Presented by
Tim Lycurgus, PhD

Clustered SMARTs

Module 8

60 min

Learning Objectives

- You will learn about what’s different in a clustered SMART with respect to:
  - Data analytics
  - Sample size considerations
  - Software

Outline

Clustered adaptive interventions [CAI]

Clustered SMARTs

New data analytics

Beta software

Outline

Clustered adaptive interventions [CAI]

Clustered SMARTs

New data analytics

Beta software

Review: Standard Clustered Interventions

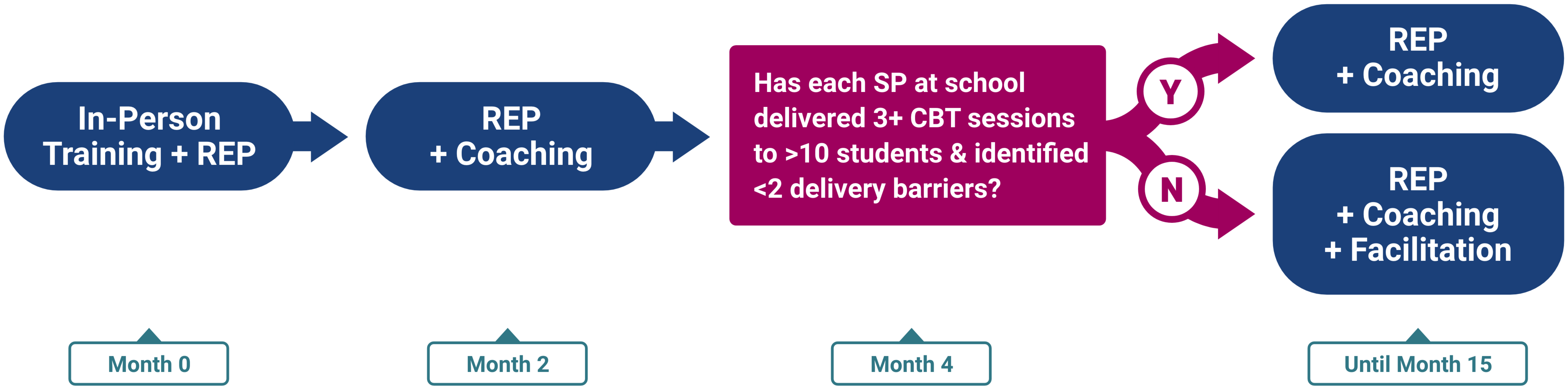
- In education, interventions often take place at the cluster (or organization) level (e.g. schools or classrooms).
- These cluster-level interventions are often designed to improve outcomes at units that are nested within each cluster.
- Examples:
    - Coaching school professionals to help children with mental health problems
    - Schoolwide Positive Behavior Supports in general education
- Clustered interventions can also be adaptive.

Clustered Adaptive Interventions

- Adapting and re-adapting cluster-level interventions may improve outcomes for the greatest number of units nested within a cluster.
- A clustered adaptive intervention (CAI) is a pre-specified set of decision rules that guides how best to serve baseline and ongoing needs of clusters from a pre-specified population.

PI: Amy Kilbourne

A Clustered AI for Implementing CBT in Schools



What are the intervention options in this example CAI?

Name	Description
Replicating Effectiveness Program (REP)	Didactic training in CBT for all school mental health staff and as-needed, ongoing technical assistance for school professionals.
Coaching	Provides live training to improve competence in providing CBT delivery.
Facilitation	Provides schools with opportunities to discuss and address barriers to CBT delivery with a “facilitator” who hosts monthly discussions with school staff responsible for coordinating and delivering CBT.

Clustered Adaptive Interventions

As with regular adaptive interventions, there are five components of a clustered AI:

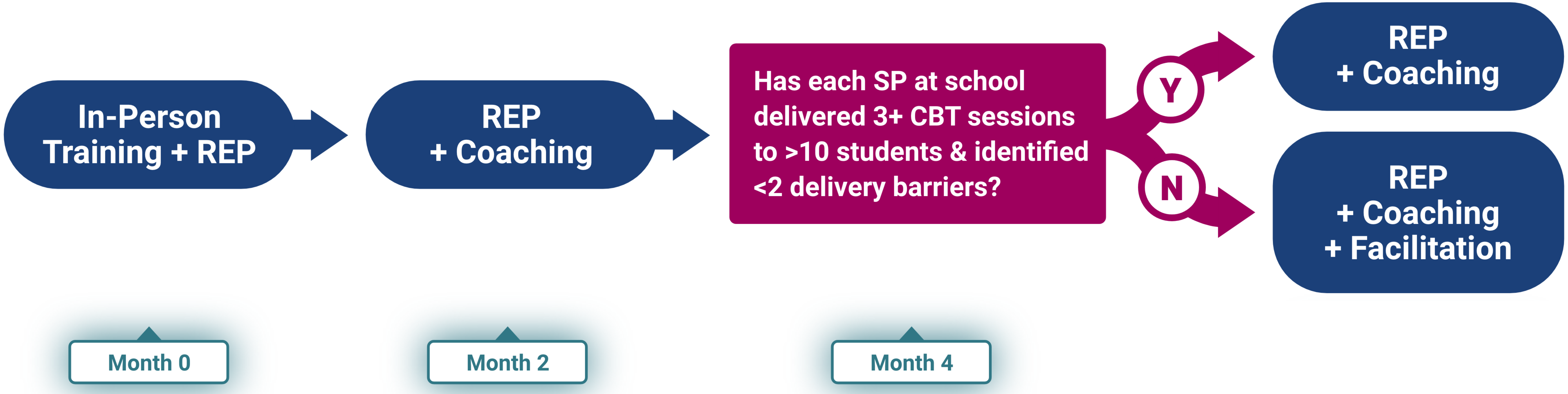
1) Proximal and distal outcomes
2) Decision points
3) Intervention options
4) Tailoring variables
5) Decision rules

But, there are different design considerations due to the nested nature of the AI.

