

# OPTIMIZING BEHAVIORAL INTERVENTIONS: AN INTEGRATION OF METHODOLOGICAL PERSPECTIVES FROM THE BEHAVIORAL AND ENGINEERING SCIENCES

Linda M. Collins, Ph.D.  
The Methodology Center and  
Department of Human Development & Family Studies  
Penn State

Presented at the Modern Modeling Methods Conference  
University of Connecticut  
May 21, 2013

# WHY BE INTERESTED IN OPTIMIZING BEHAVIORAL INTERVENTIONS?

- Every year
  - 443,000 deaths in the US due to smoking
  - 400,000 deaths in the US due to poor diet/inactivity
  - 35,000,000 deaths worldwide from noncommunicable diseases that stem from largely behavioral causes
  - And much additional morbidity, mortality, and economic loss directly or indirectly due to behavioral factors
- Behavioral interventions can save lives!

# OUTLINE

- Definitions
- What's wrong with business as usual?
- What is MOST? What is optimization?
- OK, how do you do this? Three examples
- Concluding remarks

# WHAT IS A BEHAVIORAL INTERVENTION?

- Definition: *A program aimed at modifying behavior for the purpose of treating or preventing disease, promoting health, improving academic achievement, and/or enhancing well-being*
- Examples:
  - School-based drug abuse prevention
  - Clinic-based smoking cessation
  - Weight loss treatment
- Most behavioral interventions are made up of multiple components.

# WHAT IS AN INTERVENTION COMPONENT?

- Definition: *Any aspect of an intervention that can be separated out for study*
  - Parts of intervention content
    - e.g., topics in a curriculum
  - Features that promote compliance/adherence
    - e.g., phone calls to parents to remind them to do home exercises with kids
  - Features aimed at improving fidelity
    - e.g., 800 number for teachers to call with questions

# WHAT IS AN INTERVENTION COMPONENT?

- Can impact efficacy, effectiveness, cost-effectiveness
- Some components may be pharmaceutical (e.g. ADHD treatment)
- Components can be defined at any level: individual, family, school, etc.

# OUTLINE

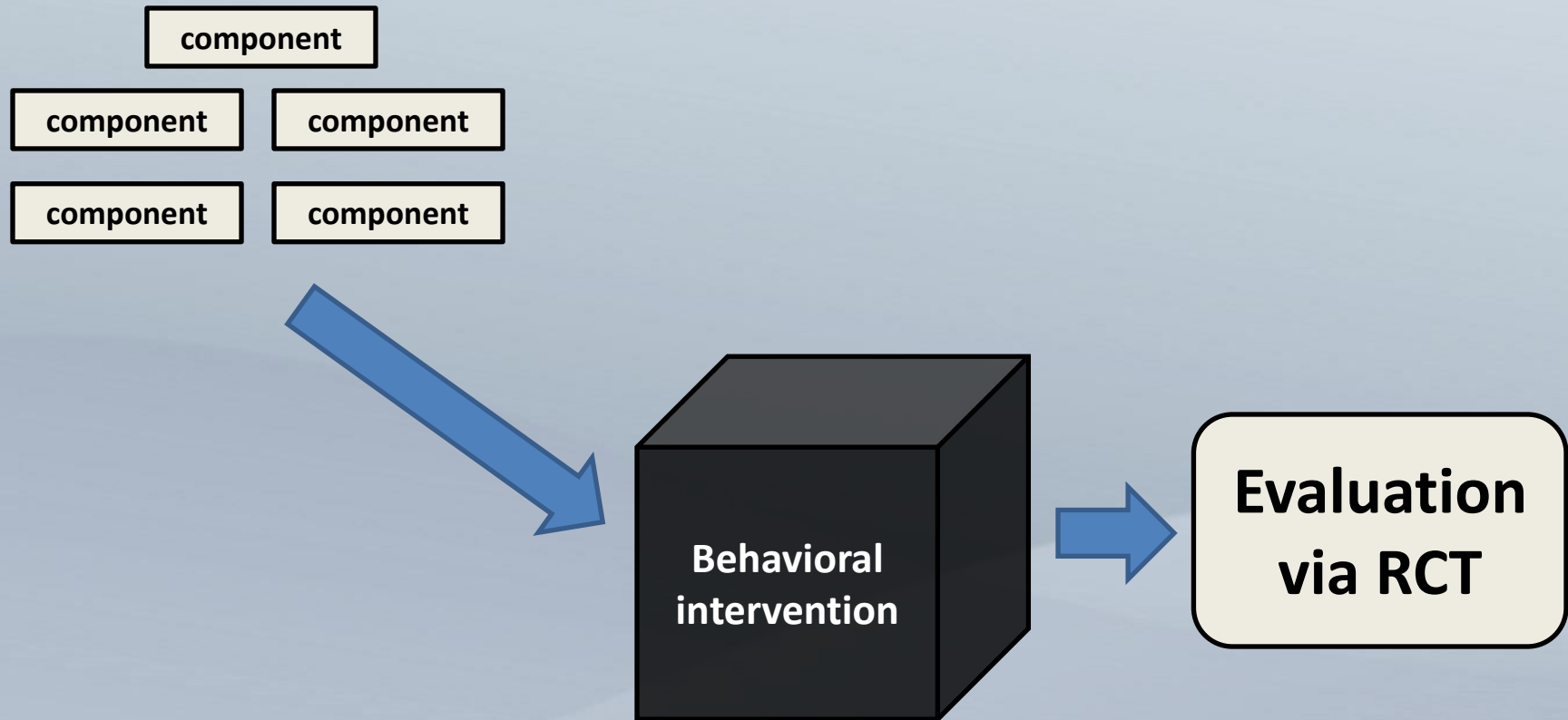
- Definitions
- What's wrong with business as usual?
- What is MOST? What is optimization?
- OK, how do you do this? Three examples
- Concluding remarks

# HOW BEHAVIORAL INVENTIONS ARE TYPICALLY DEVELOPED AND EVALUATED

- Intervention components are chosen based on scientific theory, clinical experience, etc.
- Combined into a package
- Package is evaluated via a randomized controlled trial (RCT)
- Let's call this the treatment package approach



# TREATMENT PACKAGE APPROACH



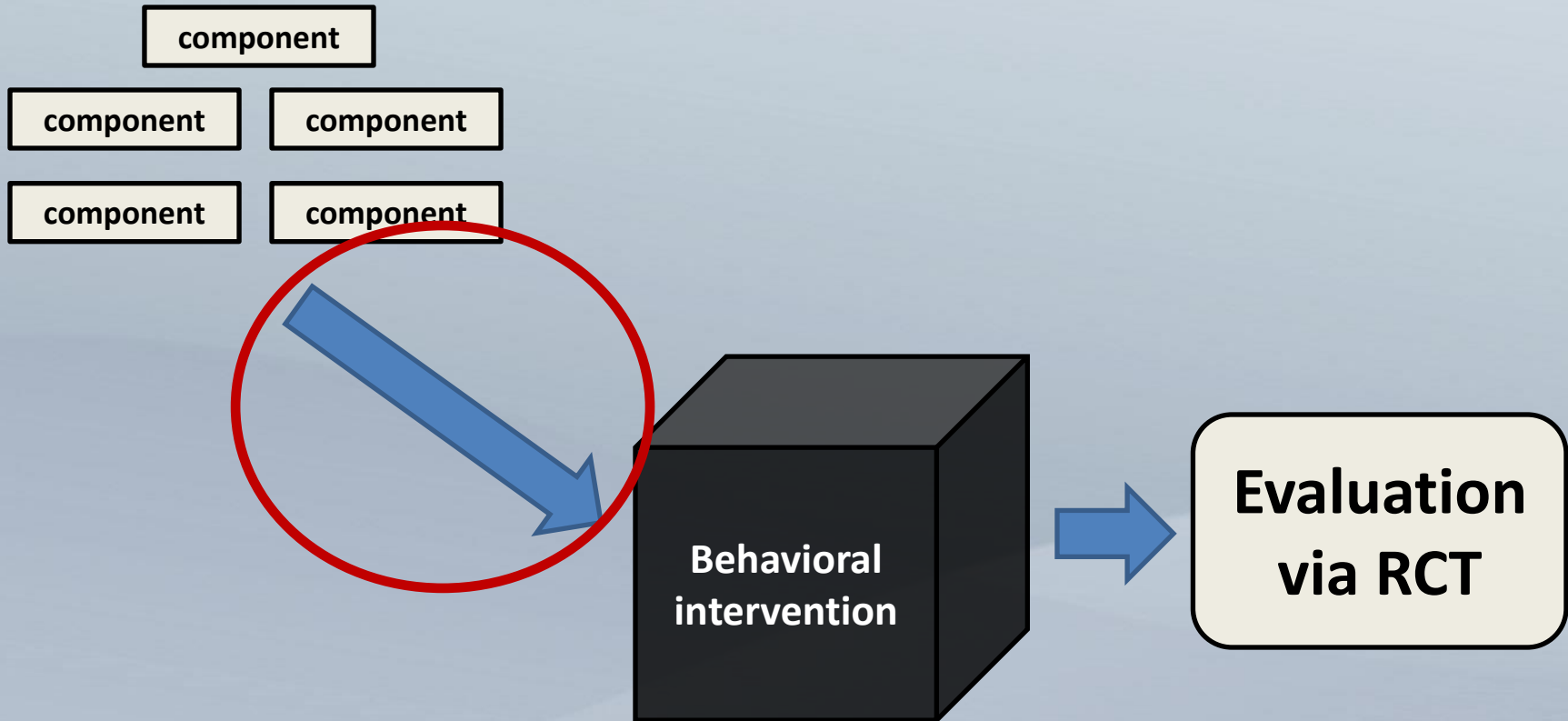
**What's wrong with evaluating a  
treatment package via an RCT?**

**Absolutely nothing!**

# THE RCT IS BEST SUITED FOR

- Determining whether a treatment package performs better than
  - A control or comparison group
  - An alternative intervention

# TREATMENT PACKAGE APPROACH



# WHAT THE RCT CANNOT NOT TELL US

An RCT that finds a significant effect DOES NOT tell us

- Which components are making positive contributions to overall effect
- Whether the inclusion of one component has an impact on the effect of another
- Whether a component's contribution offsets its cost
- Whether all the components are really needed
- How to make the intervention more effective, efficient, and scalable

# WHAT THE RCT CANNOT NOT TELL US

An RCT that finds a non-significant effect DOES NOT tell us

- Whether any components are worth retaining
- Whether one component had a negative effect that offset the positive effect of others
- Specifically what went wrong and how to do it better the next time

