A Handbook for Designing, Implementing, and Evaluating Successful Electric Utility Pilots

Peter Cappers
C. Anna Spurlock

September 18, 2020

What do we even mean by “Pilot”?

Pilot

An activity undertaken as an **experiment** before introducing something more widely

What do we even mean by “Pilot”?

Pilot

An activity undertaken as a test before introducing something more widely

What do we even mean by “Pilot”?

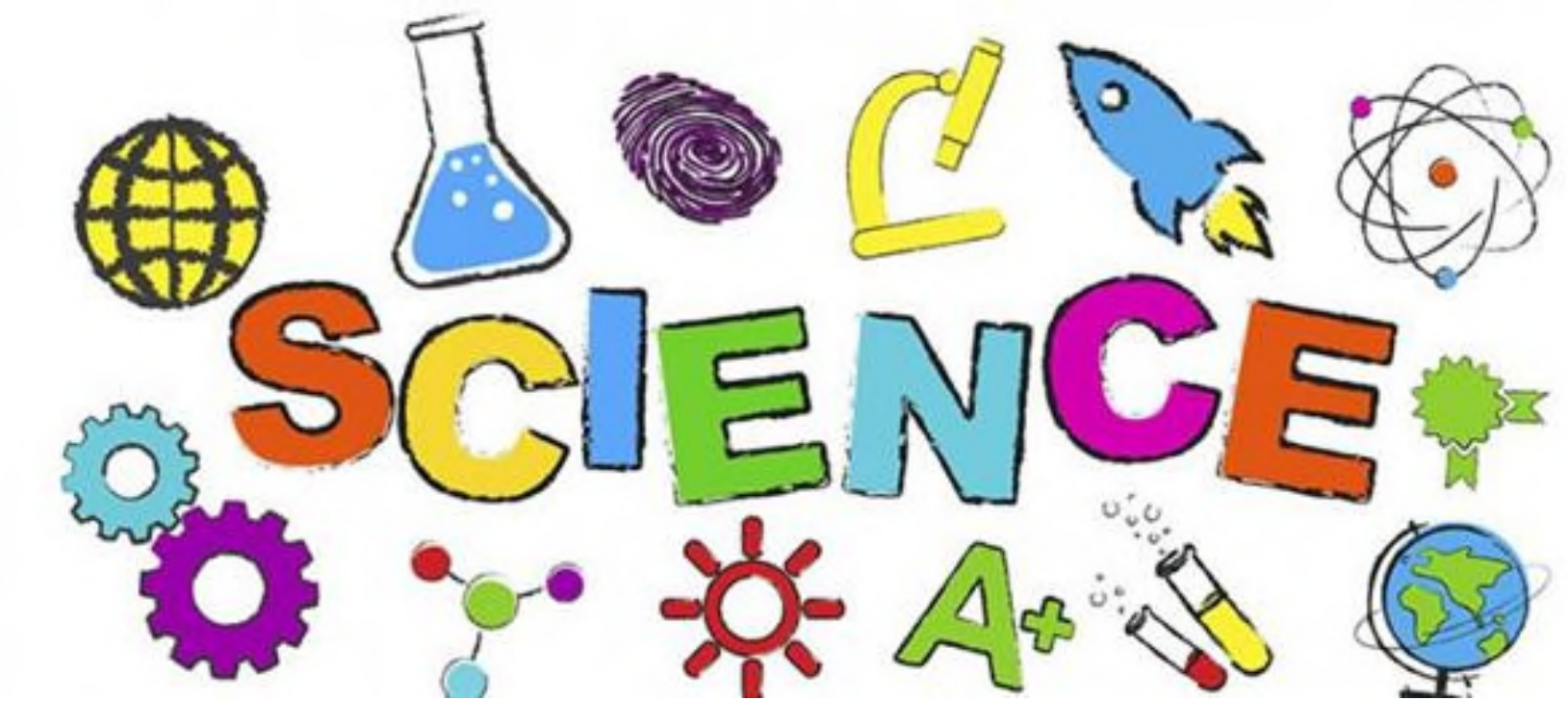
Pilot

An activity undertaken as an **experiment** before introducing something more widely

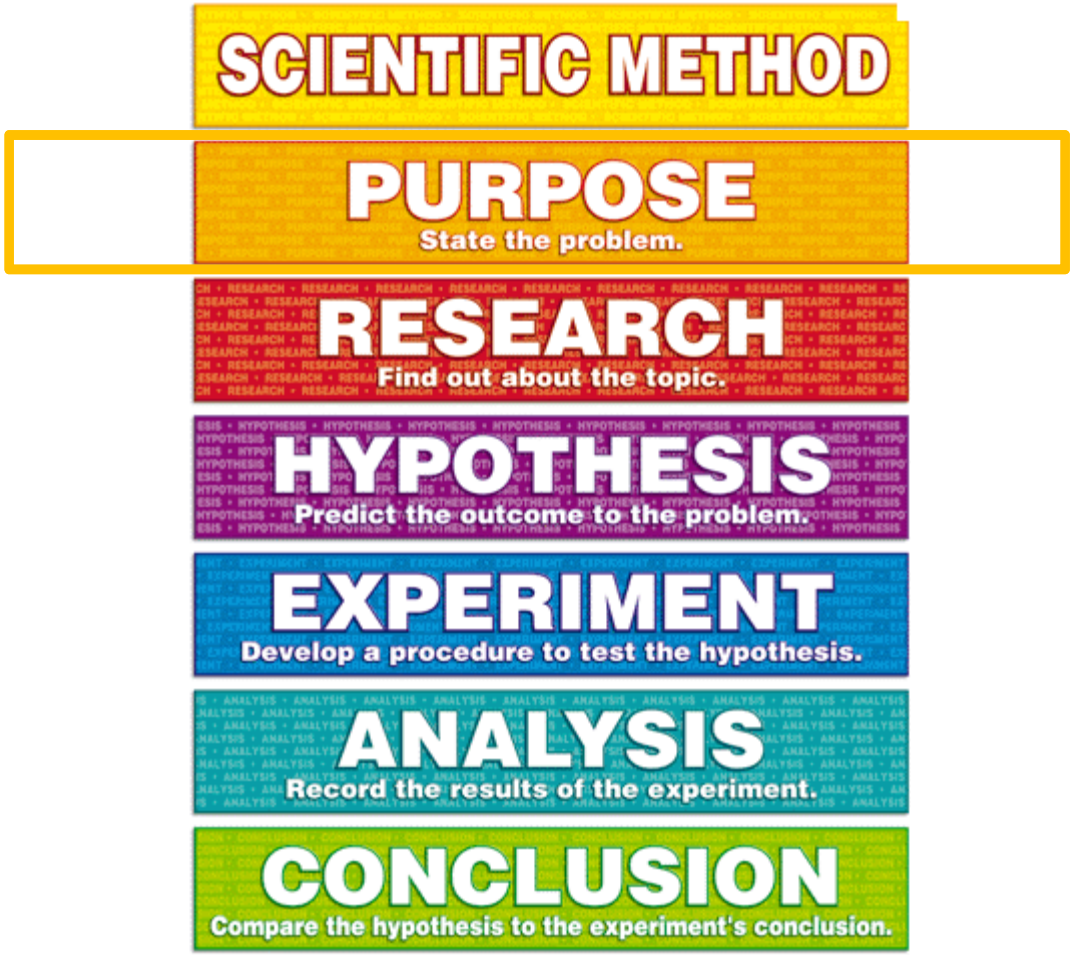
An activity undertaken as a test before introducing something more widely

Process for Pursuing a Pilot & Ensuring Success

Remember 7th Grade Science Class?



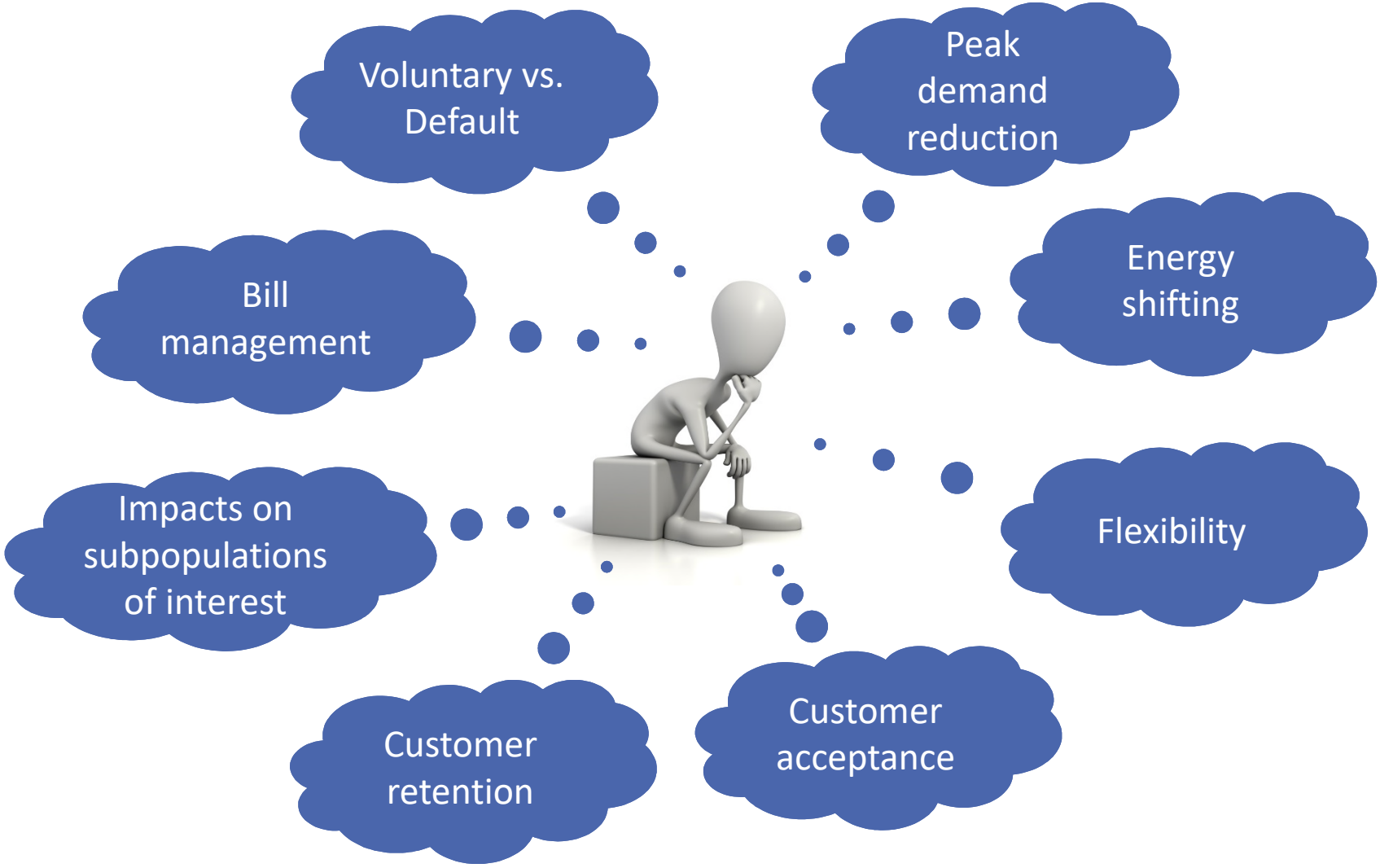
This is the Basic Formula for a Successful Pilot!


SCIENTIFIC METHOD
PURPOSE
State the problem.
RESEARCH
Find out about the topic.
HYPOTHESIS
Predict the outcome to the problem.
EXPERIMENT
Develop a procedure to test the hypothesis.
ANALYSIS
Record the results of the experiment.
CONCLUSION
Compare the hypothesis to the experiment's conclusion.

Process for Deciding if a Pilot is Necessary

What do you want to learn about?

What Do You Want to Learn About?



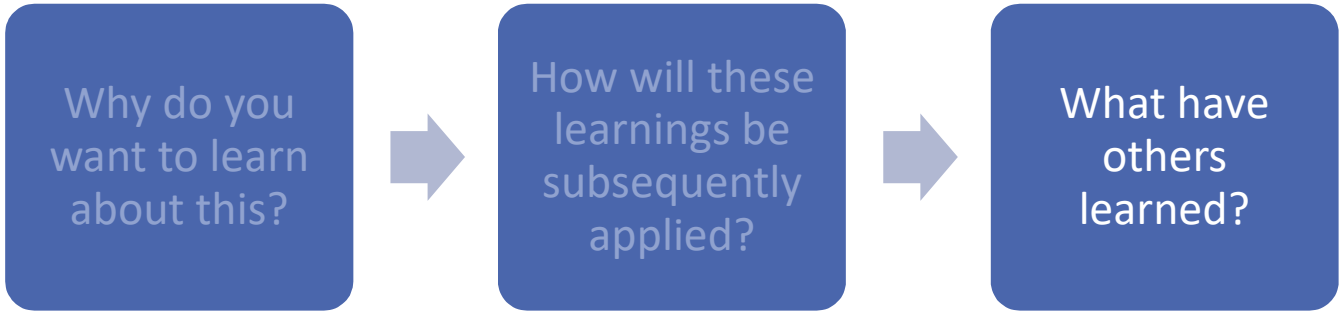
Process for Deciding if a Pilot is Necessary

What do you want to learn about?