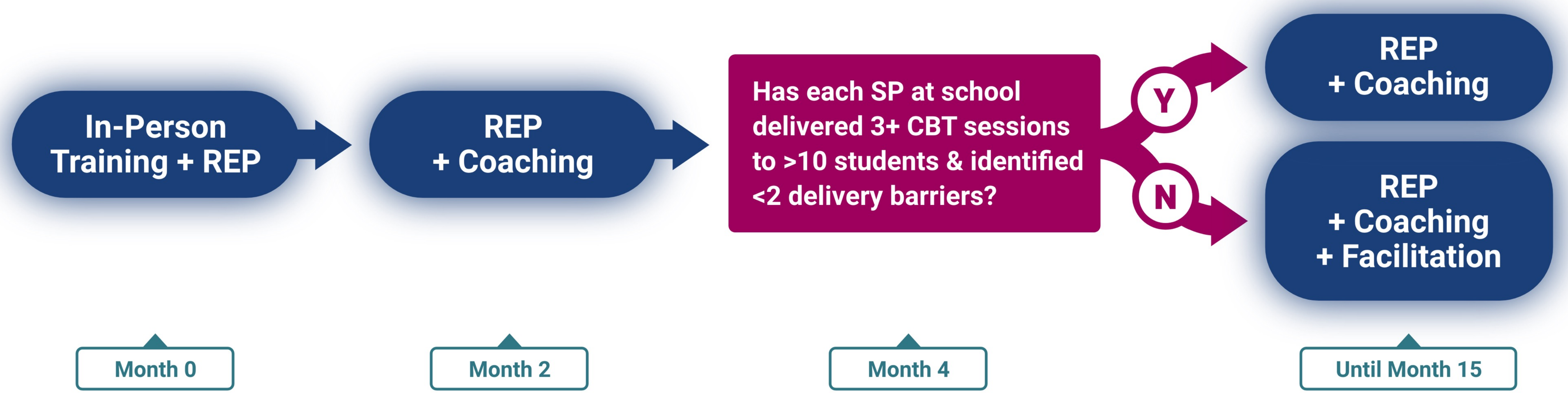
Proximal and Distal Outcomes of this CAI

- Proximal outcome (Primary goal) at the SP level: increase the number of CBT sessions delivered by school professionals within the school.
- Distal outcome at the SP level: improve CBT knowledge, perception, skills, etc. of school professionals within schools.
- Distal outcome at the student level: improvement in anxiety levels among students identified to be in need of CBT.