# RELIANCE ON THE TREATMENT PACKAGE APPROACH HAS ENCOURAGED...

...stuffing the behavioral intervention with many components to get a significant effect

...downplaying considerations such as

- efficiency
- cost-effectiveness and time-effectiveness
- participant burden
- scalability

# RELIANCE ON THE TREATMENT PACKAGE APPROACH HAS ENCOURAGED...

...focusing primarily on attaining statistical significance

...paying insufficient attention to meeting substantively meaningful criteria



# WHAT'S THE ALTERNATIVE?

- When engineers build products they take an approach that is
  - Systematic
  - Efficient
  - Focused on the clear objective of optimizing the product
- We can integrate methodological perspectives from the behavioral and engineering sciences...
- ... and build optimized behavioral interventions

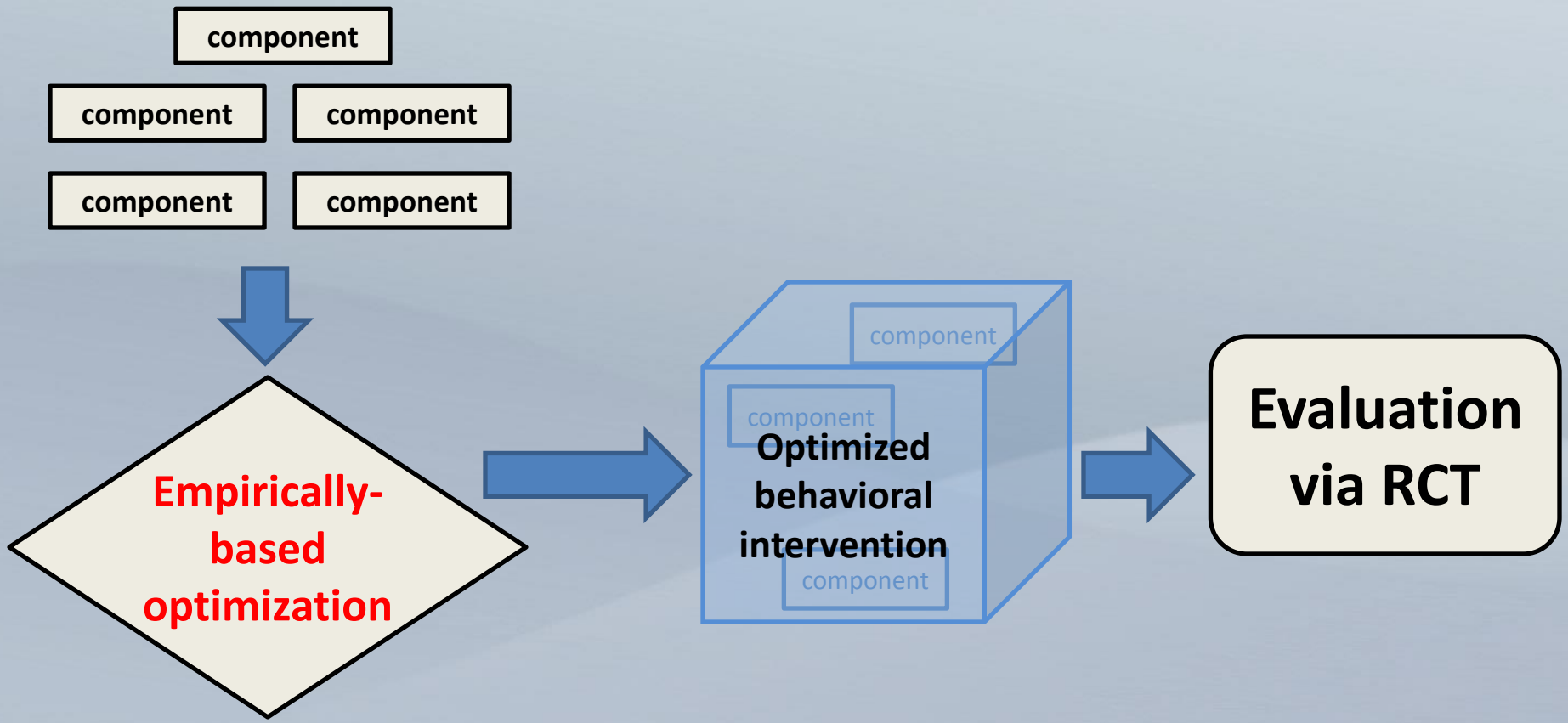
# OUTLINE

- Definitions
- What's wrong with business as usual?
- What is MOST? What is optimization?
- OK, how do you do this? Three examples
- Concluding remarks

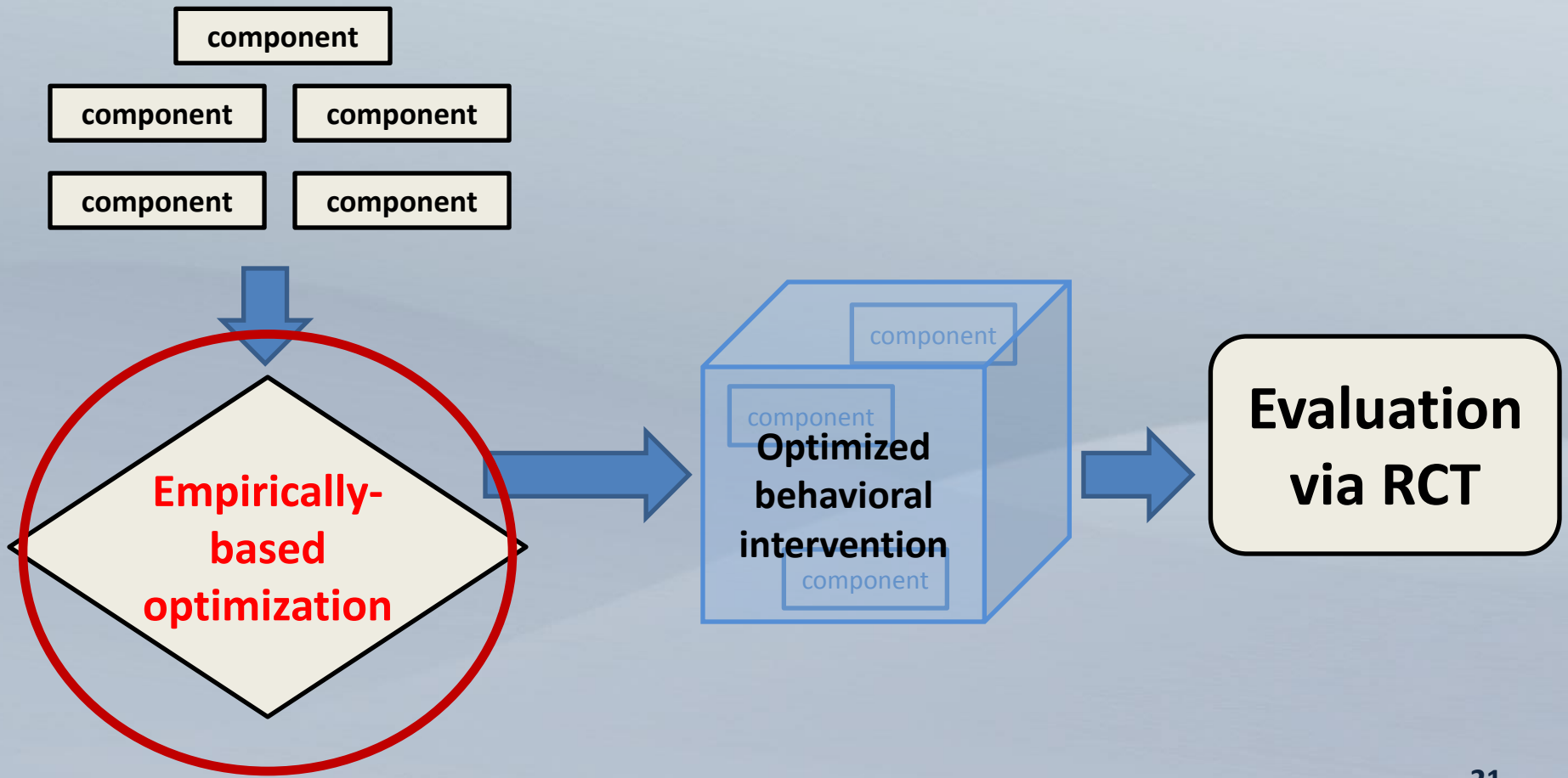
# THE MULTIPHASE OPTIMIZATION STRATEGY (MOST)

- An engineering-inspired framework for development, optimization, and evaluation of behavioral interventions
- Using MOST it is possible to engineer an intervention to meet specific criteria

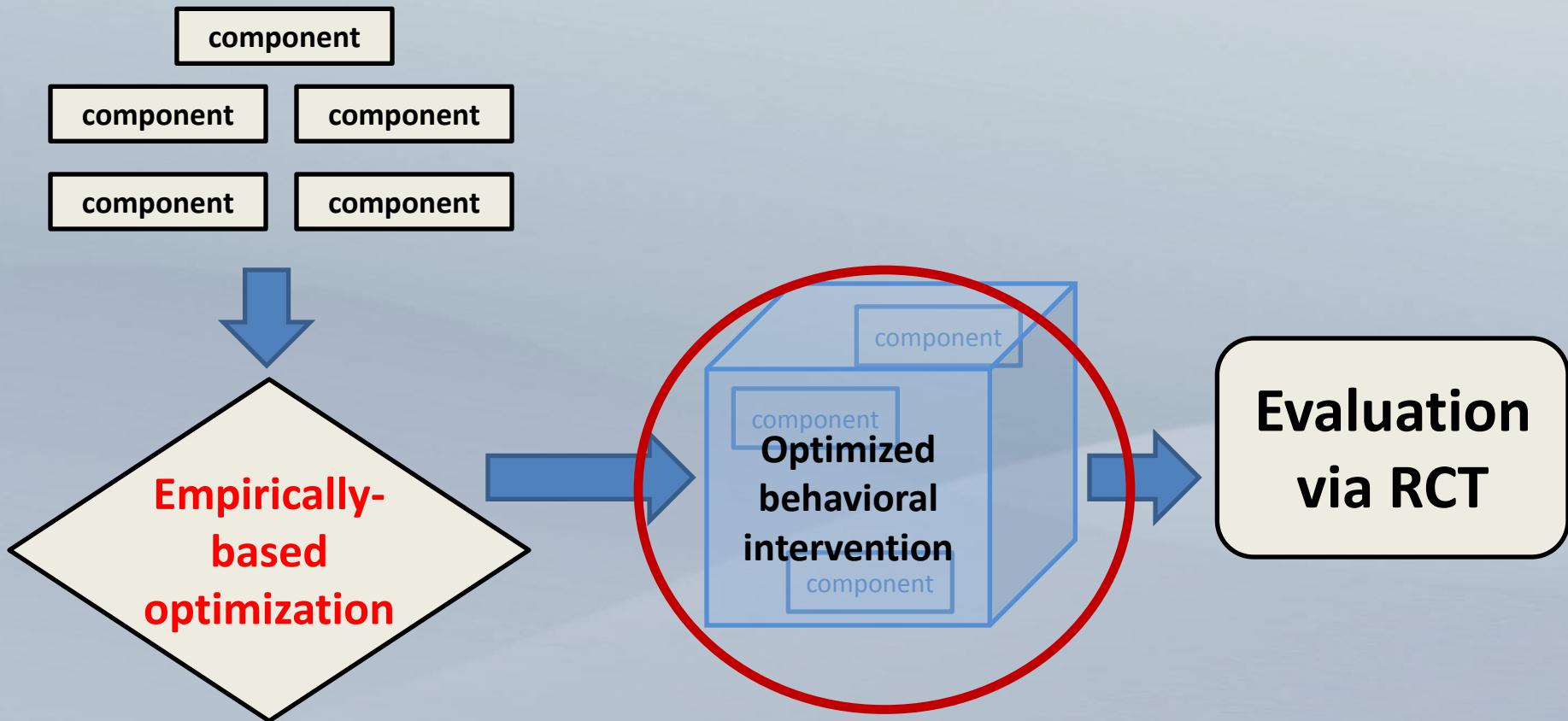
# MULTIPHASE OPTIMIZATION STRATEGY (MOST)



