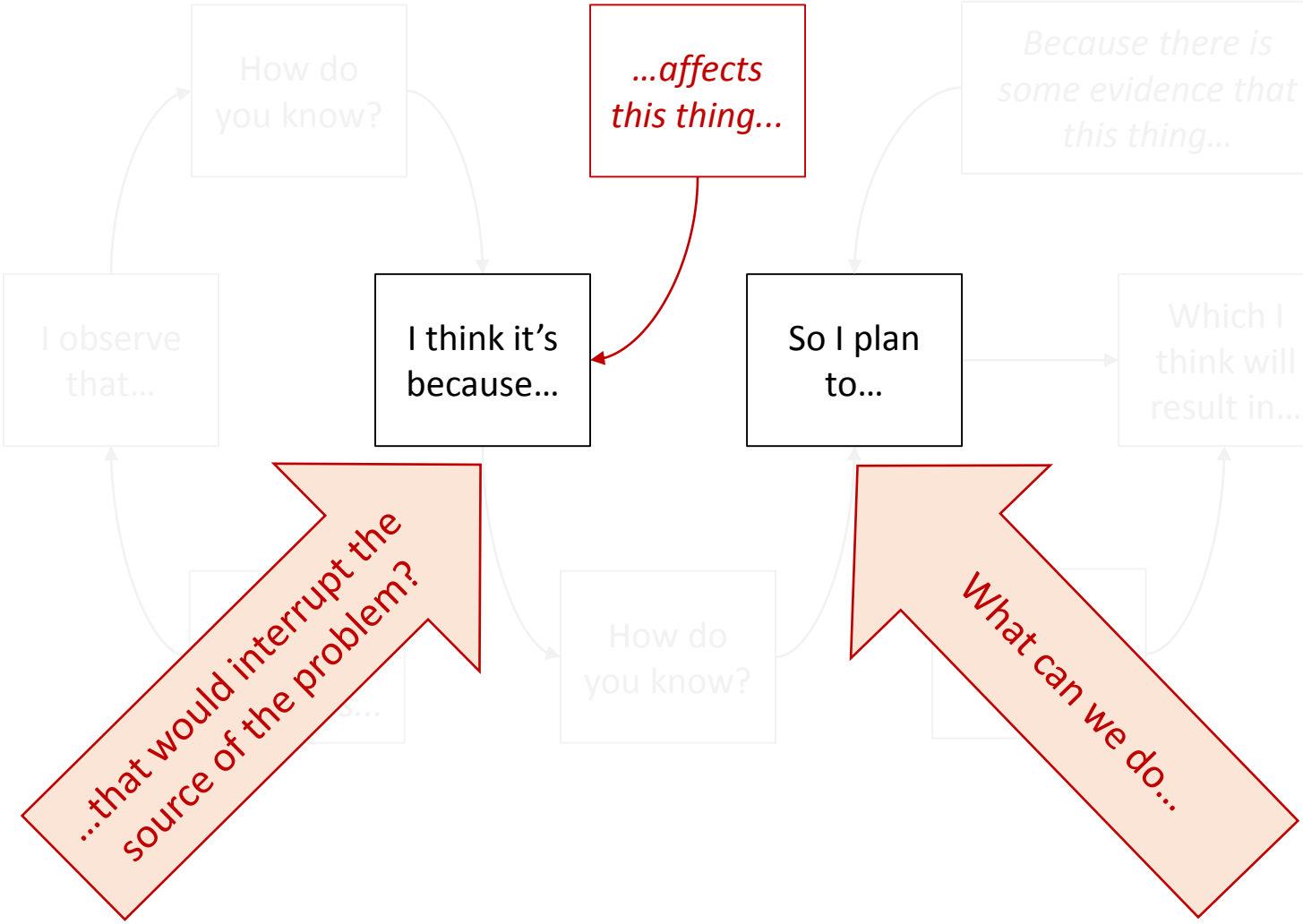
- Your “I observe that…” statement along with its supporting evidence
- Your “I think it’s because…” hypothesis
- An explanation of how you explored that hypothesis (i.e., an explanation of what questions you asked of whom)
- Your main findings from that exercise
- A statement about whether your hypothesis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

So I plan to…

Theory of change



Developing an intervention…



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?

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.

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.

Due dates

August 17	Email your assignment to Jennifer, Daniel and Lily
August 31	Receive feedback from Jennifer, Daniel and Lily
September 7	Email your revised/final assignment to Jennifer and Lily so they can be shared with Renee in advance of Session 3 (September 13).