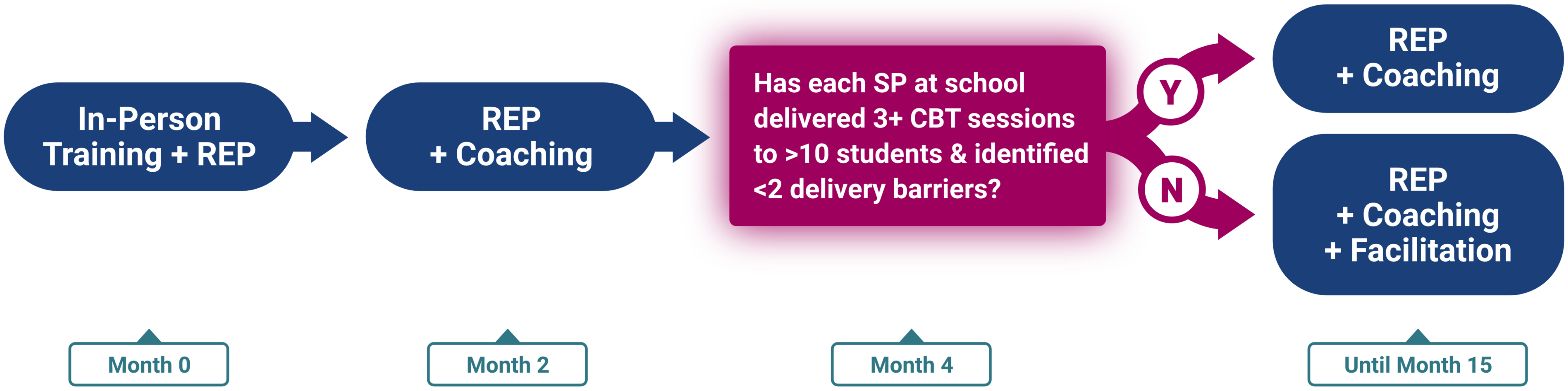
Decision Points



Intervention Options

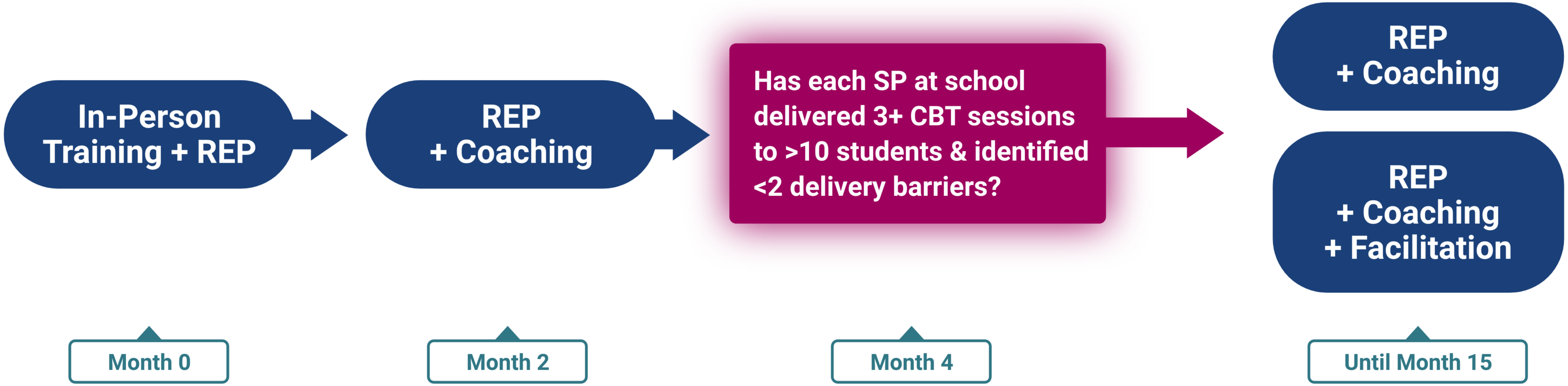


Tailoring Variables



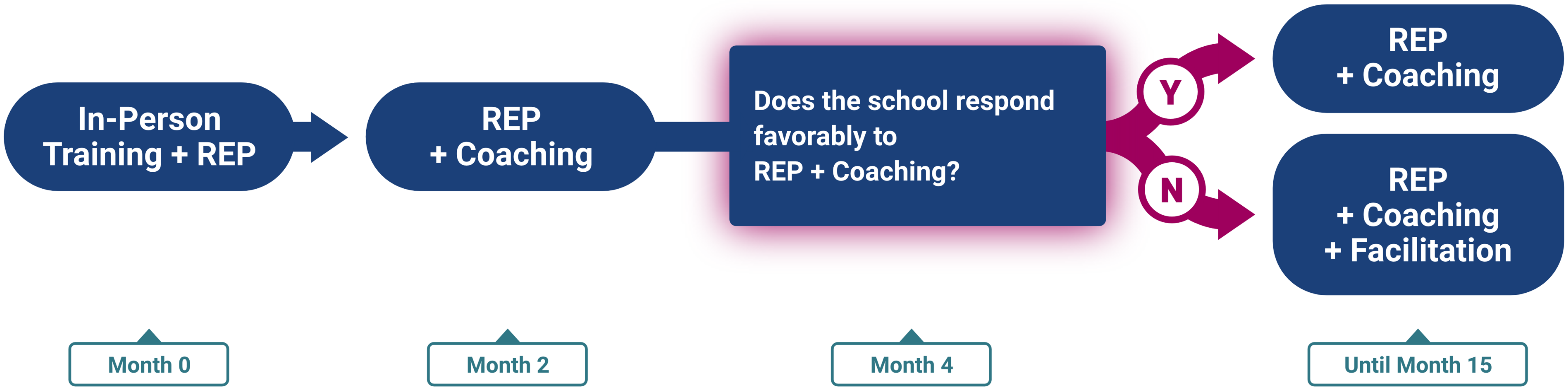
Tailoring Variables

- A tailoring variable is forward looking, in that it tailors an AI with the intention of improving end-of-study outcomes.

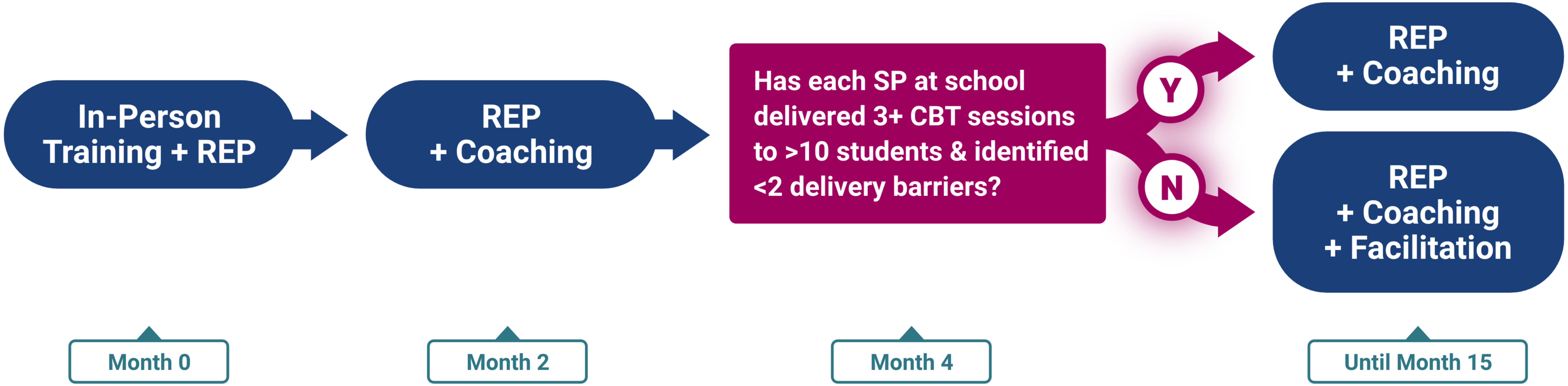


Tailoring Variables

- A tailoring variable may also be affected by prior intervention.



Decision Rules



But as education scientists, we may have many scientific questions about how best to optimize a CAI.

Optimization Questions

Can we do better than the CAI I showed earlier?

We might ask some of the following to optimize the clustered AI:

1. What is the effect of offering versus not offering Coaching in the first stage of the intervention?

2. Among schools who do not respond favorably to the initial intervention, should we augment their intervention with Facilitation?

3. Should we offer both Coaching and Facilitation in the first two stages of the intervention or should we offer neither of the two?

Other Optimization Questions

We could also ask questions other optimization questions such as:

- Example: Is it possible that first-stage interventions have no effect in the short-run, but have beneficial effects in the long-run when followed by a particular second-stage strategy?

- Example: Should Facilitation only be offered to sub-optimally responding schools within the lowest resourced school districts?

Outline

Clustered adaptive interventions [CAI]

Clustered SMARTs

New data analytics

Beta software

Clustered SMARTs

- Clustered SMART can be designed to answer such optimization questions.

- They are similar to SMARTs but with randomizations at the cluster level and outcomes at a nested level

- Clustered SMARTs are used to optimize clustered AIs.