# MULTIPHASE OPTIMIZATION STRATEGY (MOST)



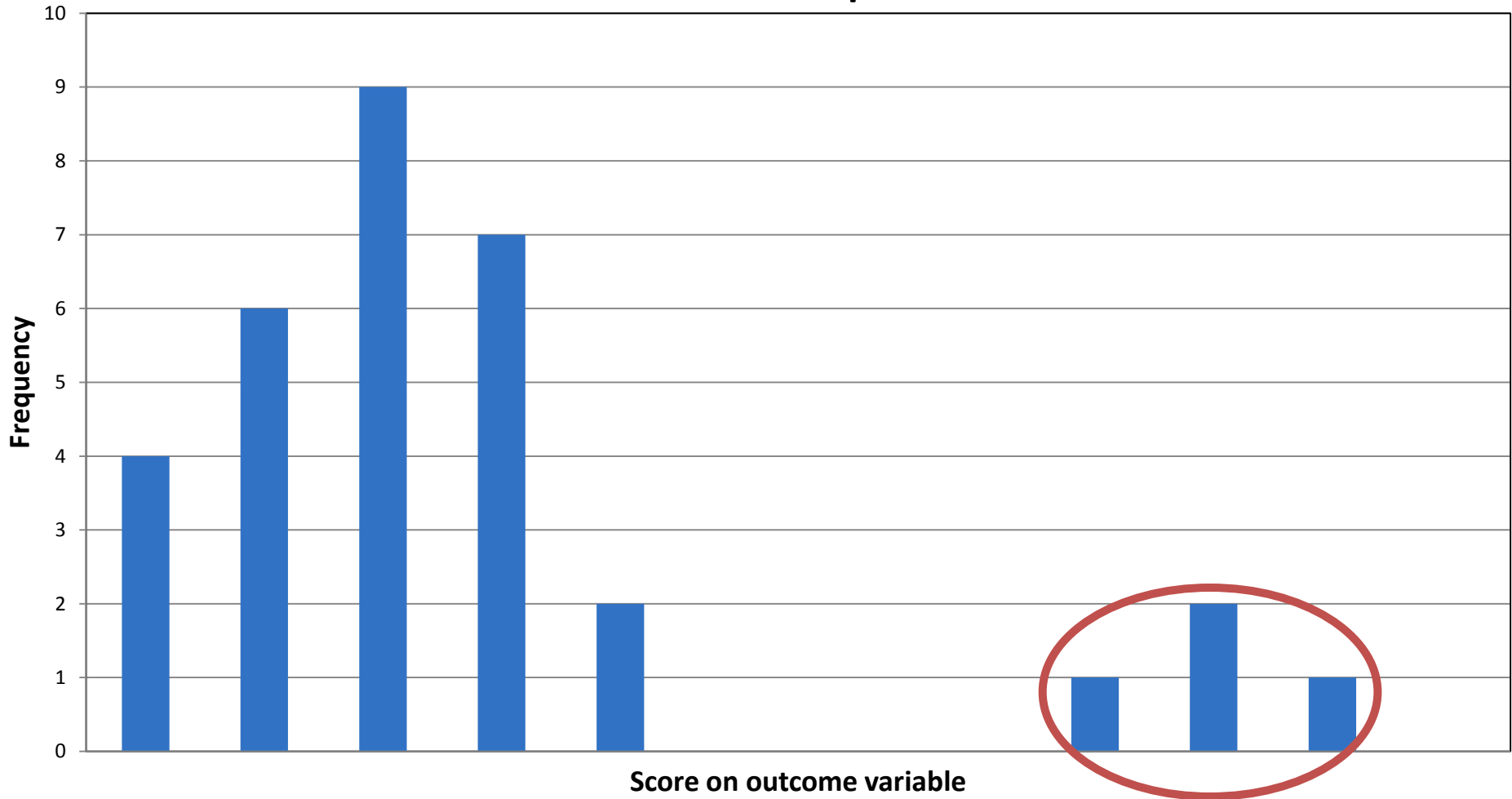
# MULTIPHASE OPTIMIZATION STRATEGY (MOST)



# LOGIC BEHIND MOST

- Logic: Objective is NOT to identify the ONE BEST combination of intervention components
  - To do this definitively would be prohibitively expensive
- Objective IS to identify ONE OF THE TWO OR THREE BEST approaches

### Hypothetical distribution of outcome variable for different combinations of components





# LOGIC BEHIND MOST

- Logic: Objective is to identify ONE OF THE TWO OR THREE BEST approaches
- Use the Resource Management Principle

# RESOURCE MANAGEMENT PRINCIPLE

- Conduct research to gain the most scientific information relevant to the research questions at hand, without exceeding available resources.
  - This is what I need to find out: \_\_\_\_\_
  - These are the resources I have: \_\_\_\_\_
  - HOW CAN I MANAGE MY RESOURCES STRATEGICALLY TO FIND OUT WHAT I NEED TO KNOW?

# LOGIC BEHIND MOST

- Logic: Objective is to identify ONE OF THE TWO OR THREE BEST approaches
- Use the Resource Management Principle
- Manage research resources strategically
  - Decide what information most important, and target resources there
- Choose efficient experimental designs
- Take calculated risks to move science forward faster

# LOGIC BEHIND MOST

- Note that the starting point for the Resource Management Principle is the resources you have
- By definition, MOST does not require an increase in research resources
  - But may require a *realignment* of research resources

# MOST: PREPARATION, OPTIMIZATION, EVALUATION

- Preparation
  - Purpose: to lay groundwork for optimization
    - Review prior research, take stock of clinical experience, conduct secondary analyses, etc.
    - Derive theoretical model
    - Select intervention components to examine
    - Conduct pilot/feasibility work
    - Identify clearly operationalized optimization criterion

# DEFINITION: OPTIMIZATION

- *“The process of finding the best possible solution... subject to given constraints”* (The Concise Oxford Dictionary of Mathematics)
  - Optimized does not mean best in an absolute or ideal sense
  - Instead, realistic because it includes constraints
- Optimization always involves a clearly stated optimization criterion

# SELECTING AN OPTIMIZATION CRITERION

- Your definition of “best possible, given constraints”
- This is the goal you want to achieve
- Once achieved, it is the bar that sets a standard for later efforts

# ONE POSSIBLE OPTIMIZATION CRITERION:

- Efficient intervention with no “dead wood”
- CONSIDER a school-based drug abuse prevention program.
  - Suppose the investigators want to be confident that every component is necessary to reduce waste of time and money.
  - Achieve this by selecting only active intervention components.



# ANOTHER POSSIBLE OPTIMIZATION CRITERION

- Most effective intervention that can be delivered for  $\leq$  some \$\$
- CONSIDER a school-based drug abuse prevention program.
  - Suppose school districts say they can afford a program that costs no more than \$25/child to implement, including materials and staff time.
  - Achieve this by selecting set of components that represents the most effective intervention that can be delivered for  $\leq$  \$25/child.

# ANOTHER POSSIBLE OPTIMIZATION CRITERION

- Most effective intervention that can be delivered in  $\leq$  some amount of time
- CONSIDER a school-based drug abuse prevention program.
  - Suppose interviews with teachers suggest that the program has the best chance of being implemented well if it takes no more than 3 classroom hours to deliver.
  - Achieve this by selecting set of components that represents the most effective intervention that can be delivered in  $\leq 3$  classroom hours.

# OTHER POSSIBLE OPTIMIZATION CRITERIA

- Cost-effectiveness
- A criterion based on a combination of cost and time
- Most effective without exceeding a specified level of participant burden
- Or any other relevant criterion

# MOST: PREPARATION, OPTIMIZATION, EVALUATION

- Preparation
  - Purpose: to lay groundwork for optimization
    - Review prior research, take stock of clinical experience, conduct secondary analyses, etc.
    - Derive theoretical model
    - Select intervention components to examine
    - Conduct pilot/feasibility work
    - Identify optimization criterion

# MOST: PREPARATION, OPTIMIZATION, EVALUATION

- Optimization
  - Objective: To form a treatment package that meets the optimization criterion
    - Collect and analyze empirical data on performance of individual intervention components relying on efficient randomized experiments
    - Based on information gathered, select components and levels that meet optimization criterion.