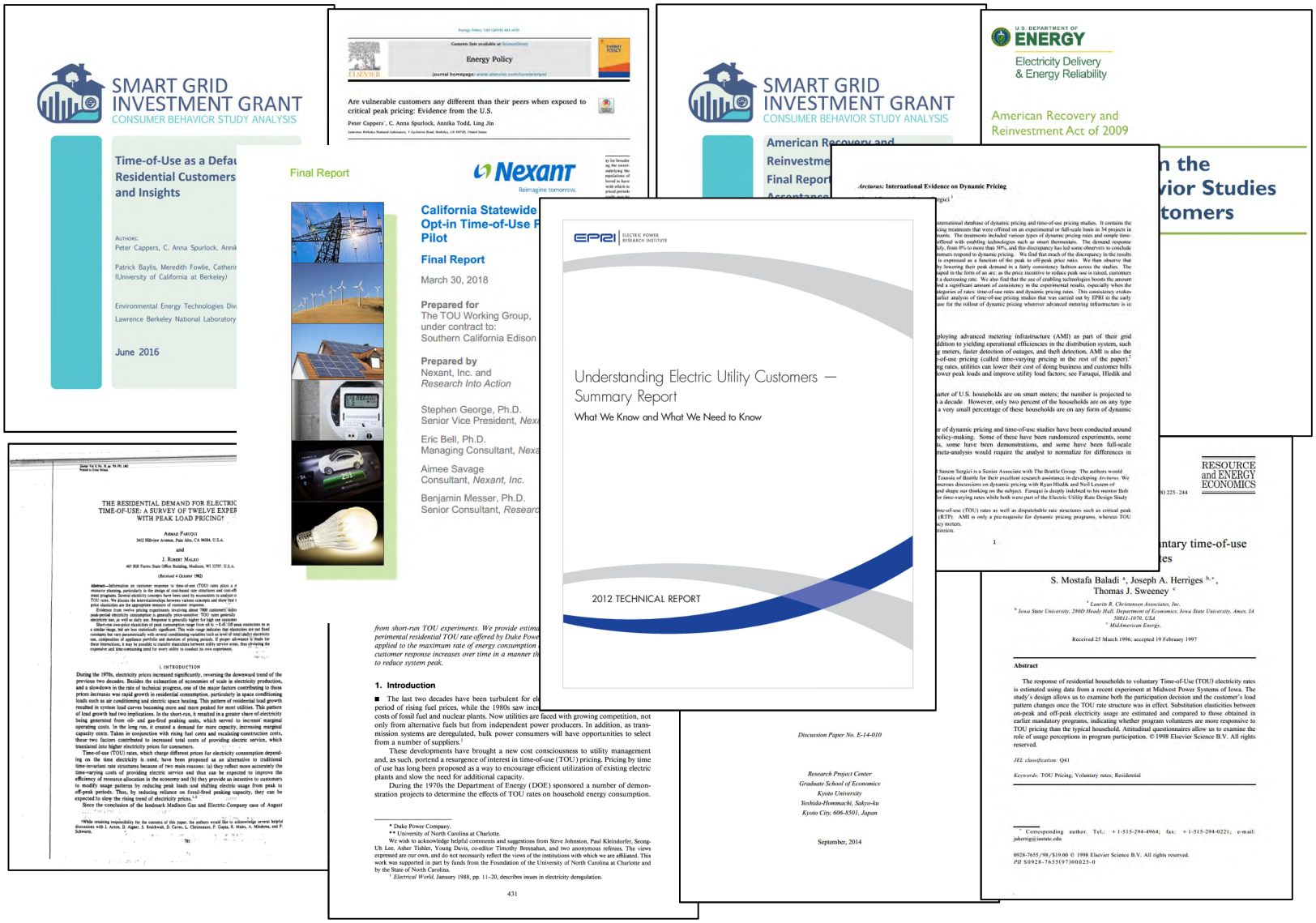
How will these learnings be subsequently applied?


SCIENTIFIC METHOD
PURPOSE
State the problem.
RESEARCH
Find out about the topic.
HYPOTHESIS
Predict the outcome to the problem.
EXPERIMENT
Develop a procedure to test the hypothesis.
ANALYSIS
Record the results of the experiment.
CONCLUSION
Compare the hypothesis to the experiment's conclusion.

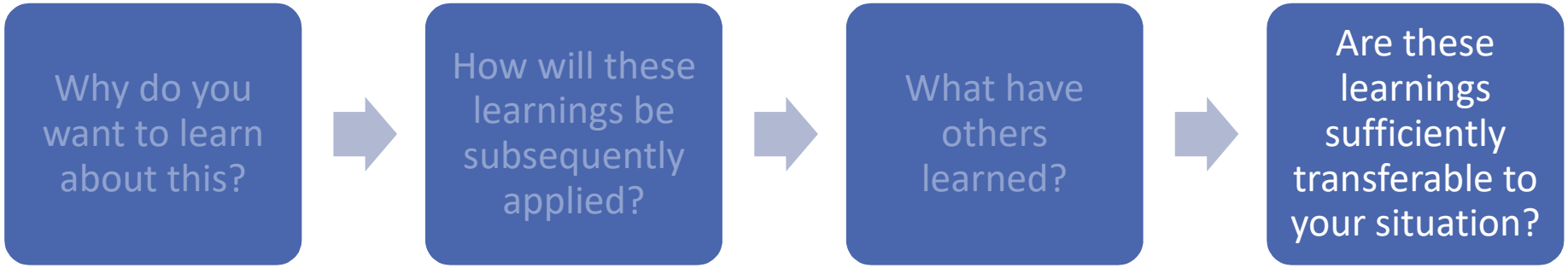
Process for Deciding if a Pilot is Necessary



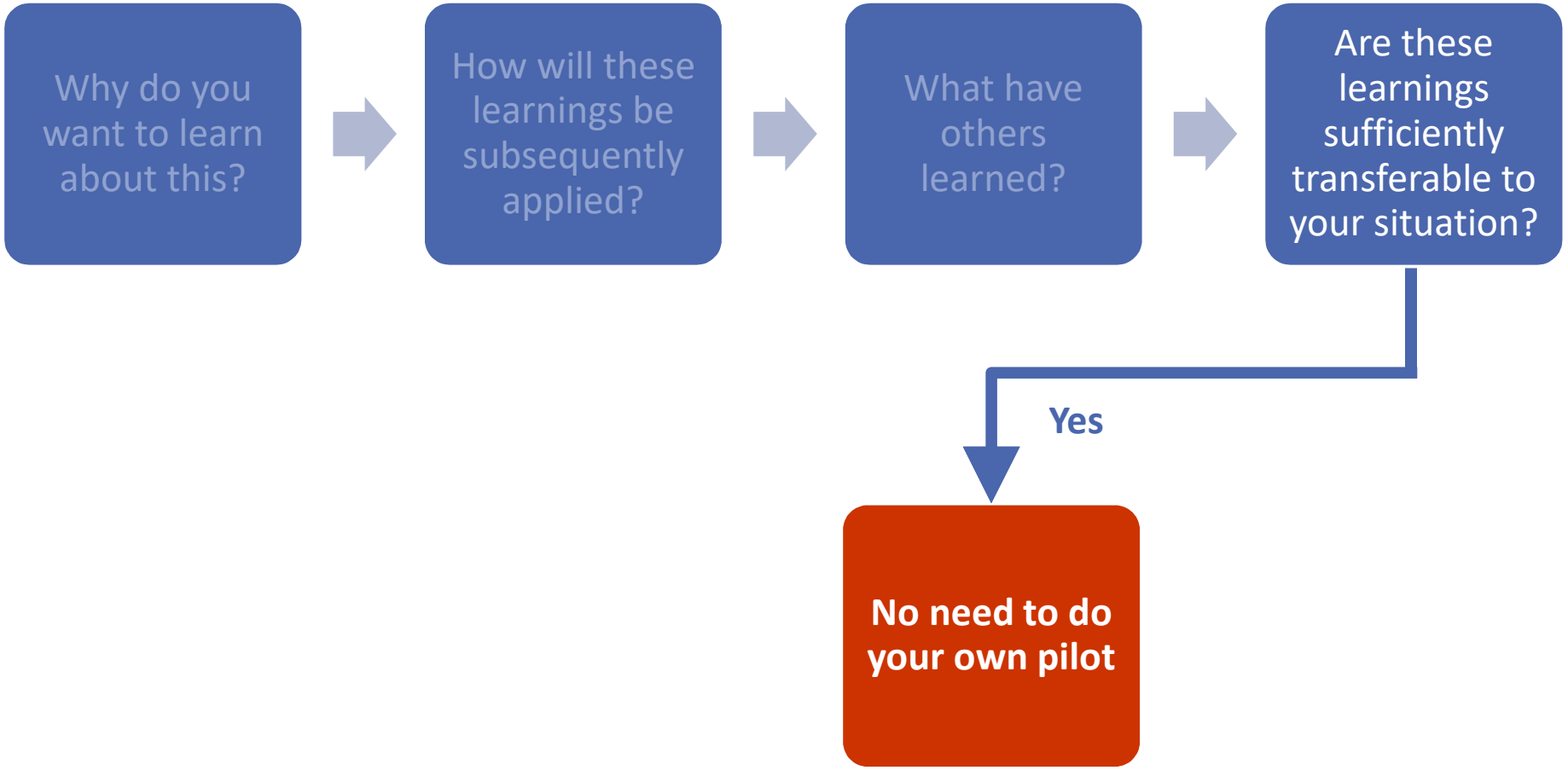
Literature Review



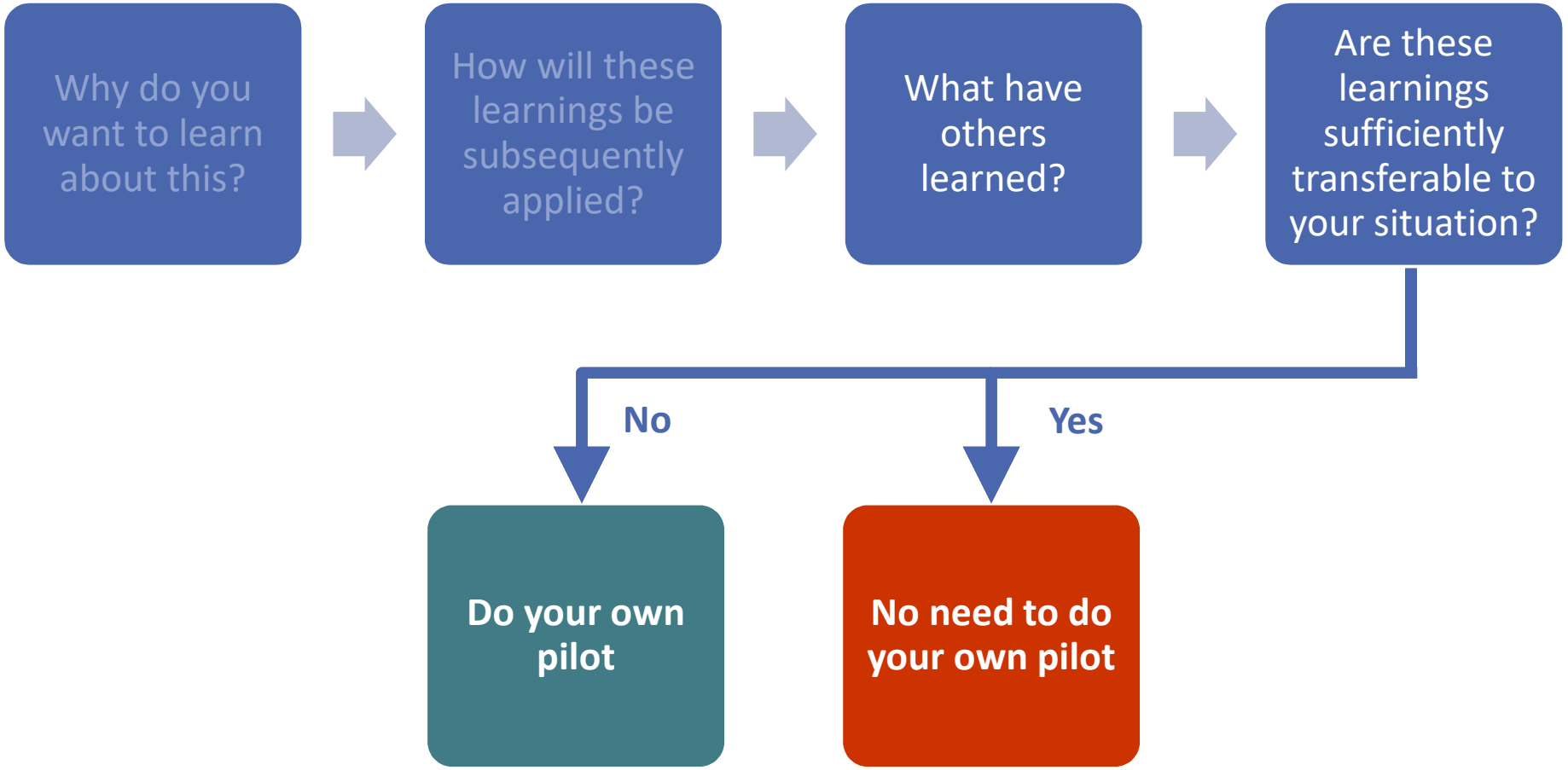
Process for Deciding if a Pilot is Necessary



Process for Deciding if a Pilot is Necessary



Process for Deciding if a Pilot is Necessary

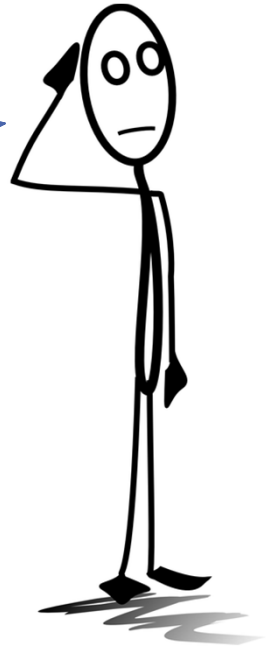



SCIENTIFIC METHOD
PURPOSE
State the problem.
RESEARCH
Find out about the topic.
HYPOTHESIS
Predict the outcome to the problem.
EXPERIMENT
Develop a procedure to test the hypothesis.
ANALYSIS
Record the results of the experiment.
CONCLUSION
Compare the hypothesis to the experiment's conclusion.