Motivating Example SMART: ASIC

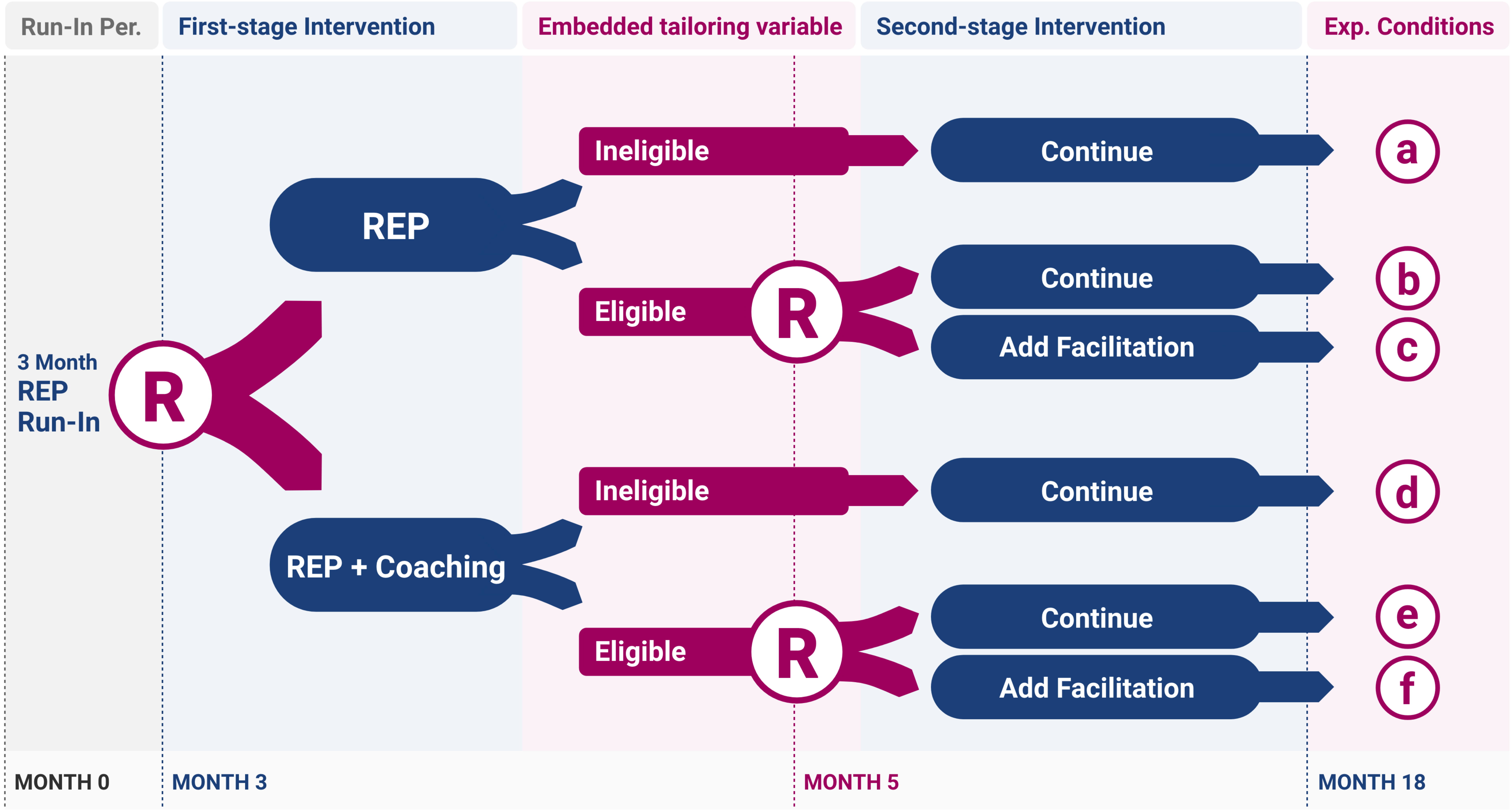
Adaptive School-based Implementation of CBT (ASIC) is a clustered SMART designed to optimize CAIs aimed at increasing delivery of CBT in school settings.

- ASIC takes place in 94 high schools in Michigan, constituting 200 school professionals.
- There are between 1 and 3 school professionals at each high school in Michigan.
- The primary outcome is the number of CBT sessions delivered by each school professional nested within each school.

SMART Example

ASIC: School-Based Implementation of CBT

PI: Kilbourne N=200



REP →

Replicating Effective Programs; low-level implementation strategy that provides manualization of intervention (e.g., CBT), didactic training, & technical assistance

Coaching →

In-person coaching during CBT groups at the school for a minimum 12 weeks

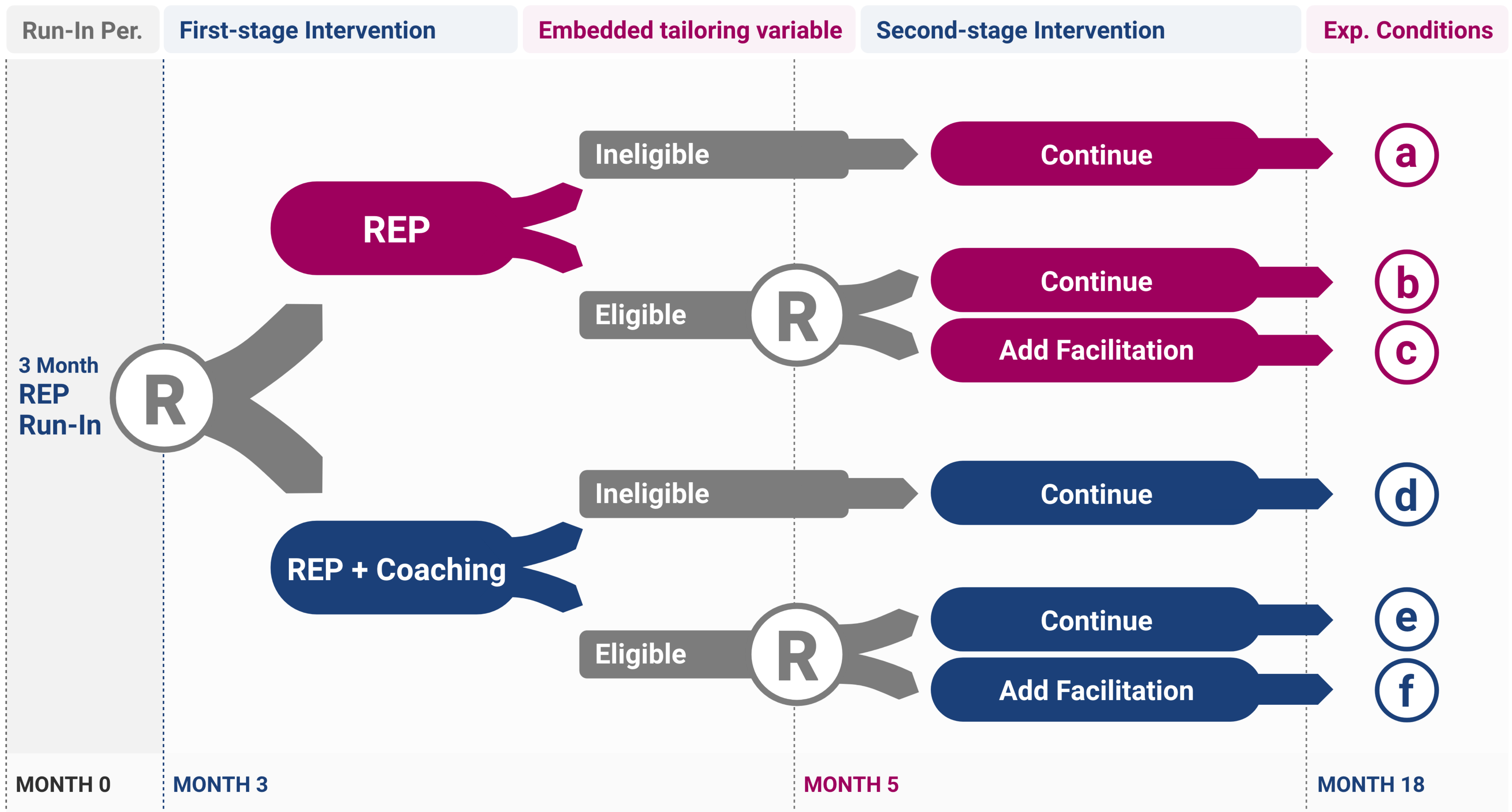
Facilitation →

Phone calls with an expert in CBT & strategic thinking for a minimum 12 weeks.