# MOST: PREPARATION, OPTIMIZATION, EVALUATION

- Evaluation
  - Objective: To establish whether the optimized intervention has a statistically significant effect compared to a control or alternative intervention
    - Conduct an RCT

**Evaluation and optimization:  
Both important;  
not the same thing.**

**Evaluation:  
Is the intervention's effect  
*statistically significant?***

**Optimization:  
Is the intervention  
the *best possible,*  
*given constraints?***

No

Yes

No

May wish to  
optimize to  
improve effect  
size

Intervention can  
probably be  
improved

Yes

Different  
intervention  
strategy needed

What we should  
be aiming for

# OUTLINE

- Definitions
- What's wrong with business as usual?
- What is MOST? What is optimization?
- OK, how do you do this? Three examples
- Concluding remarks



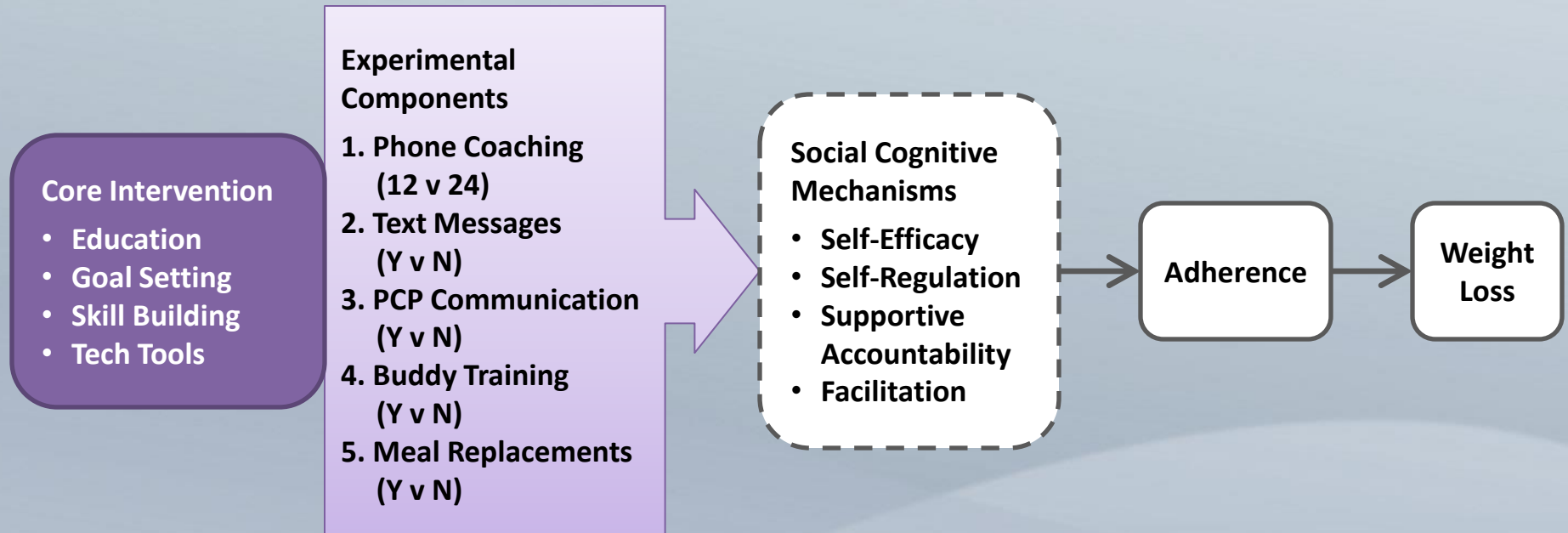
# EXAMPLE 1: “OPT-IN” WEIGHT REDUCTION INTERVENTION STUDY

Objective: Develop a highly effective weight reduction intervention that can be delivered for  $\leq$  \$500/person

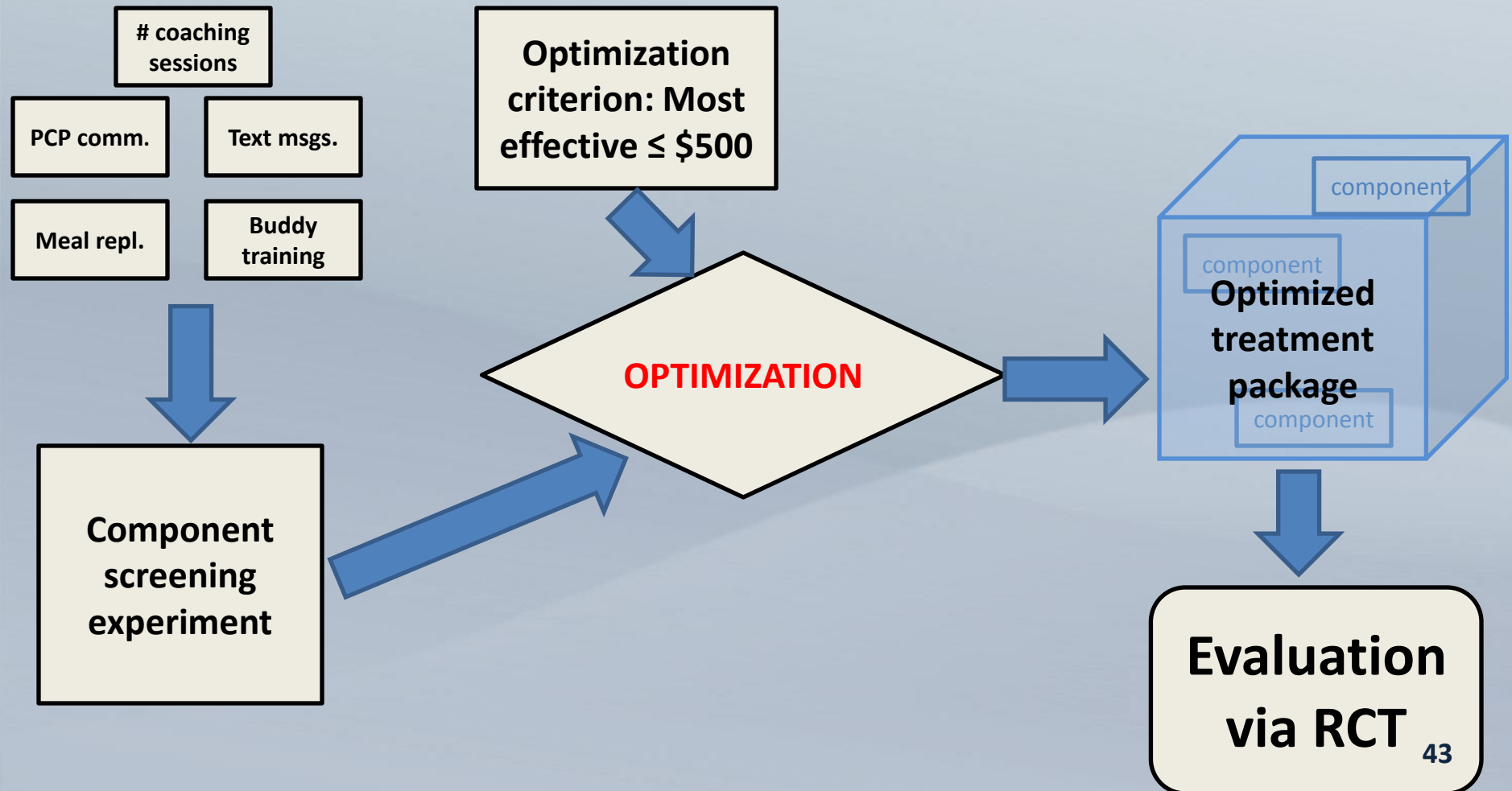
Investigative team: Bonnie Spring (NWU), Linda Collins (PSU), Christine Pellegrini (NWU)

Funded by the National Institute of Diabetes and Digestive and Kidney Diseases

# OPT-IN THEORETICAL MODEL



# MOST AS IMPLEMENTED IN OPT-IN



# THE COMPONENT SCREENING EXPERIMENT

- Purpose: efficient screening of intervention components
  - Weed out underperforming components
  - Get a sense of magnitude of each component's effect
  - Examine whether effect of a component is augmented or reduced in presence of another
- This information is then used to optimize the intervention

# CHOICE OF DESIGN FOR COMPONENT SCREENING EXPERIMENT IS CRITICAL

- Any experimental design is a possibility BUT...
- ...must be selected based on Resource Management Principle!!!

# RESOURCE MANAGEMENT PRINCIPLE

- To select a design, consider several, and examine
  - The scientific information each will provide
    - And whether it is what you want!
  - What each design costs
    - Number of subjects
    - Number of experimental conditions
- NOTE that the starting point is the resources you have

# EXPERIMENTAL DESIGN POSSIBILITY 1

- Conduct an experiment for each component

<b>Experiment 1</b>	<b>12 coaching sessions</b>	<b>24 coaching sessions</b>
<b>Experiment 2</b>	<b>No text messages</b>	<b>Text messages</b>
<b>Experiment 3</b>	<b>No PCP communication</b>	<b>PCP communication</b>
<b>Experiment 4</b>	<b>No buddy training</b>	<b>Buddy training</b>
<b>Experiment 5</b>	<b>No meal replacements</b>	<b>Meal replacements</b>

# EXPERIMENTAL DESIGN POSSIBILITY 2

- Comparative treatment experiment

<p><b>24 coaching sessions</b></p> <p>All others set to low</p>	<p><b>Text messages</b></p> <p>All others set to low</p>	<p><b>PCP communication</b></p> <p>All others set to low</p>	<p><b>Buddy training</b></p> <p>All others set to low</p>	<p><b>Meal replacements</b></p> <p>All others set to low</p>	<p><b>All set to low:</b>                  12 coaching sessions                  No text mgs                  No PCP                  No buddy training                  No meals</p>
---	--	--	---	--	---



# COMPARISON OF EXPERIMENTAL DESIGN POSSIBILITIES 1 AND 2

Design	N to achieve power $\geq .8$	Number of experimental conditions	Can interactions be examined?
Option 1: Five individual experiments	2,800	10	No
Option 2: Comparative treatment	1,680	6	No

# EXPERIMENTAL DESIGN POSSIBILITY 3

- Factorial experiment
- The Opt-In study would require a 2X2X2X2X2, or  $2^5$ , factorial experiment
- This would involve 32 experimental conditions.

# COMPARISON OF EXPERIMENTAL DESIGN POSSIBILITIES 1-3

Design	N to achieve power $\geq .8$	Number of experimental conditions	Can interactions be examined?
Option 1: Five individual experiments	2,800	10	No
Option 2: Comparative treatment	1,680	6	No
Option 3: Factorial experiment	560	32	Yes, all

- BUT we felt we could not handle more than 16 conditions.