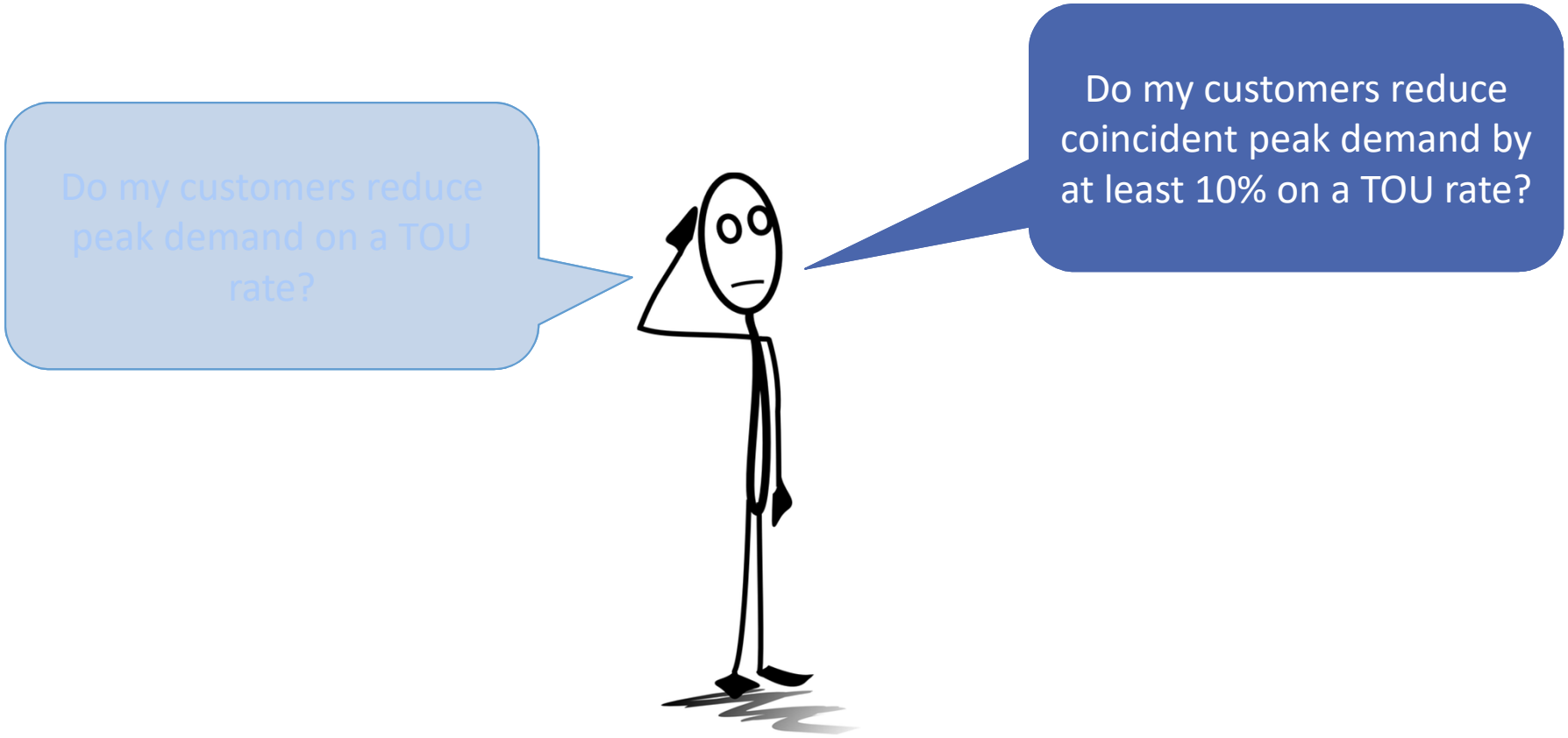
Brainstorm Research Questions

Be Specific

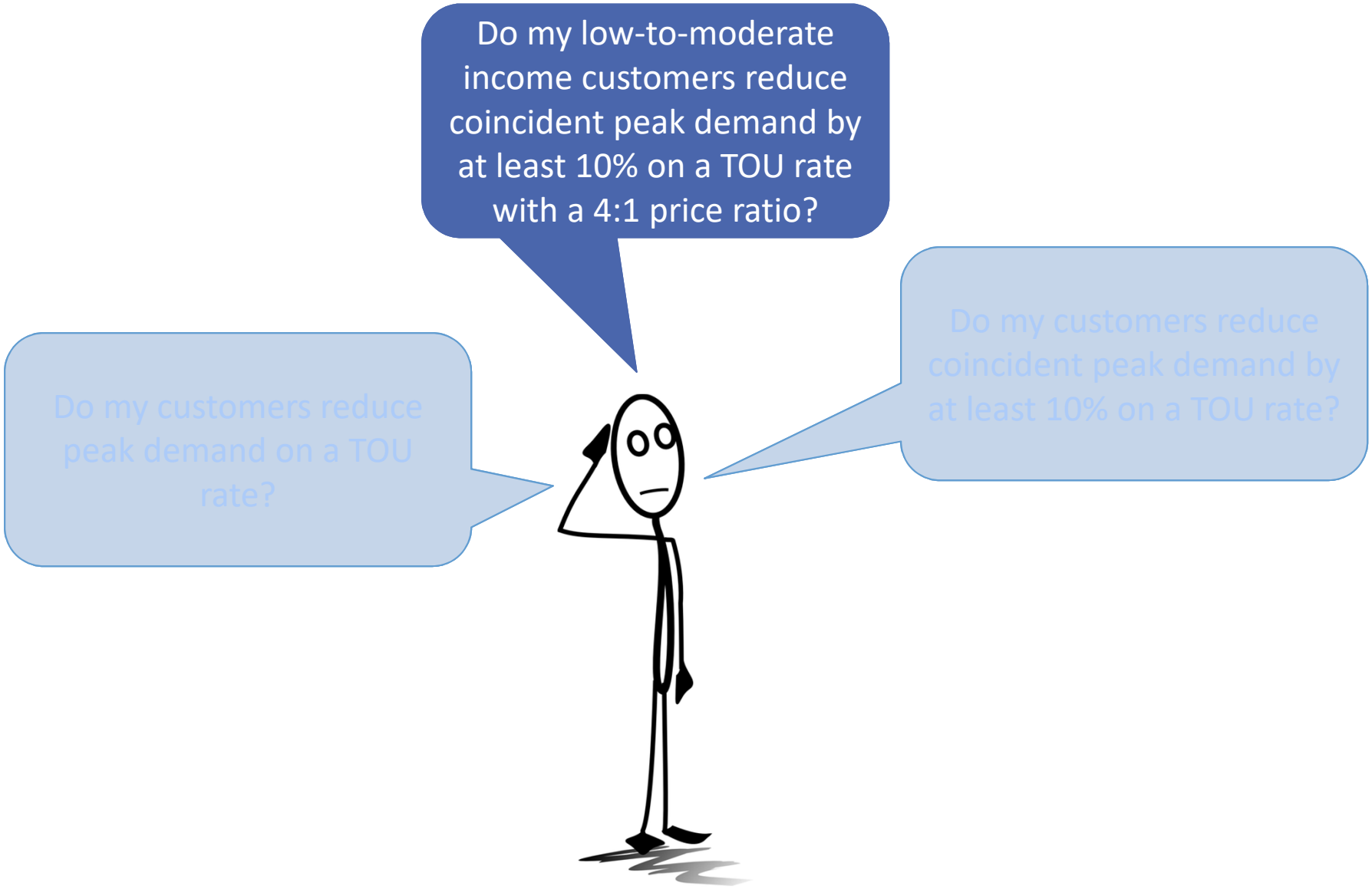
Do my customers reduce peak demand on a TOU rate?



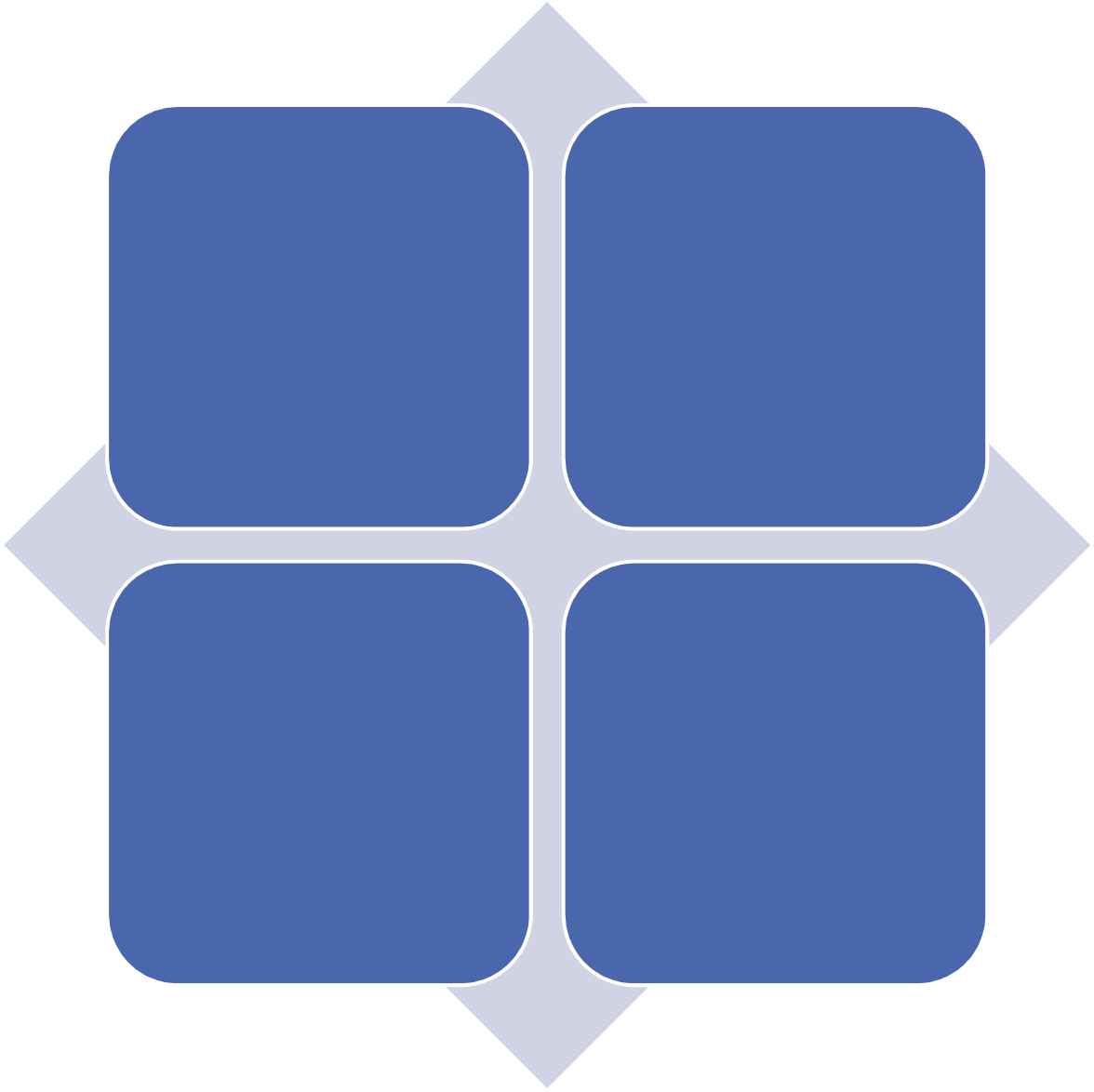
Be Specific


Do my customers reduce peak demand on a TOU rate?
Do my customers reduce coincident peak demand by at least 10% on a TOU rate?

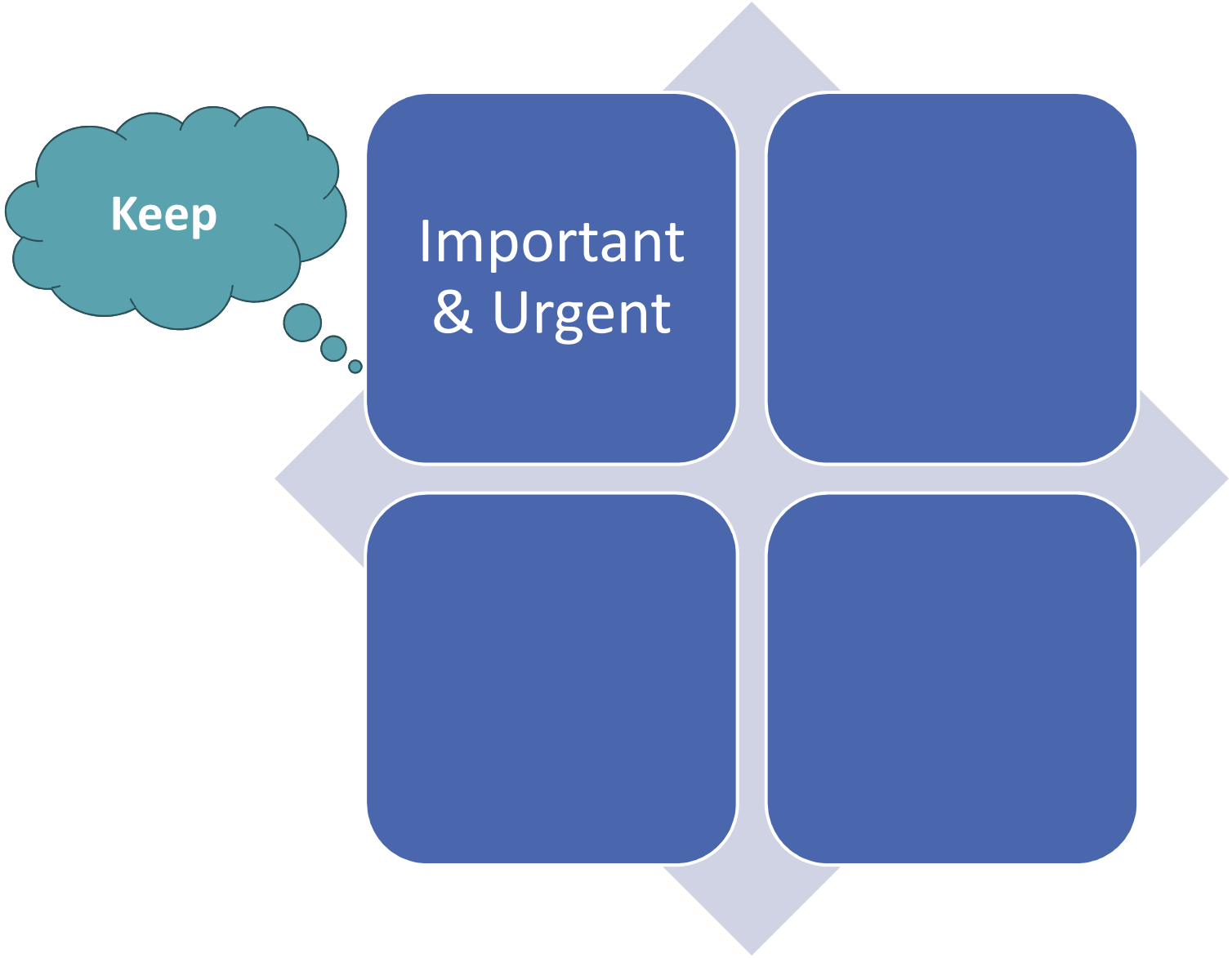
Be Specific


Do my low-to-moderate income customers reduce coincident peak demand by at least 10% on a TOU rate with a 4:1 price ratio?
Do my customers reduce peak demand on a TOU rate?
Do my customers reduce coincident peak demand by at least 10% on a TOU rate?