Three Common Primary Aims in a Clustered SMART

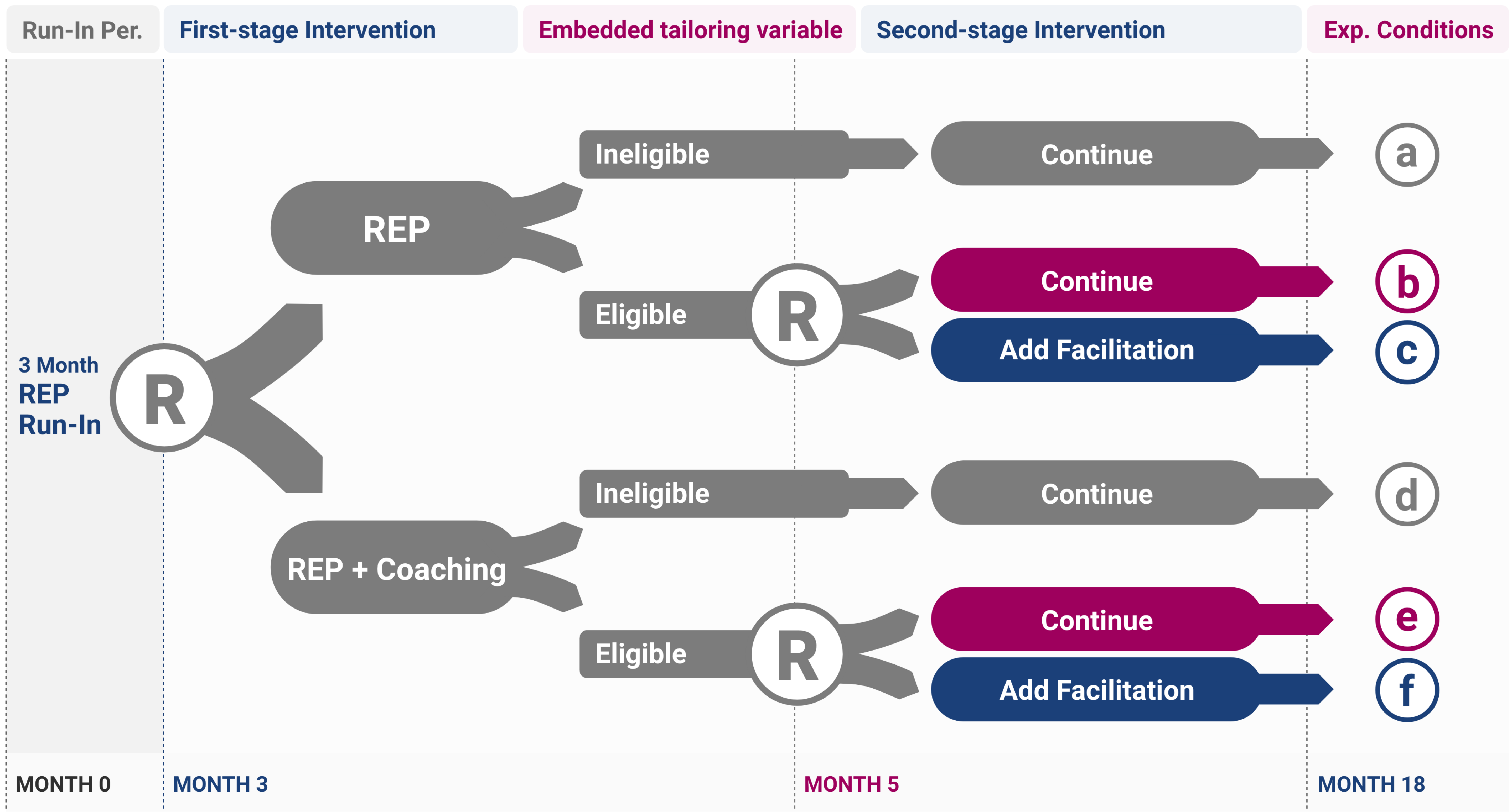
As before, there are three:

1. Main effect of the first stage of the intervention

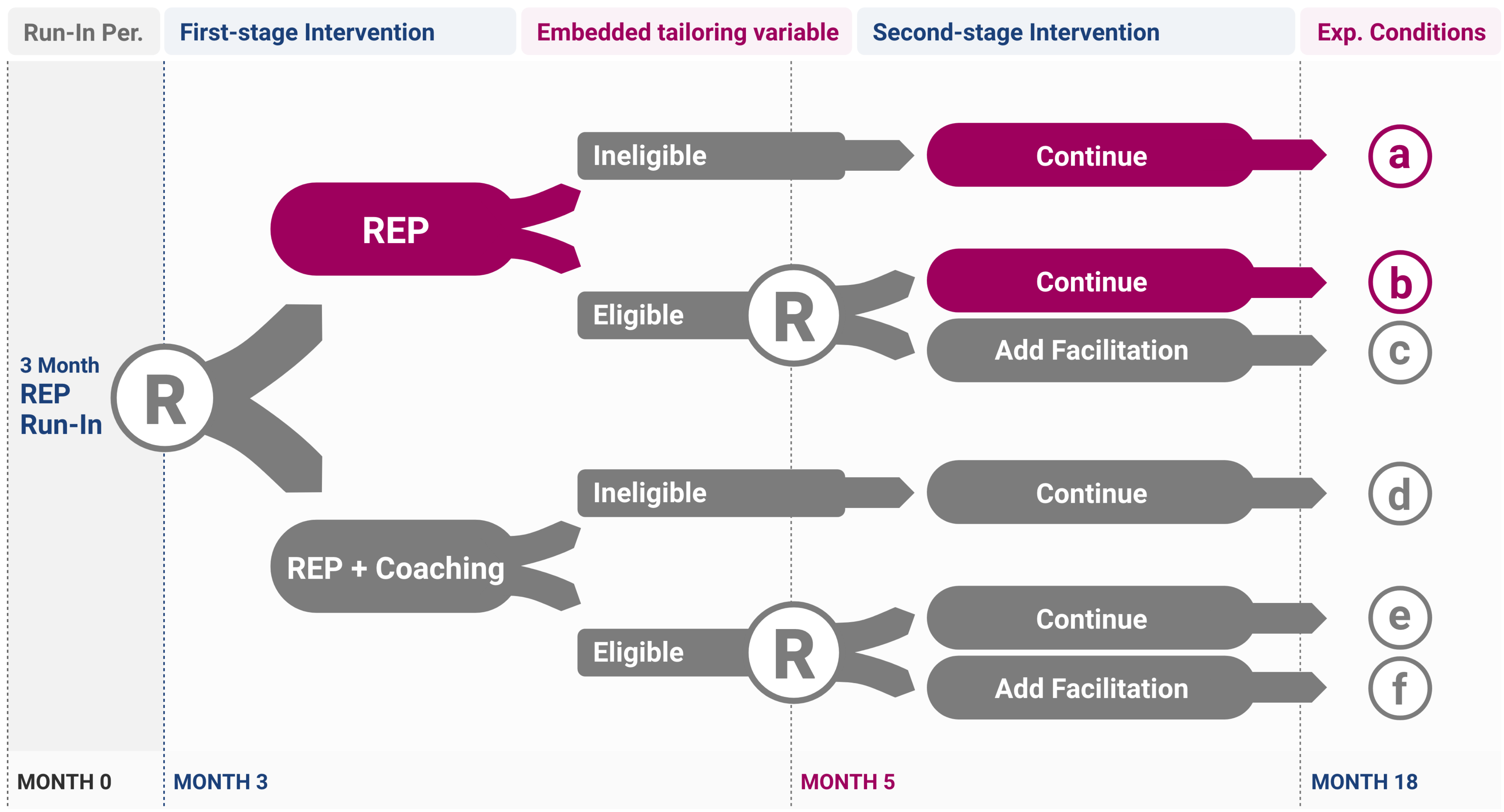


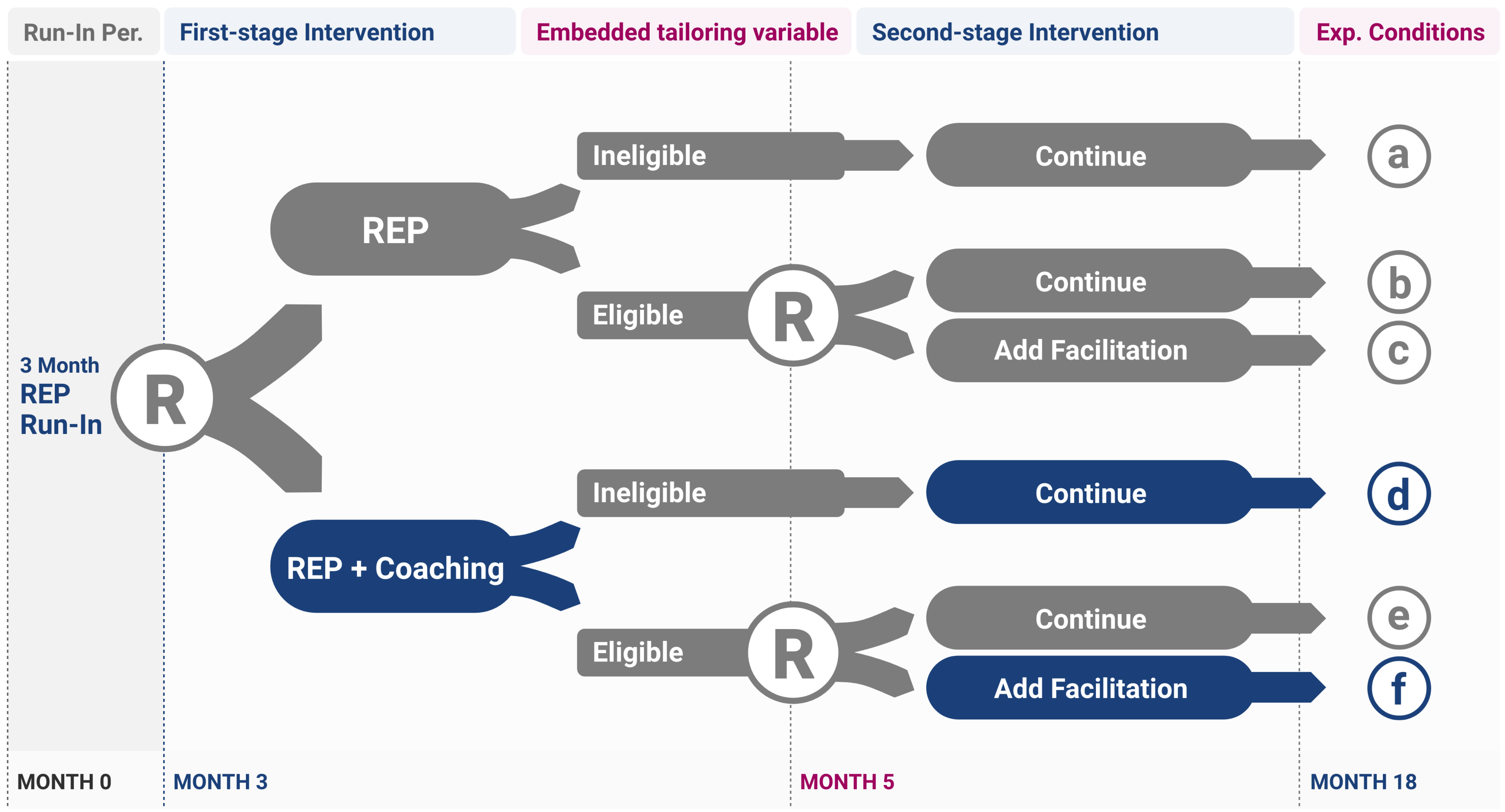
Three Common Primary Aims in a Clustered SMART

2. Main effect of the second stage of the intervention among non-responders



ASIC’s Primary Aim





3. Primary Aim: To test REP only (not a CAI) vs a CAI where schools receive REP + Coaching in the first stage, and then non-responding schools receive REP + Coaching + Facilitation and responders continue with REP + Coaching.

Outline

Clustered adaptive interventions [CAI]

Clustered SMARTs

New data analytics

Beta software

What’s new with methods in Clustered SMARTS?

- Multilevel Modeling Considerations
- Estimation Considerations
- Sample Size Formulae

Methodological innovations are all are due to the multilevel structure

- In standard SMART designs, everything occurs at the same level.
- In clustered SMARTs, we have units nested within clusters. In ASIC, we have school professionals nested within schools.
- In practice, this means:
    - We need to account for the correlation structure within clusters.
    - We need new software that allows us to account for this nested structure.