# EXPERIMENTAL DESIGN POSSIBILITY 4

- Fractional factorial experiment
- Special type of factorial experiment
  - Used commonly in engineering research
  - A fraction of the experimental conditions are run
  - Powered exactly the same as an ordinary factorial experiment
  - Important trade-offs that we will discuss shortly

# COMPARISON OF EXPERIMENTAL DESIGN POSSIBILITIES 1-4

Design	N to achieve power $\geq .8$	Number of experimental conditions	Can interactions be examined?
Option 1: Five individual experiments	2,800	10	No
Option 2: Comparative treatment	1,680	6	No
Option 3: Factorial experiment	560	32	Yes, all
Option 4: Fractional factorial experiment	560	16	Yes, selected

## Design for Opt-In Component Screening Experiment

Experimental Condition	Core Intervention	# Phone Coaching Sessions	PCP Communication	Text Messages	Meal Replacements	Buddy Training
1	Yes	12	Yes	No	No	No
2	Yes	12	Yes	No	Yes	Yes
3	Yes	12	Yes	Yes	No	Yes
4	Yes	12	Yes	Yes	Yes	No
5	Yes	12	No	No	No	Yes
6	Yes	12	No	No	Yes	No
7	Yes	12	No	Yes	No	No
8	Yes	12	No	Yes	Yes	Yes
9	Yes	24	Yes	No	No	No
10	Yes	24	Yes	No	Yes	Yes
11	Yes	24	Yes	Yes	No	Yes
12	Yes	24	Yes	Yes	Yes	No
13	Yes	24	No	No	No	Yes
14	Yes	24	No	No	Yes	No
15	Yes	24	No	Yes	No	No
16	Yes	24	No	Yes	Yes	Yes

Overall  $N=560$ , per-condition  $n=35$ .

How can 35 per condition be enough?

## Design for Opt-In Component Screening Experiment

Experimental Condition	Core Intervention	# Phone Coaching Sessions	PCP Communication	Text Messages	Meal Replacements	Buddy Training
1	Yes	12	Yes	No	No	No
2	Yes	12	Yes	No	Yes	Yes
3	Yes	12	Yes	Yes	No	Yes
4	Yes	12	Yes	Yes	Yes	No
5	Yes	12	No	No	No	Yes
6	Yes	12	No	No	Yes	No
7	Yes	12	No	Yes	No	No
8	Yes	12	No	Yes	Yes	Yes
9	Yes	24	Yes	No	No	No
10	Yes	24	Yes	No	Yes	Yes
11	Yes	24	Yes	Yes	No	Yes
12	Yes	24	Yes	Yes	Yes	No
13	Yes	24	No	No	No	Yes
14	Yes	24	No	No	Yes	No
15	Yes	24	No	Yes	No	No
16	Yes	24	No	Yes	Yes	Yes

Main effect of # Phone Coaching Sessions *based on overall N of 560*

## Design for Opt-In Component Screening Experiment

Experimental Condition	Core Intervention	# Phone Coaching Sessions	PCP Communication	Text Messages	Meal Replacements	Buddy Training
1	Yes	12	Yes	No	No	No
2	Yes	12	Yes	No	Yes	Yes
3	Yes	12	Yes	Yes	No	Yes
4	Yes	12	Yes	Yes	Yes	No
5	Yes	12	No	No	No	Yes
6	Yes	12	No	No	Yes	No
7	Yes	12	No	Yes	No	No
8	Yes	12	No	Yes	Yes	Yes
9	Yes	24	Yes	No	No	No
10	Yes	24	Yes	No	Yes	Yes
11	Yes	24	Yes	Yes	No	Yes
12	Yes	24	Yes	Yes	Yes	No
13	Yes	24	No	No	No	Yes
14	Yes	24	No	No	Yes	No
15	Yes	24	No	Yes	No	No
16	Yes	24	No	Yes	Yes	Yes

Main effect of PCP Communication *based on overall N of 560*



## Design for Opt-In Component Screening Experiment

Experimental Condition	Core Intervention	# Phone Coaching Sessions	PCP Communication	Text Messages	Meal Replacements	Buddy Training
1	Yes	12	Yes	No	No	No
2	Yes	12	Yes	No	Yes	Yes
3	Yes	12	Yes	Yes	No	Yes
4	Yes	12	Yes	Yes	Yes	No
5	Yes	12	No	No	No	Yes
6	Yes	12	No	No	Yes	No
7	Yes	12	No	Yes	No	No
8	Yes	12	No	Yes	Yes	Yes
9	Yes	24	Yes	No	No	No
10	Yes	24	Yes	No	Yes	Yes
11	Yes	24	Yes	Yes	No	Yes
12	Yes	24	Yes	Yes	Yes	No
13	Yes	24	No	No	No	Yes
14	Yes	24	No	No	Yes	No
15	Yes	24	No	Yes	No	No
16	Yes	24	No	Yes	Yes	Yes

Main effect of Text Messages *based on overall N of 560*

## Design for Opt-In Component Screening Experiment

Experimental Condition	Core Intervention	# Phone Coaching Sessions	PCP Communication	Text Messages	Meal Replacements	Buddy Training
1	Yes	12	Yes	No	No	No
2	Yes	12	Yes	No	Yes	Yes
3	Yes	12	Yes	Yes	No	Yes
4	Yes	12	Yes	Yes	Yes	No
5	Yes	12	No	No	No	Yes
6	Yes	12	No	No	Yes	No
7	Yes	12	No	Yes	No	No
8	Yes	12	No	Yes	Yes	Yes
9	Yes	24	Yes	No	No	No
10	Yes	24	Yes	No	Yes	Yes
11	Yes	24	Yes	Yes	No	Yes
12	Yes	24	Yes	Yes	Yes	No
13	Yes	24	No	No	No	Yes
14	Yes	24	No	No	Yes	No
15	Yes	24	No	Yes	No	No
16	Yes	24	No	Yes	Yes	Yes

Main effect of Meal Replacements *based on overall N of 560*

## Design for Opt-In Component Screening Experiment

Experimental Condition	Core Intervention	# Phone Coaching Sessions	PCP Communication	Text Messages	Meal Replacements	Buddy Training
1	Yes	12	Yes	No	No	No
2	Yes	12	Yes	No	Yes	Yes
3	Yes	12	Yes	Yes	No	Yes
4	Yes	12	Yes	Yes	Yes	No
5	Yes	12	No	No	No	Yes
6	Yes	12	No	No	Yes	No
7	Yes	12	No	Yes	No	No
8	Yes	12	No	Yes	Yes	Yes
9	Yes	24	Yes	No	No	No
10	Yes	24	Yes	No	Yes	Yes
11	Yes	24	Yes	Yes	No	Yes
12	Yes	24	Yes	Yes	Yes	No
13	Yes	24	No	No	No	Yes
14	Yes	24	No	No	Yes	No
15	Yes	24	No	Yes	No	No
16	Yes	24	No	Yes	Yes	Yes

Main effect of Buddy Training *based on overall N of 560*

# FRACTIONAL FACTORIAL DESIGNS: TRADE-OFFS

- There are no free lunches in statistics
- WHAT WE GAIN USING A FRACTIONAL FACTORIAL DESIGN IN OPT-IN:
  - Reduce number of experimental conditions by half
  - Ability to examine five components instead of four
- WHAT WE GIVE UP:
  - Certain effects are estimated as a “bundle” with certain other effects. THIS IS CALLED ALIASING.

# FRACTIONAL FACTORIAL DESIGNS: TRADE-OFFS

- This design is a Resolution V
  - Each main effect is aliased with one four-way interaction
  - Each two-way interaction is aliased with one three-way interaction
  
- Logic:
  - Our theoretical model does not predict large three-way and four-way interactions
  - Therefore, the aliased estimate will be primarily due to main effect or two-way interaction

# FRACTIONAL FACTORIAL DESIGNS: TRADE-OFFS

	Higher-order effects negligible	Higher-order effects large
Complete factorial chosen (four components)	Resources wasted; cannot investigate important research questions	Move science forward faster
Fractional factorial chosen (five components)	Move science forward faster	Possibility of some incorrect decisions about component selection

- Sometimes maximizing efficiency calls for taking calculated risks
- There are opportunity costs associated with the “less risky” option
- *This is the Resource Management Principle in action*

# FRACTIONAL FACTORIAL DESIGNS: TRADE-OFFS

- Designs like the comparative treatment do not avoid aliasing

<b>24 coaching sessions</b>	<b>Text messages</b>	<b>PCP communication</b>	<b>Buddy training</b>	<b>Meal replacements</b>	<b>All set to low: 12 coaching sessions No text mgs No PCP No buddy training No meals</b>
<b>All others set to low</b>	<b>All others set to low</b>	<b>All others set to low</b>	<b>All others set to low</b>	<b>All others set to low</b>	

- Example: this aliases the main effect of Coaching with ALL interactions involving Coaching
  - Collins, Dziak, & Li (2009)

# USING DATA FROM THE EXPERIMENT TO OPTIMIZE

- Conduct an analysis of variance, obtain estimates of effects of each of the components
- Use this information to select components
  - Discard components that do not perform adequately
  - Use size of effects in combination with other data (e.g. cost) to select components that will make up optimized intervention



# STATUS OF OPT-IN PROJECT

- Finalizing intervention protocols and manual of operations
- Expect to start the experiment late summer/early fall

# AFTER WE HAVE OPTIMIZED AND EVALUATED THIS INTERVENTION

- We will have set a bar for weight loss interventions
- Our work will establish which components work
- Future work (by us or others) can build on this to develop
  - equally effective for less money
  - OR more effective for \$500
  - OR “Here is a more meaningful optimization criterion: \_\_\_\_\_”
- Continuous optimization principle in action

## EXAMPLE 2: MY PLAYBOOK: DRUG ABUSE PREVENTION FOR NCAA ATHLETES

Objective: Engineer an intervention made out of components each demonstrating an effect of  $d \geq .3$