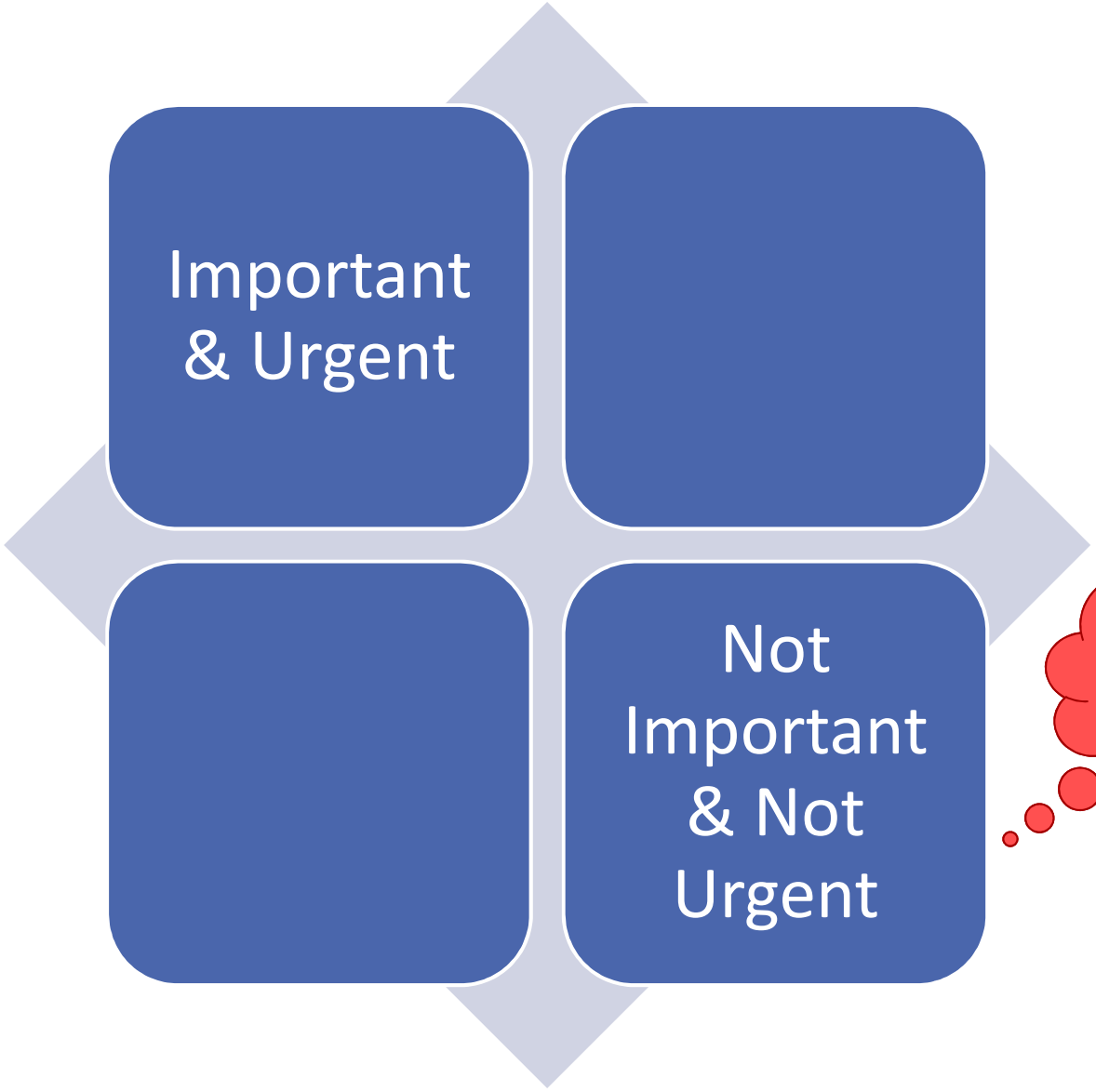
Prioritize: Assess Importance & Urgency



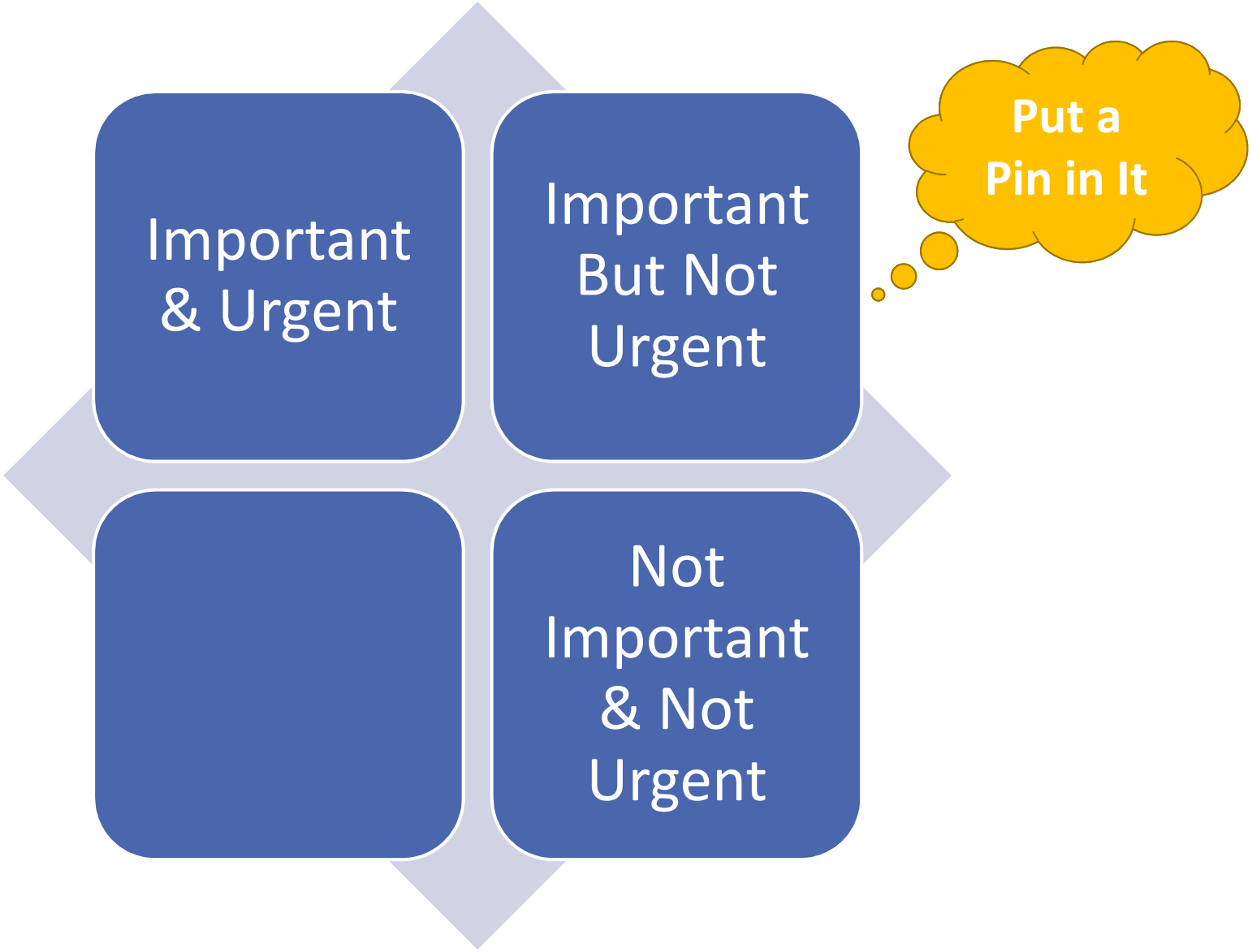
Prioritize: Assess Importance & Urgency



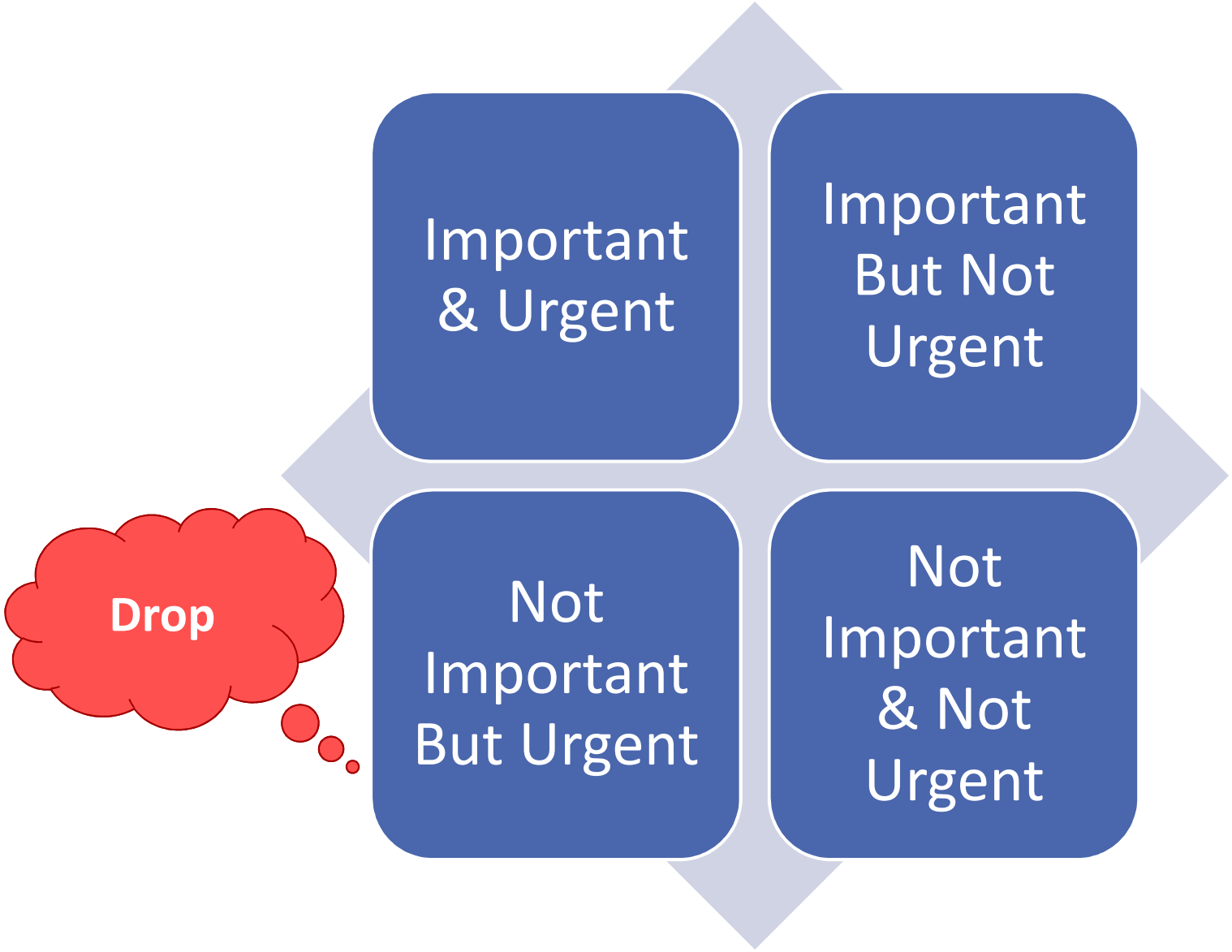
Prioritize: Assess Importance & Urgency



Prioritize: Assess Importance & Urgency



Prioritize: Assess Importance & Urgency



State as a Testable Hypothesis

Research Questions to be Answered	Hypothesis
How does an opt-in TOU rate without a free enabling technology offer affect participant summer, daily, and event load for residential customers?	1. During the test period, average daily energy use for residential customers on the opt-in TOU rate without a free technology offer is lower for the treatment group than for the control group. 2. During the test period, peak energy use for residential customers on the opt-in TOU rate without a free technology offer is lower for the treatment group than for the control group. 3. On event days, peak demand for residential customers on the opt-in TOU rate without a free technology offer is lower for the treatment group than for the control group.

Source: Jimenez, L. R., Potter, J. M. and George, S. S. (2013) Smart Pricing Options Interim Evaluation. Sacramento Municipal Utility District. Prepared for U.S. Department of Energy,. October 2013.

Cautionary Tale: Lack of Specificity in Hypotheses

"A well designed education and outreach program based on individual and social behavioral leading practices on top of the existing inverted rate could induce customer energy efficiency and demand response behavior"

Cautionary Tale: Lack of Specificity in Hypotheses

"A well designed education and outreach program based on individual and social behavioral leading practices on top of the existing inverted rate could induce customer energy efficiency and demand response behavior"

- Uses conditional language instead of declarative language
- Induces a result relative to what
- Joint hypotheses combined into a single statement


SCIENTIFIC METHOD
PURPOSE
State the problem.
RESEARCH
Find out about the topic.
HYPOTHESIS
Predict the outcome to the problem.
EXPERIMENT
Develop a procedure to test the hypothesis.
ANALYSIS
Record the results of the experiment.
CONCLUSION
Compare the hypothesis to the experiment's conclusion.