Intraclass Correlated (ICC) Outcomes

- With clustered SMARTs, we expect outcomes to be correlated within schools
- Intraclass Correlation (ICC): a measure of how similar outcomes of units are to one another within a given cluster.

Why do we have correlated outcomes?

In the context of CBT delivery within schools, the amount of CBT delivered by a SP within a school is likely to be correlated with the amount delivered by the other SPs in that school.

New Modeling Considerations

For Example

$$Y_{ij}(a_1, a_2) = \underbrace{\beta_0 + \beta_0 a_1 + \beta_2 a_2 + \beta_3 a_1 a_2}_{\text{Marginal structural Mean model: } \mu(a_1, a_2; \beta)} + \underbrace{e_{ij}}_{\text{Total Error}}$$

New Modeling Considerations

For Example

$$Y_{ij}(a_1, a_2) = \beta_0 + \beta_0 a_1 + \beta_2 a_2 + \beta_3 a_1 a_2 + e_{ij}$$

$$Y_{ij}(a_1, a_2) = \underbrace{\beta_0 + \beta_0 a_1 + \beta_2 a_2 + \beta_3 a_1 a_2}_{\text{Marginal structural Mean model: } \boldsymbol{\mu(a_1, a_2; \beta)}} + \underbrace{(\eta_j(a_1, a_2) + \epsilon_{ij})}_{\text{Total Error using Random Effects Modeling}}$$

where,

$$Mean(\eta_j(a_1, a_2)) = 0, Var\left(\eta_j(a_1, a_2)\right) = \sigma^2_{sch}$$
$$Mean(\epsilon_{ij}) = 0, Var\left(\epsilon_{ij}\right) = \sigma^2_{res}$$

New Modeling Considerations

For Example

$$Y_{ij}(a_1, a_2) = \beta_0 + \beta_0 a_1 + \beta_2 a_2 + \beta_3 a_1 a_2 + e_{ij}$$

$$Y_{ij}(a_1, a_2) = \underbrace{\beta_0 + \beta_0 a_1 + \beta_2 a_2 + \beta_3 a_1 a_2}_{\textbf{Marginal structural Mean model: } \boldsymbol{\mu(a_1, a_2; \beta)}} + \underbrace{(\eta_j(a_1, a_2) + \epsilon_{ij})}_{\textbf{Total Error using Random Effects Modeling}}$$

where,

$$Mean(\eta_j(a_1, a_2)) = 0, Var\left(\eta_j(a_1, a_2)\right) = \sigma_{sch}^2$$

$$Mean(\epsilon_{ij}) = 0, Var\left(\epsilon_{ij}\right) = \sigma_{res}^2$$

Now,

$$Var\left(Y_{ij}\right) = \sigma_{sch}^2 + \sigma_{res}^2 = \sigma_T^2$$

$$Cov\left(Y_{ij}, Y_{kj}\right) = \sigma_{sch}^2$$

$$\left.Cor\left(Y_{ij}, Y_{kj}\right) = \frac{Cov(Y_{ij}, Y_{kj})}{\sqrt{Var\left(Y_{ij}\right) Var\left(Y_{kj}\right)}} = \frac{\sigma_{sch}^2}{\sigma_{sch}^2 + \sigma_{res}^2} = \rho \right\} \boxed{\textbf{ICC!}}$$

New Modeling Considerations
For Example

Our decision below determines the structure of our working marginal variance model, $V(a_1, a_2)$:

$$Mean(\eta_j(a_1, a_2)) = 0, Var\left(\eta_j(a_1, a_2)\right) = \sigma^2_{sch}$$
$$Mean(\epsilon_{ij}) = 0, Var(\epsilon_{ij}) = \sigma^2_{res}$$

Leads to:

ICC!

$$Cor(Y_{ij}, Y_{kj}) = \frac{\sigma^2_{sch}}{\sigma^2_{sch} + \sigma^2_{res}} = \rho$$

Leads to a working Marginal variance model

$$\sigma^2_T \begin{pmatrix} 1 & \rho & \rho \\ \rho & 1 & \rho \\ \rho & \rho & 1 \end{pmatrix} = \mathbf{V}(a_1, a_2)$$

Note: this is for a school with 3 SPs

New Modeling Considerations

For Example

What happens to $V(a_1, a_2)$ if we assume independence of school professionals within schools?

$$Mean(\eta_j(a_1, a_2)) = 0, Var\left(\eta_j(a_1, a_2)\right) = 0$$
$$Mean(\epsilon_{ij}) = 0, Var(\epsilon_{ij}) = \sigma^2_{res}$$

Leads to:

ICC is zero here!

$$Cor\left(Y_{ij}, Y_{kj}\right) = \frac{0}{0 + \sigma^2_{res}} = 0$$

Leads to a working Marginal variance model

$$\sigma^2_T \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \mathbf{V}(a_1, a_2)$$

New Modeling Considerations

For Example

What happens to $V(a_1, a_2)$ if we assume the correlation of school professionals within schools depends on the CAI they receive?

$$Mean(\eta_j(a_1,a_2)) = 0, Var\left(\eta_j(a_1,a_2)\right) = \sigma^2_{sch}(a_1,a_2)$$
$$Mean(\epsilon_{ij}) = 0, Var(\epsilon_{ij}) = \sigma^2_{res}$$

Leads to:

ICC depends on the CAI here!

$$Cor(Y_{ij}, Y_{kj})(a_1, a_{2NR}) = \frac{\sigma^2_{sch}(a_1,a_2)}{\sigma^2_{sch}(a_1,a_2) + \sigma^2_{res}} = \rho_{a_1,a_{2NR}}$$

Leads to a working Marginal variance model

$$\sigma^2_T(a_1,a_2)\begin{pmatrix} 1 & \rho_{a_1,a_{2NR}} & \rho_{a_1,a_{2NR}} \\ \rho_{a_1,a_{2NR}} & 1 & \rho_{a_1,a_{2NR}} \\ \rho_{a_1,a_{2NR}} & \rho_{a_1,a_{2NR}} & 1 \end{pmatrix} = \mathbf{V}(a_1,a_2)$$

New Estimation Method

$$0 = \sum_{j=1}^{N} \sum_{a_1, a_{2NR}} I_j(a_1, a_{2NR}) W_j \boldsymbol{D}(a_1, a_{2NR})^T \boldsymbol{V_j^{-1}}(a_1, a_{2NR}) \left(\boldsymbol{Y_j} - \boldsymbol{\mu}(a_1, a_{2NR})\right)^2$$

Y_j is the vector of outcomes for school *j*: $\mathbf{Y_j} = \begin{pmatrix} Y_{1j}, & Y_{2j}, & Y_{3j} \end{pmatrix}^T$

- Even if you get the working variance model incorrect, it'll be okay! We've developed methods ensuring the causal effects are still unbiased and your hypothesis test is still valid!
- We also have easy to use software that will provide estimates by solving this equation for you!