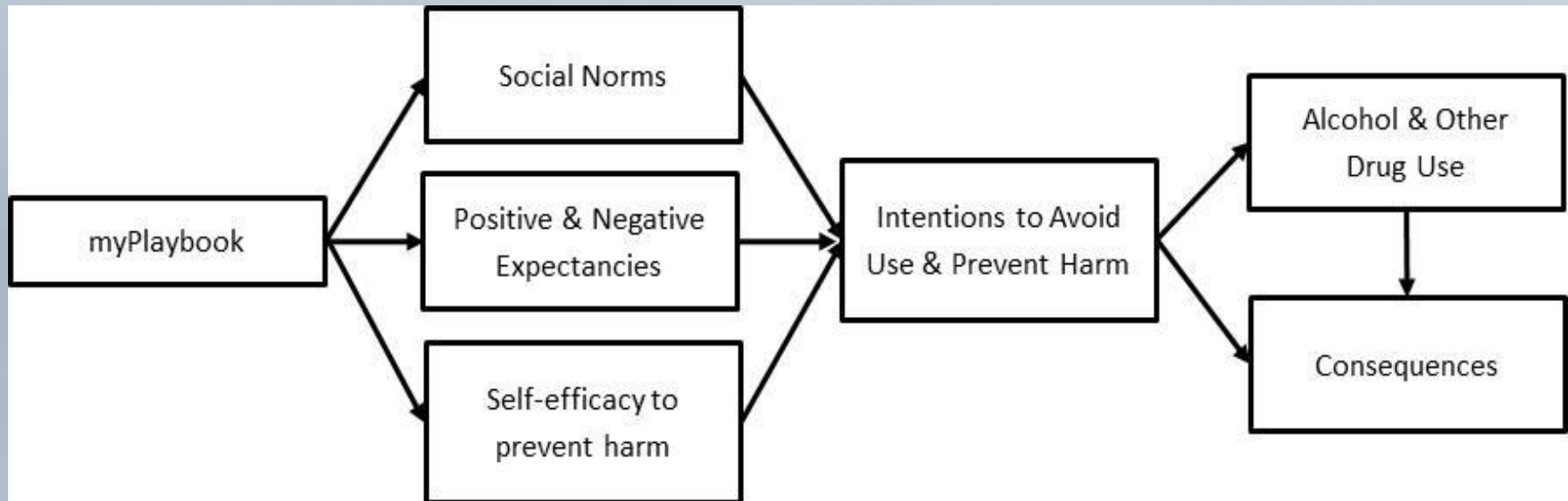
Investigative team: David Wyrick (UNCG), Melodie Fearnow-Kenney (Prevention Strategies, LLC), Kelly Rullison (UNCG), Linda Collins (PSU)

Funded by the National Institute on Drug Abuse

# SOME INTERESTING FEATURES OF EXAMPLE 2

- Intervention is delivered online
- Experimental design is a cluster-randomized fractional factorial
- Takes an iterative approach

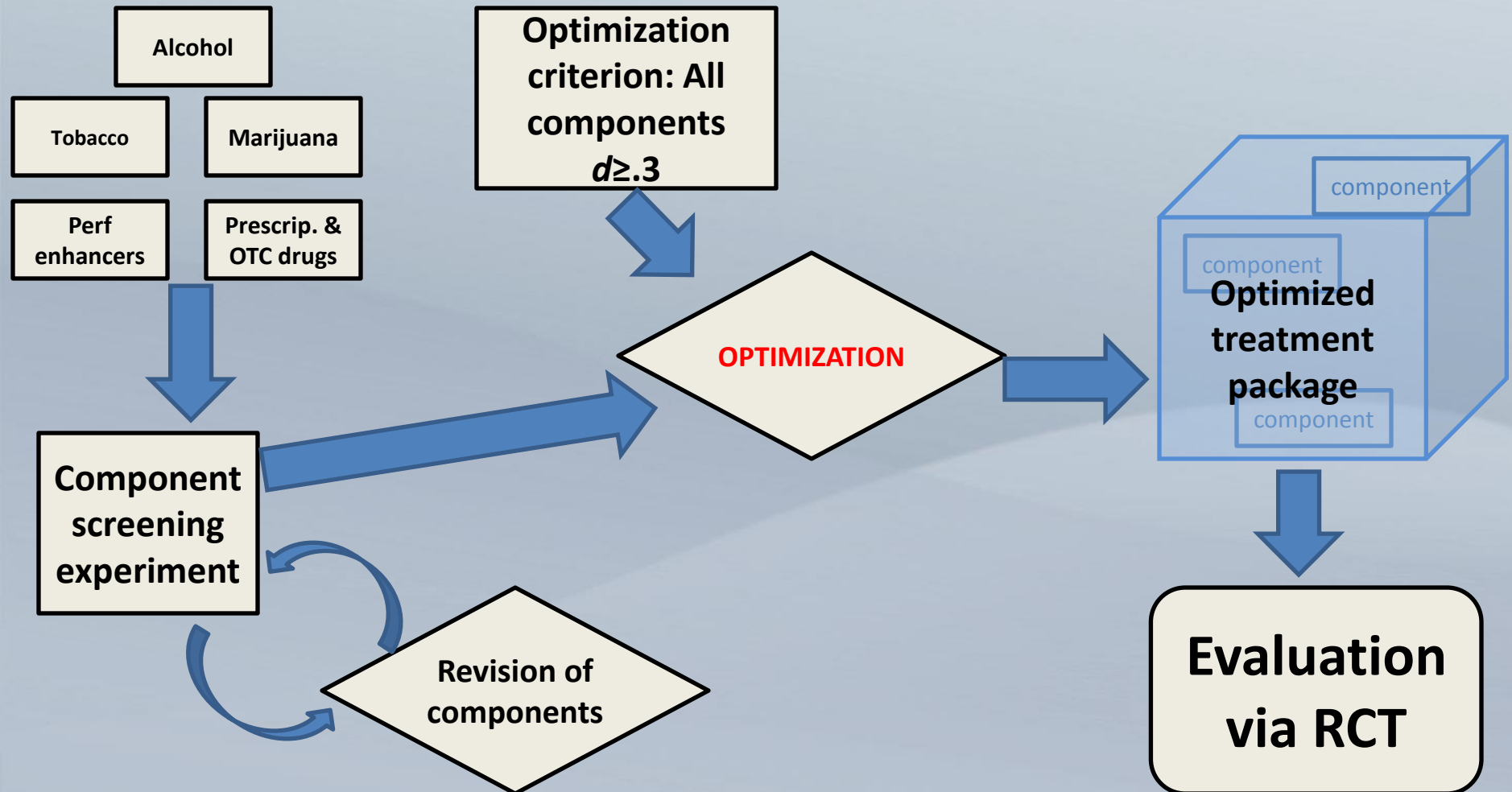
# THEORETICAL MODEL FOR MY PLAYBOOK STUDY



# INTERVENTION COMPONENTS

- Five components, each corresponding to a substance:
  - I. Alcohol
  - II. Tobacco
  - III. Marijuana
  - IV. Performance enhancers
  - V. Prescription and OTC drugs
  
- Each component aimed at
  - social norms
  - positive and negative expectancies
  - intentions to avoid use and prevent harm

# MOST AS PLANNED TO BE IMPLEMENTED IN MY PLAYBOOK



# OVERVIEW OF ITERATIVE APPROACH

- The plan:
  - Experiment to evaluate individual effects of intervention components
  - Any components not showing main effect of  $d \geq .3$  revised
  - Experiment repeated
  - Any components not showing main effect of  $d \geq .3$  revised
  - Final experiment
  - Intervention assembled



# CLUSTER RANDOMIZATION IS NECESSARY

- Elite athletes in a school tend to know each other
- Concerned about contamination

# DETOUR

# DETOUR

- We want to examine 5 components, so would like to use a factorial design
- BUT we have to use cluster randomization.
- This applies to lots of other situations
  - Any school-based intervention research
  - Couples, family, or small groups intervention research
- Is a cluster-randomized factorial experiment possible? We see two issues.

# FACTORIAL EXPERIMENTS AND CLUSTER RANDOMIZATION: TWO ISSUES

- #1: Number of clusters available for randomization
  - Suppose you want to conduct a  $2^5$  factorial
  - You might have a large enough  $N$  BUT fewer than 32 units!
- A fractional factorial design might be the only option

# FACTORIAL EXPERIMENTS AND CLUSTER RANDOMIZATION: TWO ISSUES

- #2: Power
  - Will standard methods of estimating power work?
    - Cluster randomization approaches come out of educational research methods
    - Fractional factorial designs come out of engineering
      - Cluster randomization not used much in engineering
  - Is it feasible to have enough power?

# POWER IN CLUSTER RANDOMIZED FACTORIAL EXPERIMENTS

- Design effect

$$D = 1 + (n - 1)\rho_X\rho_{Y|X}$$

- $n$  = number in each cluster
  - $\rho_X$  = intraclass correlation of  $X$
  - $\rho_{Y|X}$  = intraclass correlation of  $Y$  given  $X$
- If you need  $N$  individually sampled subjects, you will need  $DN$  clustered subjects

# POWER IN CLUSTER RANDOMIZED FACTORIAL EXPERIMENTS

- It's complicated
- In addition to noncentrality,  $\alpha$ , and total  $N$ , necessary to consider
  - $\rho_{Y|X}$
  - Number of clusters
  - $n$  per cluster

# POWER IN CLUSTER RANDOMIZED FACTORIAL EXPERIMENTS

- We\* investigated power via simulation
  - To check that standard formulae work (they do)
  - To see how feasible factorial experiments would be

\*Dziak, Nahum-Shani, & Collins, 2012



# POWER FOR BETWEEN-CLUSTER ASSIGNMENT FOR “SMALL” ( $d = .2$ ) EFFECTS, 5 FACTORS, $P_{Y|X} = .05$

#Clusters	#Per cluster	Predicted power	Observed power
25*	20	0.62	0.59
	100	0.90	0.89
30*	20	0.73	0.74
	100	0.96	0.96
40	20	0.87	0.89
	100	0.99	0.99
50	20	0.94	0.96
	100	1.00	1.00

\* Fractional factorial required

# POWER FOR BETWEEN-CLUSTER ASSIGNMENT FOR “SMALL” ( $d = .2$ ) EFFECTS, 5 FACTORS, $P_{Y|X} = .15$

#Clusters	#Per cluster	Predicted power	Observed power
25*	20	0.40	0.38
	100	0.52	0.49
30*	20	0.49	0.50
	100	0.64	0.64
40	20	0.64	0.65
	100	0.78	0.79
50	20	0.74	0.77
	100	0.87	0.89

\* Fractional factorial required

# POWER FOR BETWEEN-CLUSTER ASSIGNMENT FOR “SMALL” ( $d = .2$ ) EFFECTS, 5 FACTORS, $P_{Y|X} = .30$

#Clusters	#Per cluster	Predicted power	Observed power
25*	20	0.25	0.24
	100	0.29	0.26
30*	20	0.31	0.31
	100	0.36	0.36
40	20	0.42	0.41
	100	0.48	0.47
50	20	0.51	0.52
	100	0.58	0.59

\* Fractional factorial required

# END DETOUR

# POWER IN MY PLAYBOOK EXPERIMENTS

- We expected  $P_{Y|X} = .05$
- 64 schools (4/condition) provides power  $> .9$  for  $d = .3$ ; overall  $N$  approx. 3500
  - We have enough schools to populate a  $2^5$  design BUT not much of a “cushion” in case schools drop out
- Do not need to power for less than  $d = .3$  because this is the criterion for efficacy