Determine Level of Accuracy When Testing Hypotheses…

Power & Precision

… Based on Type of Pilot and How Results Will Be Used

Power & Precision

Low ⟷ High

Pre-Pilot Evaluation

Used to Test:

- Implementation concepts
- Logistics & operational procedures
- Innovations

Pilot Process Evaluation

Used as Basis for Decisions Regarding:

- Program planning
- System sufficiency

Pilot Program Evaluation

Used as Basis for Decisions Regarding:

- Cost effectiveness

… Based on Desired Size and Budget

Power & Precision

Pre-Pilot Evaluation

Used to Test:

- Implementation concepts
- Logistics & operational procedures
- Innovations

Smaller sized pilot with lower total cost

Pilot Process Evaluation

Used as Basis for Decisions Regarding:

- Program planning
- System sufficiency

Moderate sized pilot with moderate total cost

Pilot Program Evaluation

Used as Basis for Decisions Regarding:

- Regulatory approval

Larger sized pilot with higher total cost

Assess acceptable chance for confounding effects

Internal Validity

Low ⟷ High

Opportunity for confounding effects

High ⟷ Low

Pilot results may be **caused by something other than** the included elements

Pilot results are **only caused by** the included elements

Determine extent to which results can be extrapolated

External Validity

Applicability of results to populations not included

Results can not be extrapolated to participants or circumstances that differ from those in the pilot

Results can be extrapolated to participants or circumstances that differ from those in the pilot

Determine the most appropriate design

Experimental

- Randomized Controlled Trial (RCT)
- Randomized Encouragement Design (RED)

Determine the most appropriate design

Experimental	Quasi-experimental
• Randomized Controlled Trial (RCT) • Randomized Encouragement Design (RED)	• Non-equivalent Groups Design • Regression Discontinuity Design

Determine the most appropriate design

Experimental	Quasi-experimental	Non-experimental observational
• Randomized Controlled Trial (RCT) • Randomized Encouragement Design (RED)	• Non-equivalent Groups Design (i.e., matched control) • Regression Discontinuity Design	• Descriptive Design • Correlational Design • Developmental Design

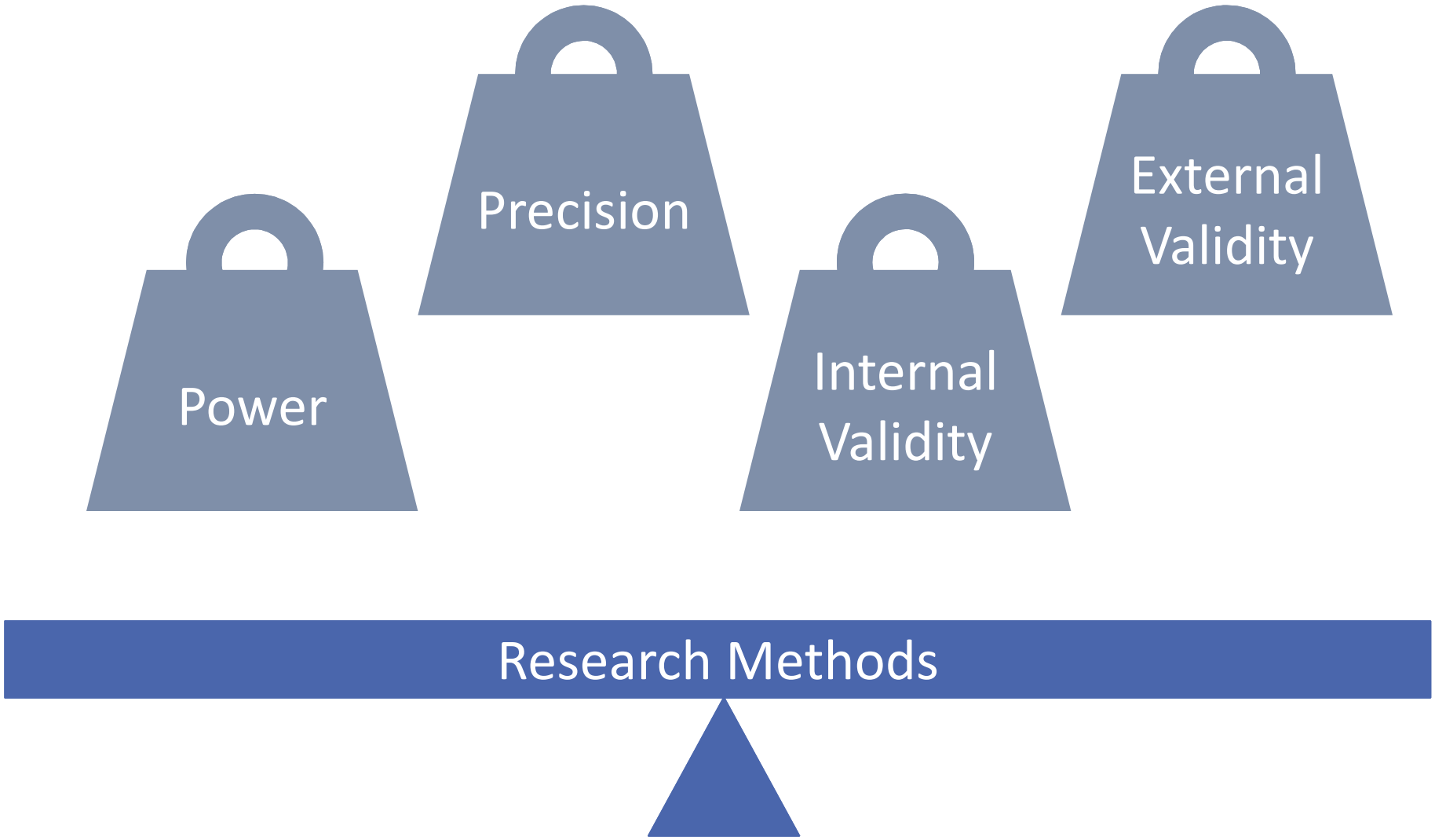
Determine the most appropriate design

Experimental	Quasi-experimental	Non-experimental observational	Non-experimental experiential
• Randomized Controlled Trial (RCT) • Randomized Encouragement Design (RED)	• Non-equivalent Groups Design (i.e., matched control) • Regression Discontinuity Design	• Descriptive Design • Correlational Design • Developmental Design	• Survey Research Design

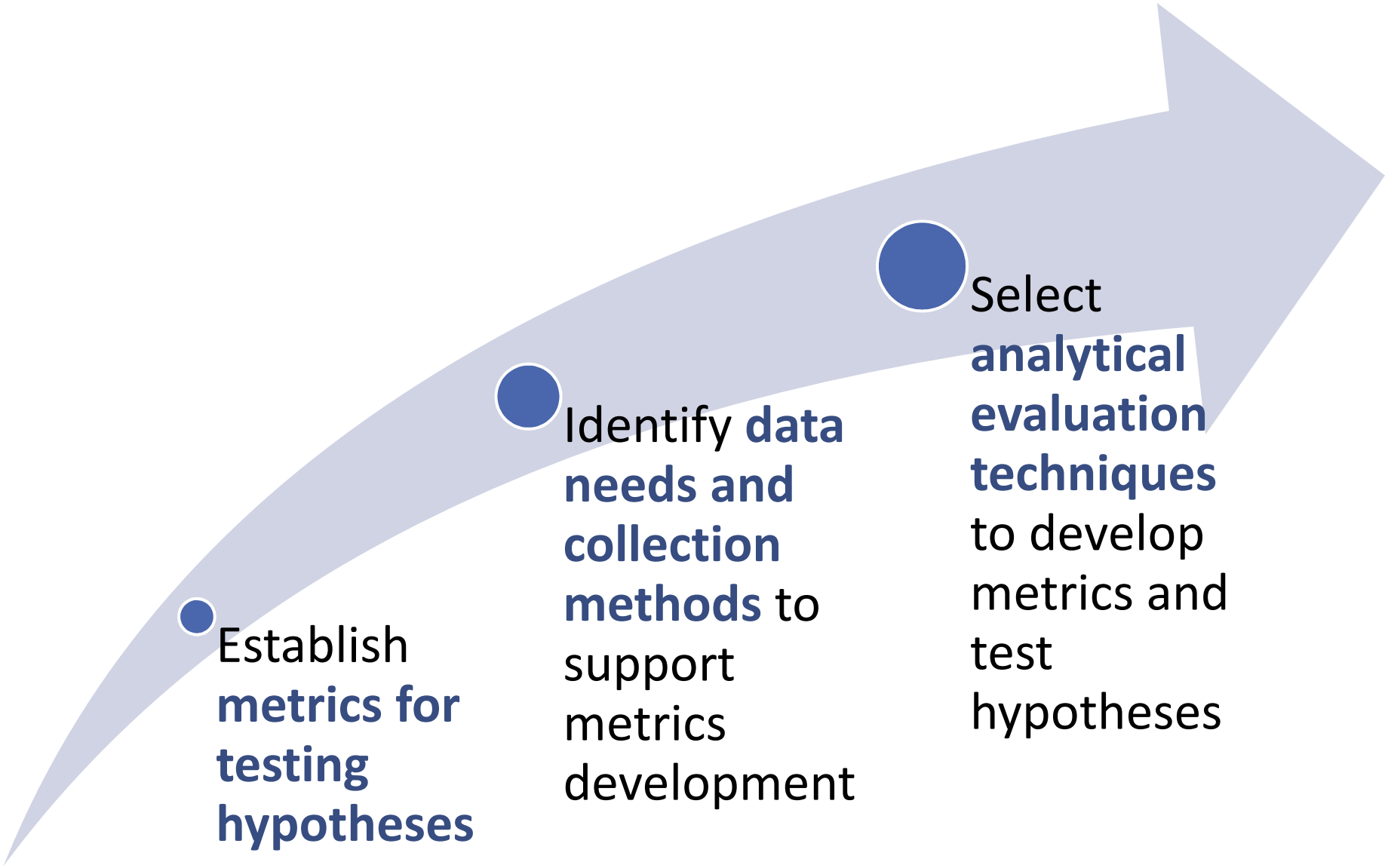
Determine the most appropriate design

Experimental	Quasi-experimental	Non-experimental observational	Non-experimental experiential
• Randomized Controlled Trial (RCT) • Randomized Encouragement Design (RED)	• Non-equivalent Groups Design (i.e., matched control) • Regression Discontinuity Design	• Descriptive Design • Correlational Design • Developmental Design	• Survey Research Design • Case Studies

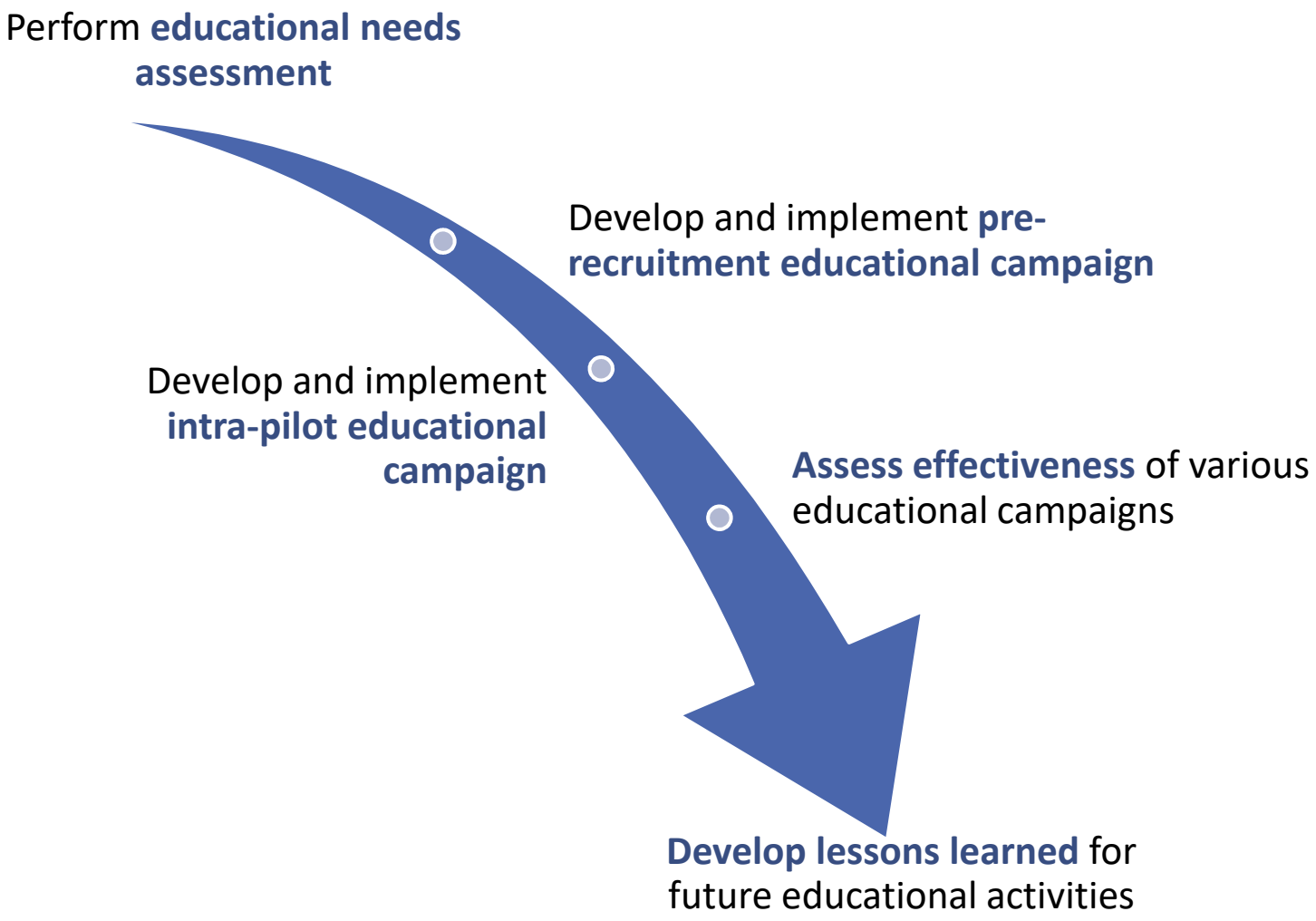
Trade-offs in design elements



Evaluation Plan



Education Plan



Marketing Plan

Perform market research to develop marketing campaign

Test your marketing messages and enrollment process through focus groups, online surveys, etc.