New Sample Size Formulae

For comparing two embedded CAIs

Inputs:

- *m* is the number of units (e.g., school professionals) within each cluster.
- δ is the standardized effect size for the comparison.
- ρ is the outcome's intra-class correlation (ICC).
- *r* is the probability of response to the first stage intervention.

Outputs:

- *N* is the total number of observations needed.
- *n* is the number of clusters (e.g., schools) needed.

New Sample Size Formulae

For comparing two embedded CAIs

$$N = n \times m = \frac{4\left(z_{1-\frac{\alpha}{2}} + z_{1-\beta}\right)^2}{\delta^2} \times (1 + (m-1)\rho) \times (2 - r)$$

Inputs:

- *m* is the number of units (e.g., school professionals) within each cluster.
- δ is the standardized effect size for the comparison.
- ρ is the outcome's intra-class correlation (ICC).
- *r* is the probability of response to the first stage intervention.

Outputs:

- *N* is the total number of observations needed.
- *n* is the number of clusters (e.g., schools) needed.

New Sample Size Formulae

For comparing two embedded CAIs

$$N = n \times m = \frac{4\left(z_{1-\frac{\alpha}{2}} + z_{1-\beta}\right)^2}{\delta^2} \times (1 + (m-1)\rho) \times (2 - r)$$

This formula differs from what you've seen before in two ways:

- Inflation factor that is a function of the ICC and the cluster size (in red)
- SMART inflation factor (in blue)

Recommendations:

- This formula assumes that *m* is the same for each cluster. In practice, you may just want to use the mean number of individuals in a cluster for *m*.
- If the response rate *r* is expected to differ by first stage intervention option, then you can either:

  1) use the adjusted formula in Necamp et al. (2017). (See Handout 3).

  2) use the smaller hypothesized response rate for *r*.

Outline

Clustered adaptive interventions [CAI]

Clustered SMARTs

New data analytics

Beta software

Beta Software is Available

- We are in the process of developing software for clustered SMARTs.

⚠️ Please use the software with caution: We have not completed our testing. ⚠️
We recommend you join our newsletter so that you are notified when the software is more fully-tested and ready for wide distribution.

Beta Software is Available

- We are in the process of developing software for clustered SMARTs.
- SMARTutils R package:

Import the library

```
install.packages("devtools")
library(devtools)
```

```
install_github("AnilBattalahalli/SMARTutils")
library(SMARTutils)
```

- Documentation is available at the following github link: https://github.com/AnilBattalahalli/SMARTutils

⚠ Please use the software with caution: We have not completed our testing. We recommend you join our newsletter so that you are notified when the software is more fully-tested and ready for wide distribution. ⚠

Clustered Analysis: Exchangeable

```
> report <- cSMART.mm(Y~a1+a2+I(a1*a2), clustered_df, verbose=T, covstr = 'EXCH')
Parameter       Estimate        Std.Err         Z Score         Pr(>|z|)
___________     --------        -------         -------         -------
(Intercept)     0.61118         0.41444         1.474699        0.1402935
a1              0.98654         0.41444         2.380382        0.01729468
a2              0.74017         0.36136         2.048283        0.04053227
I(a1 * a2)      0.66370         0.36136         1.836645        0.06626231
___________     --------        -------         -------         -------

Marginal Mean Model:  Y ~ a1 + a2 + I(a1 * a2)

Working covariance structure: 'EXCH'  (Homogeneous-Exchangeable covariance structure)
Variance        98.67547
Correlation     0.1801472
```

What can go wrong if you don't account for clustering?

- Examine regression output where there is no nested structure (e.g., we just have 200 school professionals rather than SPs nested within schools).

	Estimates	Model SE	Robust SE	wald	p
(Intercept)	0.6112	0.2873	0.2971	2.057	0.0396500
A1	0.9865	0.2873	0.2971	3.321	0.0008971
A2	0.7402	0.2873	0.2971	2.492	0.0127200
I(A1 * A2)	0.6637	0.2873	0.2971	2.234	0.0254700

```
Estimated Correlation Parameter:  0
Correlation Structure:  independence
Est. Scale Parameter:  325.7
```

What can go wrong if you don't account for clustering?

Module 4 Strategy

```
             Estimates Model SE Robust SE  wal          p
(Intercept)     0.6112   0.2873    0.2971 2.05  0.0396500
A1              0.9865   0.2873    0.2971 3.32  0.0008971
A2              0.7402   0.2873    0.2971 2.49  0.0127200
I(A1 * A2)      0.6637   0.2873    0.2971 2.23  0.0254700

 Estimated Correlation Parameter:  0
 Correlation Structure:  independence
 Est. Scale Parameter:  325.7
```

Module 8 Strategy

```
> report <- cSMART.mm(Y~a1+a2+I(a1*a2), clustered_df, verbose=T, covstr = 'EXCH')
  rameter       Estimate      Std.Err      Z Score
 ---------      --------      -------      -------      -------
( tercept)      0.61118       0.41444      1.474699     0.1402935
                0.98654       0.41444      2.380382     0.01729468
                0.74017       0.36136      2.048283     0.04053227
I  1 * a2)      0.66370       0.36136      1.836645     0.06626231
 ---------      --------      -------      -------

Marginal Mean Model:  Y ~ a1 + a2 + I(a1 * a2)

Working covariance structure: 'EXCH'  (Homogeneous-Exchangeable covariance structure)
Variance        98.67547
Correlation     0.1801472
```

Future Work

- Future Work
- We are currently working on 3- and 4-level analytic methods for clustered SMARTs
- We are also working on multi-level SMART where the sequence of randomizations occur at multiple levels
- We are excited about using CAIs to engender positive spillover effects, which can improve academic outcomes; and about new methods to quantify and better understand such spillover effects

Thank you. Questions?

- We expect Clustered and Multilevel SMARTs to have wide applicability in education, given the natural clustering that occurs in education practice settings
- Questions?

Q&A

10 min