## Experimental Conditions in My Playbook Fractional Factorial Design

		Manipulated Intervention Components				
Condition Number	Introduction	Alcohol	Tobacco	Marijuana	Performance-enhancing substances	Prescription & OTC drugs
1	On	Off	Off	Off	Off	On
2	On	Off	Off	Off	On	Off
3	On	Off	Off	On	Off	Off
4	On	Off	Off	On	On	On
5	On	Off	On	Off	Off	Off
6	On	Off	On	Off	On	On
7	On	Off	On	On	Off	On
8	On	Off	On	On	On	Off
9	On	On	Off	Off	Off	Off
10	On	On	Off	Off	On	On
11	On	On	Off	On	Off	On
12	On	On	Off	On	On	Off
13	On	On	On	Off	Off	On
14	On	On	On	Off	On	Off
15	On	On	On	On	Off	Off
16	On	On	On	On	On	On

# THIS IS A FRACTIONAL FACTORIAL DESIGN

- Design of My Playbook experiments:  $2^{5-1}$  fractional factorial
- Resolution V design
  - Main effects aliased with 4-way interactions

# EXAMPLE 3: TREATMENT OF ADHD IN CHILDREN

Objective: To develop an effective and cost-effective time-varying adaptive intervention to treat children with ADHD

Investigative team: William Pelham (FIU), Susan A. Murphy (Michigan), Inbal Nahum-Shani (Michigan) and others

Funded by National Institute of Mental Health



# TIME-VARYING ADAPTIVE BEHAVIORAL INTERVENTIONS

- Progress of individual is assessed periodically
  - Review interval
- Dose of intervention components may be changed depending on progress, according to
  - Tailoring variables
  - Decision rules
    - Collins, Murphy, & Bierman (2004)

# DEFINITION: INTERVENTION COMPONENT

- *Any aspect of an intervention that can be separated out for study. Consider an alcoholism treatment intervention:*
  - A decision rule in an adaptive intervention
    - e.g.: “If more than one heavy drinking day in a two-week period, **increase clinic visits to three per week**”
  - A tailoring variable in an adaptive intervention
    - e.g.: “If **more than one heavy drinking day in a two-week period**, increase clinic visits to three per week”
  - The review interval in an adaptive intervention
    - e.g.: “If more than one heavy drinking day in **a two-week period**, increase clinic visits to three per week”

# TREATMENT APPROACHES FOR CHILDREN WITH ADHD

- Two approaches to treatment of ADHD
  - Behavior modification (BMOD)
  - Medication
- They can be combined into a time-varying adaptive intervention
- What is the best set of decision rules?

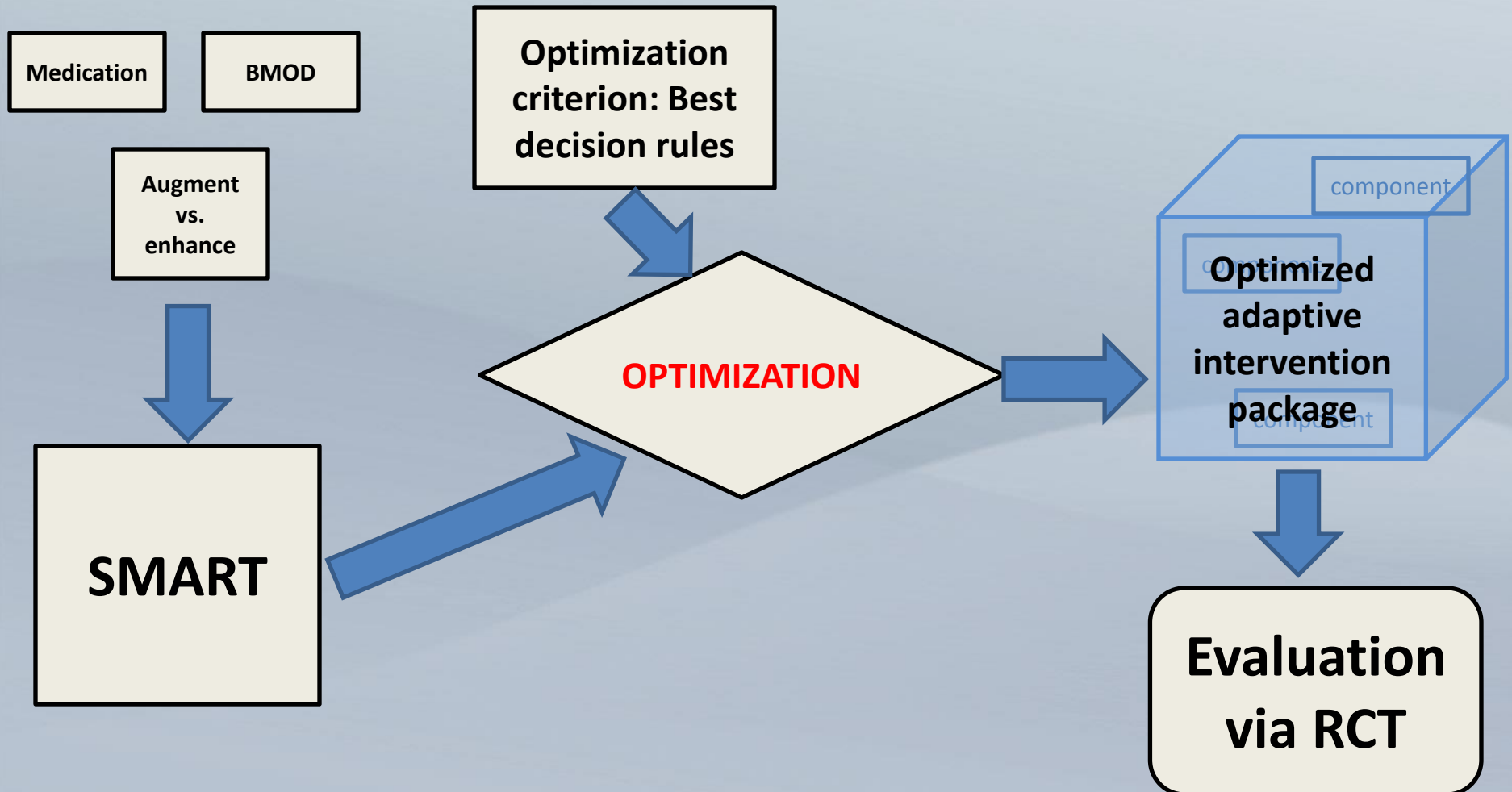
# RESEARCH QUESTIONS THAT MUST BE ADDRESSED FOR OPTIMIZATION

- Is it better to start with BMOD or medication?
  - Note: BMOD much more expensive
- For those who do not respond to initial treatment, is it better to
  - Remain with the same strategy but enhance
  - Augment with the other strategy
- What is the best overall strategy?

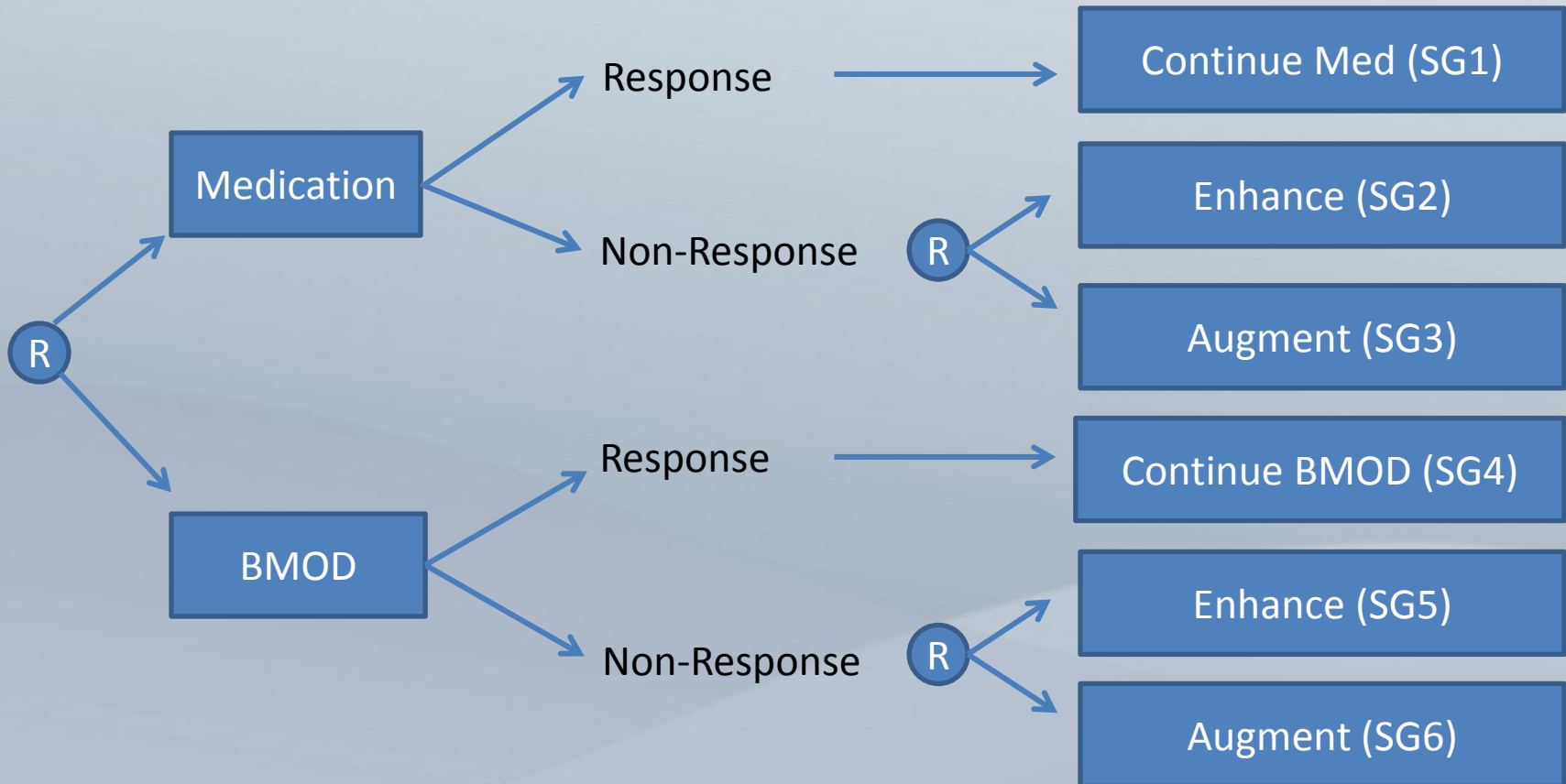
# THE SEQUENTIAL MULTIPLE-ASSIGNMENT RANDOMIZED TRIAL (SMART)