Assess effectiveness of marketing channels, touches, and enrollment process

Implement marketing campaign **first through soft-launch** and then full hard launch

Develop end-of-pilot marketing campaign to support transition of participants

Cautionary Tale: Marketing and Enrollment Gone Bad

- Utility developed its marketing strategy, enrollment process, and communications plan with a large industry consultant for a TOU pricing pilot
- After recruitment campaign ended, utility was only able to recruit about 20% of target pilot sample ➔ Forced to "cancel" pilot
- Ex-post evaluation effort revealed several deficiencies
  - Several areas for improvement in communications material were identified (e.g., initially focused on usage management instead of highlighting that a lower rate would be in effect 94% of hours)
  - A complicated and lengthy 10 step online enrollment process was used with no opportunity to enroll via phone or business reply card
- Revised marketing plan and implemented simpler and more varied enrollment options for an unplanned Phase 2 of the pilot and in so doing lost a year

Outreach Plan

Identify key stakeholders who can support achievement of a successful pilot

Develop outreach strategy with key stakeholder groups

Implement outreach strategy

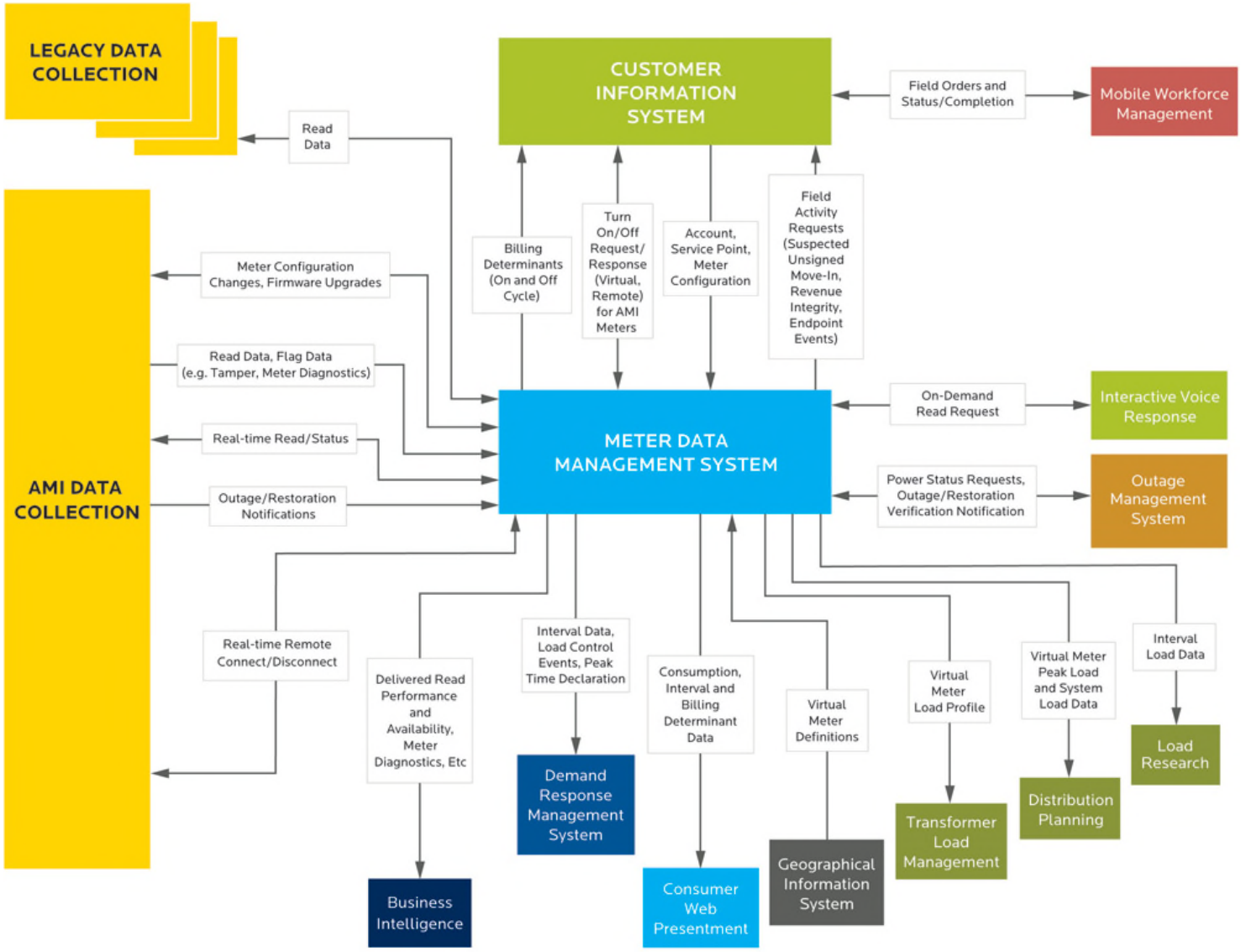
Assess effectiveness of outreach strategy

Develop lessons learned for future outreach activities

External Communication Plan

Frequency of communication	Content of communication	Audience for communication	Format of communication
• Monthly • Semi-annually • Annually • Mid-point and end of pilot	• Enrollment stats • Challenges faced and overcome • Attrition stats • Analytical results	• Regulators • Policymakers • Stakeholders • Ratepayers • Press	• Reports • Presentations • Infographics • Multimedia

Information Technology and Data Management Plan



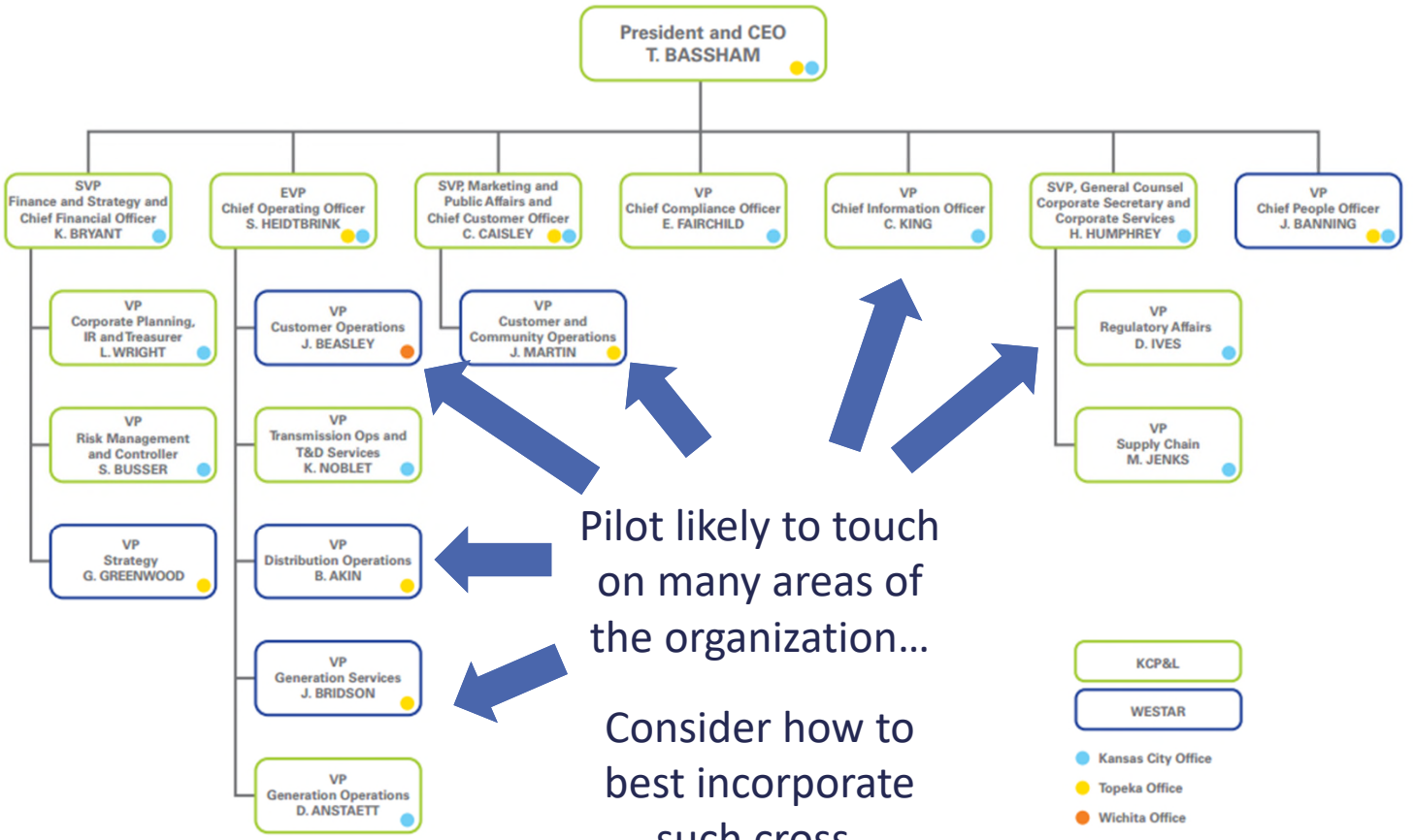
Cautionary Tale: Utility Communication with Customer Tech

- Utility pricing pilot included customer-controlled PCTs that were able to receive price signals and execute cooling strategy according to pre-programmed settings
- Utility had major problems reliably sending prices to PCTs that were correctly received and acted upon
  - Utility initially sent out the wrong price to PCTs
  - Then realized the error and corrected the problem by sending out the right (higher) price
  - PCT did not acknowledge the price increase and failed to execute the appropriate control strategy
  - A very large number of participants faced the higher price but whose PCTs failed to take the appropriate action in response
  - Utility had to provide these participants with ~$200K in total bill credits

Cautionary Tale: Internal Tracking and Data Collection

- Utility designed a pilot with specific offers of rate and technology combinations to specific customers
- During enrollment, no technology eligibility or qualification information was collected from participants by the utility
- After enrollment, the utility made primary and secondary technology assignments to participants
- Utility personnel then went out to install technology at customer premise based on those assignments
- If customer was ineligible to receive the primary technology, then the secondary technology was installed, etc.
- The data file containing Primary and Secondary technology assignments was lost, so the utility was only aware of what customers actually received
- This undermined the initial experimental design which in turn adversely affected the load impact evaluation effort

Internal (Re)Organization Plan



Pilot likely to touch on many areas of the organization…

Consider how to best incorporate such cross-functional staffing requirements

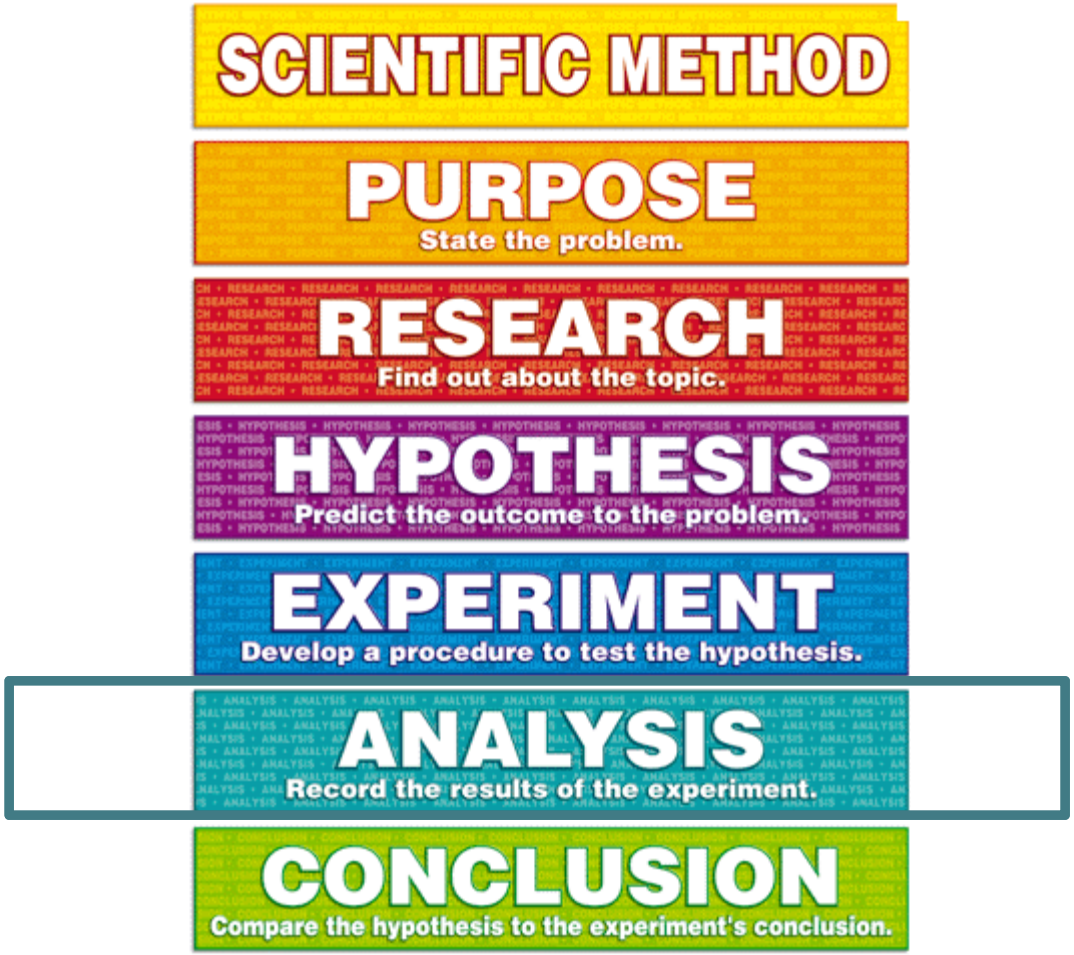
Cautionary Tale: Stress Testing Under Expected Conditions

- Utility pricing pilot included CPP with tariff language requiring day-ahead notification by a certain time for a critical event the following day
- Cross-functional team for dispatching CPP events was created
- Utility tested the event notification process for ~100 people prior to going live
- Early in the first summer of the pilot, utility declared an event for the following day
- System bogged down sending notifications to 1000s of participants most of which arrived after the deadline
  - Cross-functional team was able to quickly identify the problem, determine a solution, and execute it to avoid a substantial number of customer complaints

Contingency Plan (Informal)

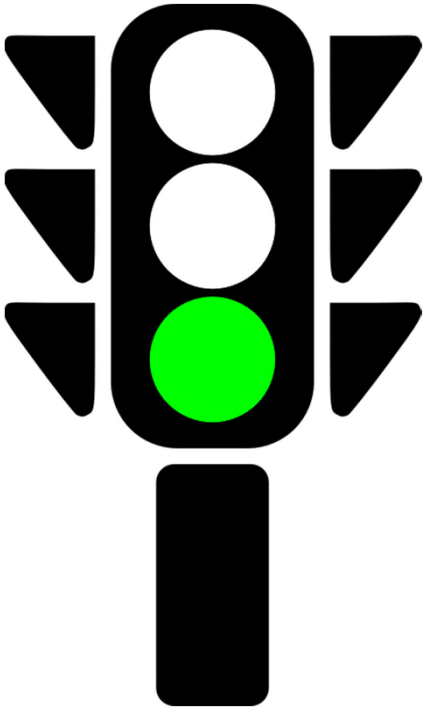
MURPHY'S LAW

WHAT CAN GO WRONG, WILL GO WRONG


SCIENTIFIC METHOD
PURPOSE
State the problem.
RESEARCH
Find out about the topic.
HYPOTHESIS
Predict the outcome to the problem.
EXPERIMENT
Develop a procedure to test the hypothesis.
ANALYSIS
Record the results of the experiment.
CONCLUSION
Compare the hypothesis to the experiment's conclusion.

Execute the Pilot, Collect Data, and Evaluate the Results

LET'S GET STARTED



Cautionary Tale: "I Just Wanted the iPad!"

- Utility implemented a recruitment effort for their pricing pilot
- Recruitment material was altered at the last minute
  - It now read as though customers were being solicited to complete a survey in order to be eligible to win an iPad
  - In so doing they would also be entered to participate in a pricing pilot – this was not obvious to some participants
- Utility confirmed participation in the pricing pilot six (6) months later (December)
  - Many participants said they did not remember signing up for the study, but vaguely remembered a survey and a chance to win an iPad
- 50% attrition rate ➔ Incapable of accurately and credibly estimate load impacts based on experimental design
- "Cancelled" the pilot and redesigned it


SCIENTIFIC METHOD
PURPOSE
State the problem.
RESEARCH
Find out about the topic.
HYPOTHESIS
Predict the outcome to the problem.
EXPERIMENT
Develop a procedure to test the hypothesis.
ANALYSIS
Record the results of the experiment.
CONCLUSION
Compare the hypothesis to the experiment's conclusion.

Apply results to determine next steps

Conclusions and Final Thoughts

- Effective and comprehensive planning should increase the transparency of the various steps in the process
- Regulators and policymakers must balance oversight of the key elements outlined above and the limitations that may place on the speed of and utility interest in innovating through the use of pilots.
- Creating an environment of utility ownership over the pilot's purpose and outcome should improve utility support and likelihood of success.

Questions/Comments

Peter Cappers

(315) 637-0513

pacappers@lbl.gov

C. Anna Spurlock

(510) 495-2072

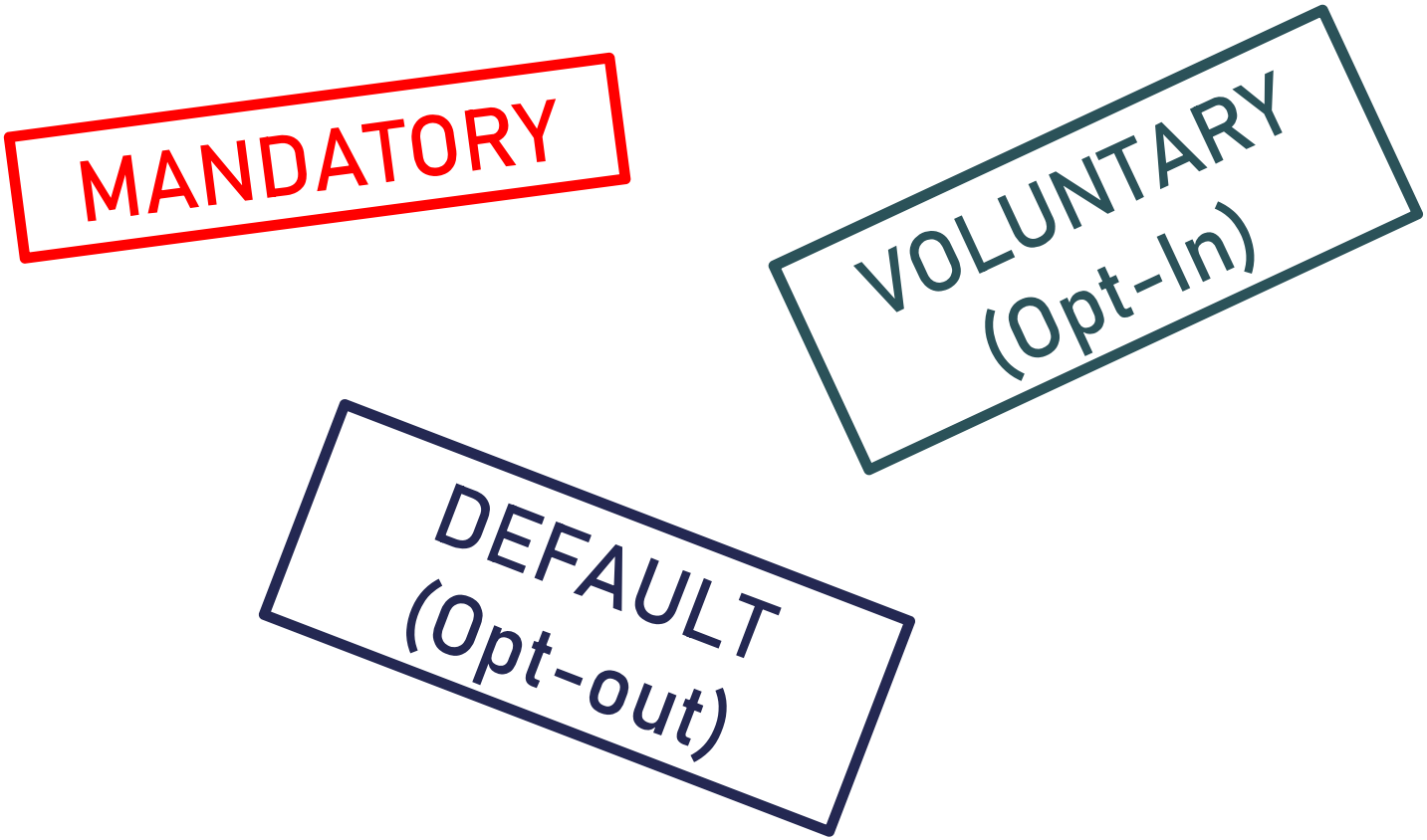
caspurlock@lbl.gov

Additional Reading Material

- Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*: Houghton Mifflin Company.
- Price, P. C. (2012). *Research Methods in Psychology*: Saylor Academy.
- Kirk, R. E. (2009). Experimental Design. In R. E. Millsap & A. Maydeu-Olivares (Eds.), *The SAGE Handbook of Quantitative Methods in Psychology* (pp. 47-72). London, England: SAGE Publishing Ltd.

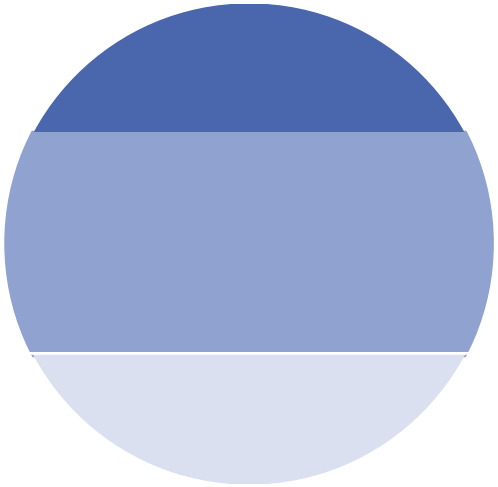
APPENDIX

Enrollment Approaches



Customer Prototypes Under Different Enrollment Approaches

Population of Interest (POI)

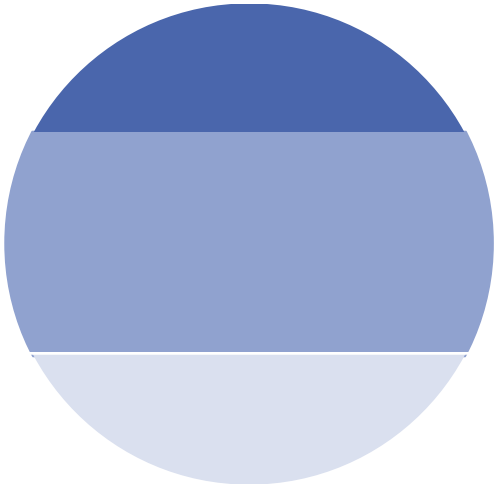


Never Takers

- Join if made mandatory
- Do not join if required to opt-out
- Do not join if required to opt-in

Customer Prototypes Under Different Enrollment Approaches

Population of Interest (POI)

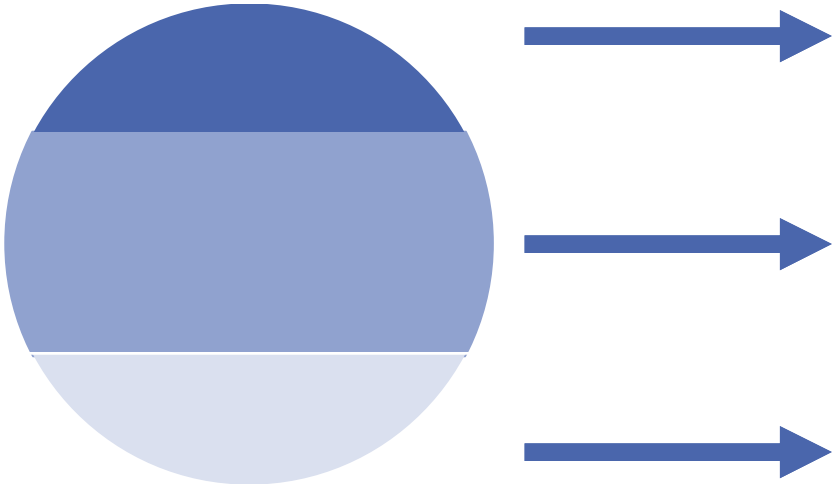


Never Takers
- Join if made mandatory
- Do not join if required to opt-out
- Do not join if required to opt-in

Complacents
- Join if made mandatory
- Join if required to opt-out
- Do not join if required to opt-in

Customer Prototypes Under Different Enrollment Approaches

Population of Interest (POI)



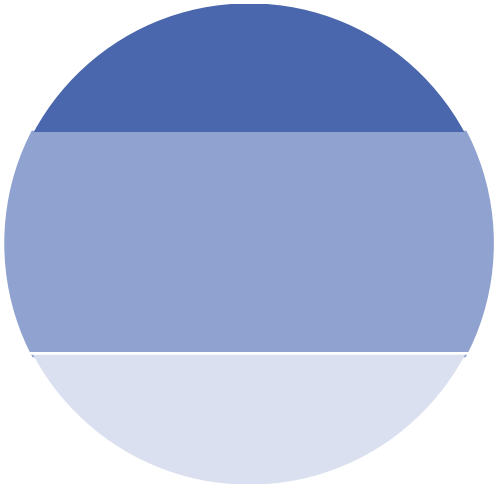
Never Takers
- Join if made mandatory
- Do not join if required to opt-out
- Do not join if required to opt-in

Complacents
- Join if made mandatory
- Join if required to opt-out
- Do not join if required to opt-in

Always Takers
- Join if made mandatory
- Join if required to opt-out
- Join if required to opt-in

Customer Prototypes Under Different Enrollment Approaches

Population of Interest (POI)



Never Takers
- Join if made mandatory
- Do not join if required to opt-out
- Do not join if required to opt-in

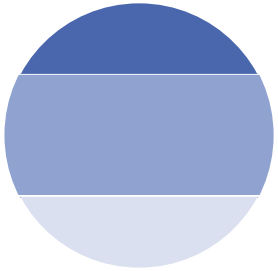
Complacents
- Join if made mandatory
- Join if required to opt-out
- Do not join if required to opt-in

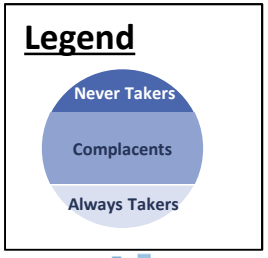
Always Takers
- Join if made mandatory
- Join if required to opt-out
- Join if required to opt-in

Critical Questions: How similar are these customers based on observable characteristics (e.g., usage, demographics, etc.)? How different are these customers in unobservable ways?

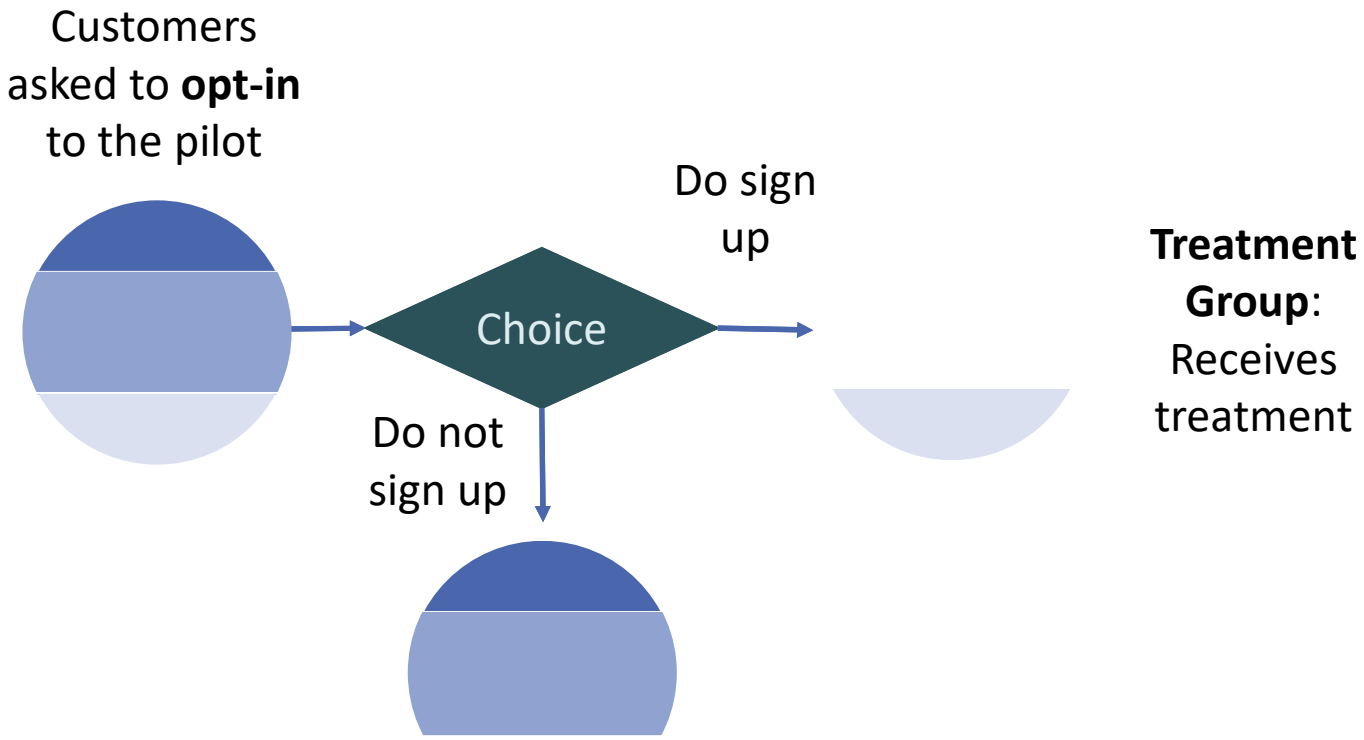
Common Study Design: Voluntary Enrollment