- A special case of the factorial experiment
- Extremely useful when selecting tailoring variables and decision rules
- Pioneered by Susan Murphy and collaborators
  - e.g., Nahum-Shani, I., Qian, M., Almirall, D., Pelham, W., Gnagy, B., Fabiano, G., ... Murphy, S. A. (2012)

# MOST AS IMPLEMENTED IN ADHD STUDY



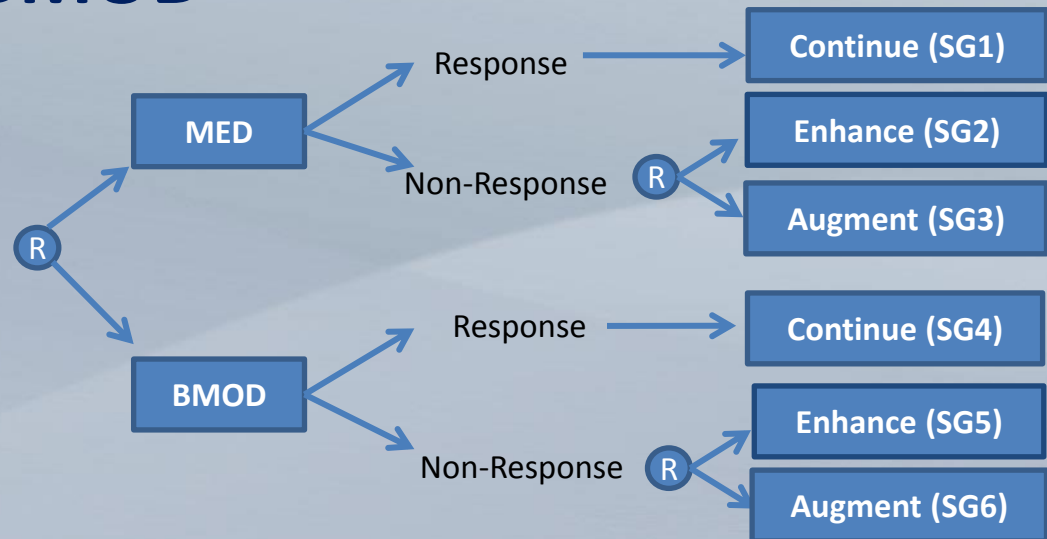
# SMART



# QUESTIONS WE CAN ADDRESS WITH SMART

- Is it better to start with BMOD or MED?
- (SG1+SG2+SG3) vs. (SG4+SG5+SG6)
- Medication vs. BMOD

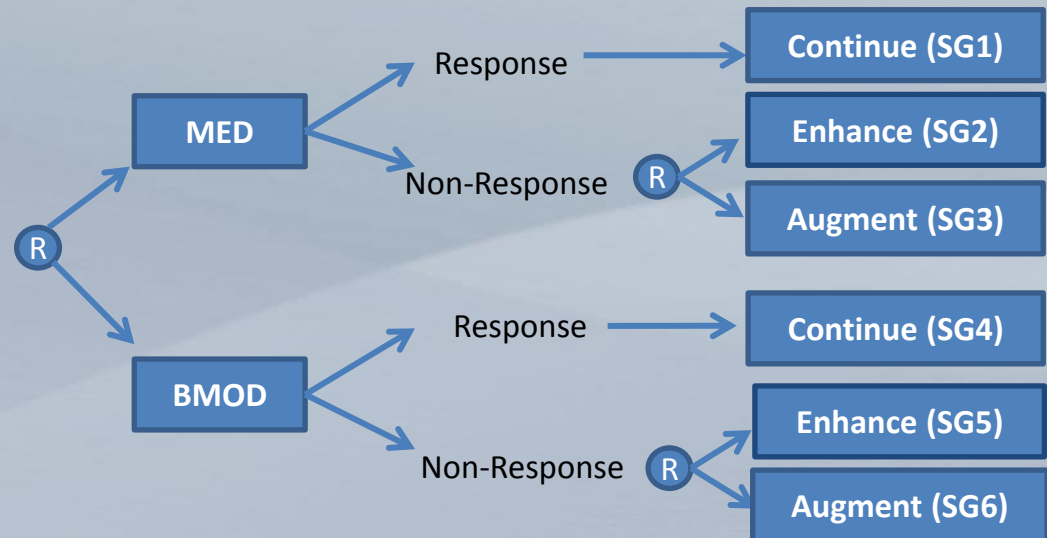
– Averaging over subsequent treatment





# QUESTIONS WE CAN ADDRESS WITH SMART

- Is it better to Enhance or Augment for non-responders?
- (SG2+SG5) vs. (SG3+SG6)
- Enhance vs. Augment



# QUESTIONS WE CAN ADDRESS WITH SMART

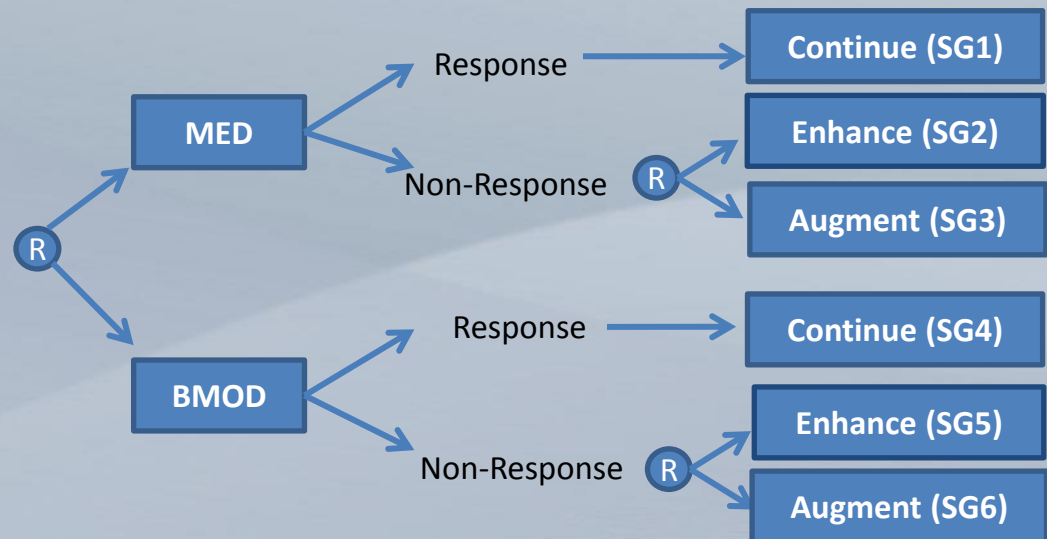
- What is the best overall strategy? There are FOUR embedded here.

Stage 1 = {MED},  
 IF response = {NO}  
 THEN stage 2 = {AUGMENT}  
 ELSE continue stage 1

Stage 1 = {BMOD},  
 IF response = {NO}  
 THEN stage 2 = {ENHANCE}  
 ELSE continue stage 1

Stage 1 = {MED},  
 IF response = {NO}  
 THEN stage 2 = {ENHANCE}  
 ELSE continue stage 1

Stage 1 = {BMOD},  
 IF response = {NO}  
 THEN stage 2 = {AUGMENT}  
 ELSE continue stage 1



# AFTER THE SMART

- Results of analyses will be used to decide on best decision rules, i.e., optimized intervention
- This does not tell us whether the optimized intervention has a statistically significant effect as compared to a control
- Must move to the Evaluation phase for that

# OUTLINE

- Definitions
- What's wrong with business as usual?
- What is MOST? What is optimization?
- OK, how do you do this? Three examples
- Concluding remarks

# SOME OPEN AREAS

- Decision making when
  - Using model other than normal (e.g. logistic, Poisson)
  - Multiple outcome variables
  - Multiple decision makers
  - Multilevel data
- What additional experimental designs can be brought in from engineering?

# SOME OPEN AREAS

- Collection of intensive longitudinal data more common
- Opens up possibilities
  - More frequent review intervals, more data
  - Intervention=dynamical system
  - Could be optimized via control engineering
    - Zafra-Cabeza, Rivera, Collins, et al., 2011
- Ecological momentary interventions
  - Extremely nimble time-varying adaptive interventions

# MOST BORROWS FROM ENGINEERING:

- Idea that behavioral interventions can be engineered
- Concept of optimization
  - Systematic, principled approach to optimization
- Conducting work in phases, each informed by prior phase
- Emphasis on efficiency, including building an efficient product
- Emphasis on resource management
- Use of efficient experimental designs

# A FEW PUBLICATIONS

- Chakraborty, B., Collins, L.M., Strecher, V., and Murphy, S.A. (2009). Developing multicomponent interventions using fractional factorial designs. *Statistics in Medicine*, 28, 2687-2708.
- Collins, L.M., Baker, T.B., Mermelstein, R.J., Piper, M.E., Jorenby, D.E., Smith, S.S., Schlam, T.R., Cook, J.W., & Fiore, M.C. (2011). The Multiphase Optimization Strategy for engineering effective tobacco use interventions. *Annals of Behavioral Medicine*, 41, 208-226.
- Collins, L.M., Dziak, J.R., & Li, R. (2009). Design of experiments with multiple independent variables: A resource management perspective on complete and reduced factorial designs. *Psychological Methods*, 14, 202-224.
- Collins, L.M., Murphy, S.A., & Bierman, K. (2004). A conceptual framework for adaptive preventive interventions. *Prevention Science*, 3, 185-196.
- Dziak, J.D., Nahum-Shani, I., & Collins, L.M. (2012). Multilevel factorial experiments for developing behavioral interventions. *Psychological Methods*, 17(2), 153-175.
- Nahum-Shani et al., (2012). Experimental design and primary data analysis for developing adaptive interventions. *Psychological Methods*, 17(4), 457-477.
- Zafra-Cabeza, A., Rivera, D.E., Collins, L.M., Ridao, M.A., & Camacho, E.F. (2011). A risk-based model predictive control approach to adaptive interventions in behavioral health. *IEEE Transactions on Control Systems Technology*, 19, 891-901.



# FOR MORE INFORMATION:

<http://methodology.psu.edu/ra/most>

This web site has

- suggested reading
- FAQ
- Advice for people writing grant proposals involving MOST

[LMCOLLINS@PSU.EDU](mailto:LMCOLLINS@PSU.EDU)