Customers asked to **opt-in** to the pilot

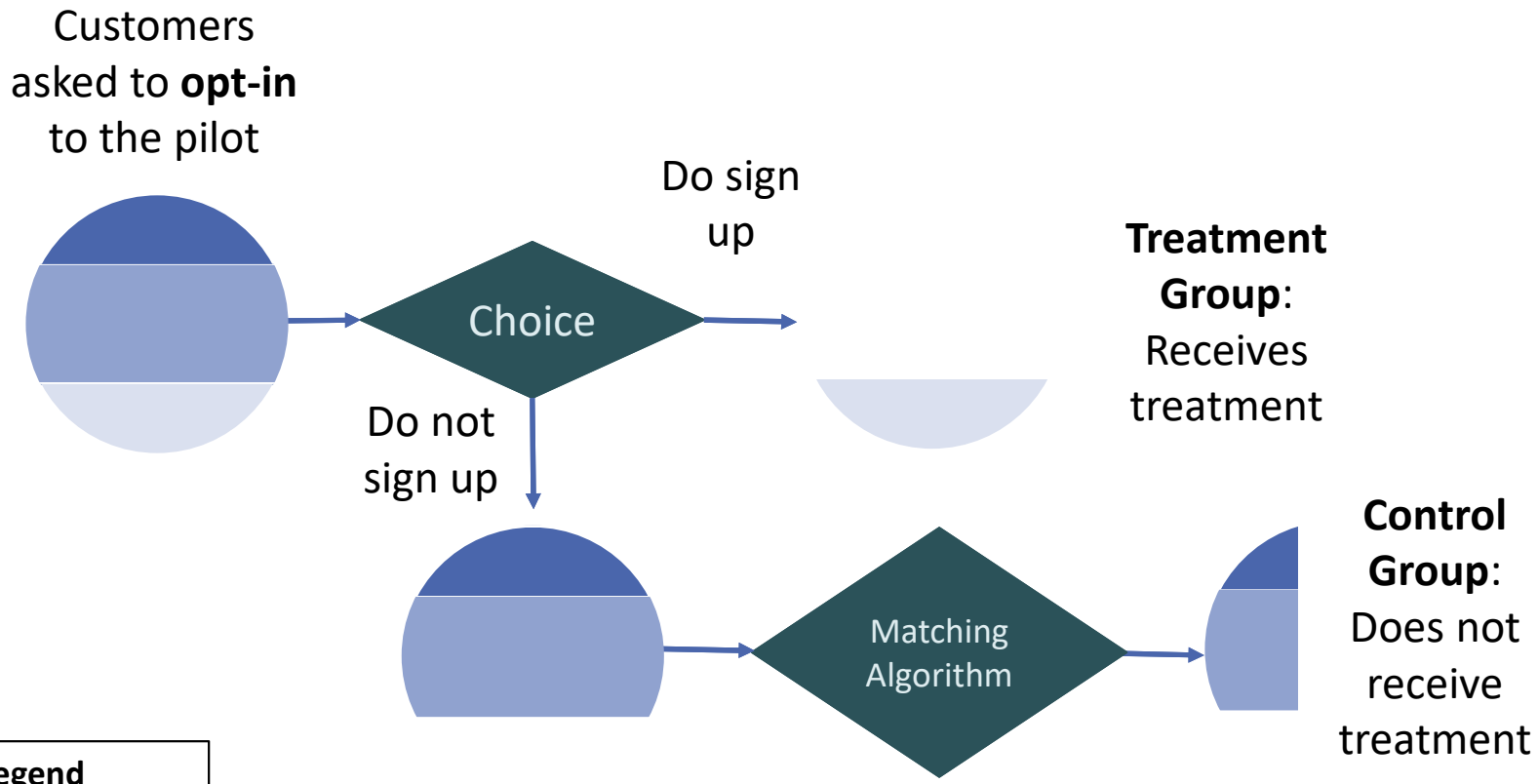




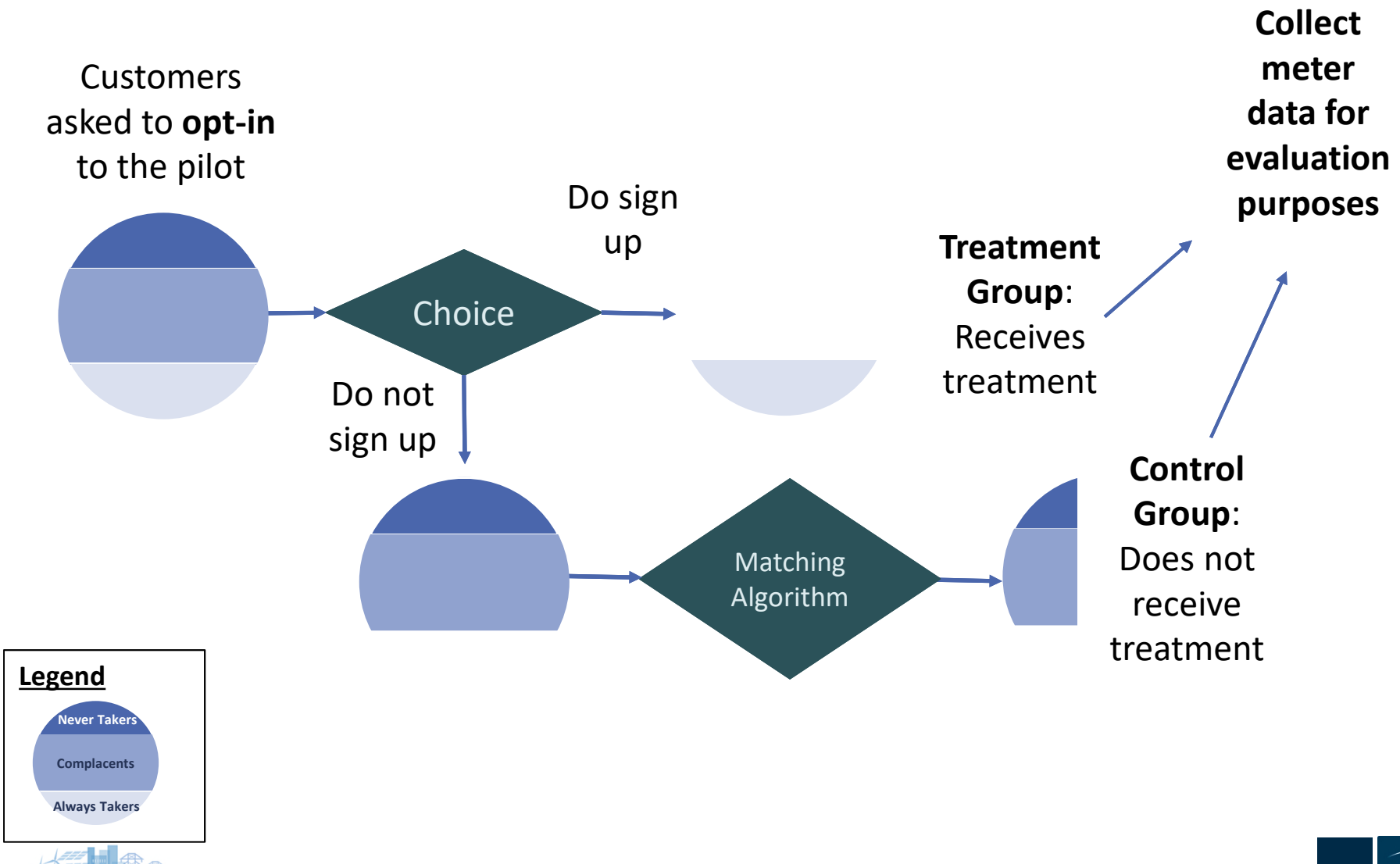
Common Study Design: Voluntary Enrollment



Common Study Design: Voluntary Enrollment

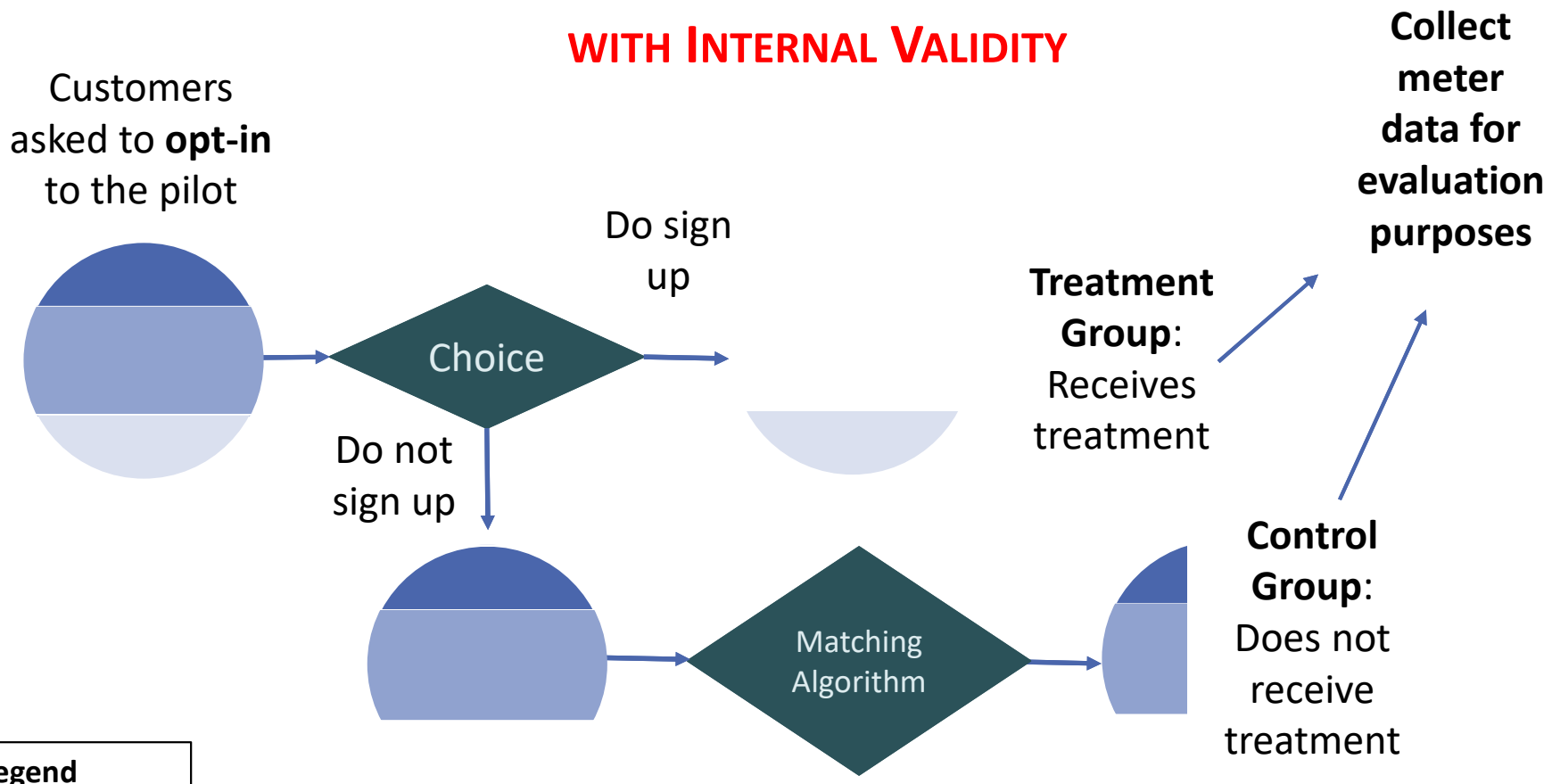


Common Study Design: Voluntary Enrollment



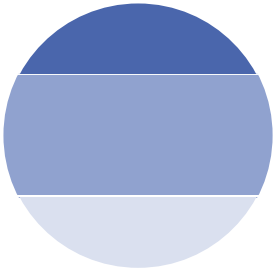
Common Study Design: Voluntary Enrollment

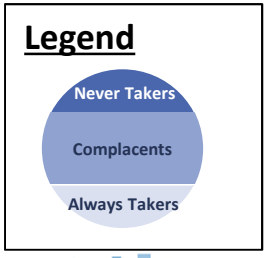
SELECTION BIAS LIKELY LEADS TO PROBLEMS WITH INTERNAL VALIDITY



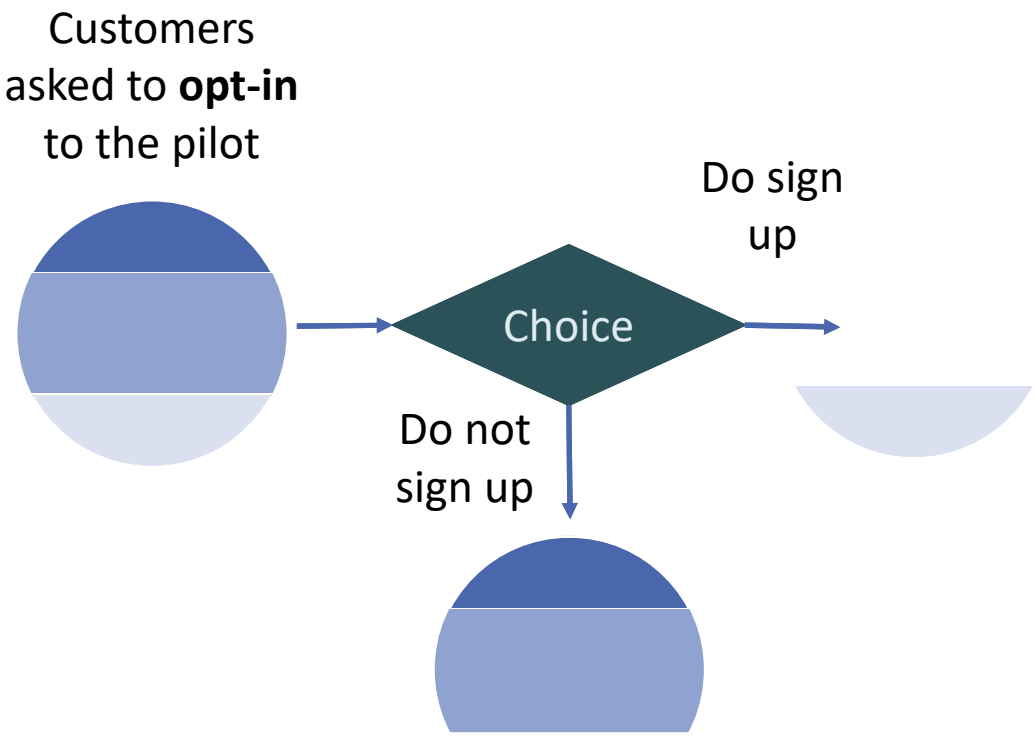
Randomized Control Trial: Voluntary Enrollment

Customers asked to **opt-in** to the pilot

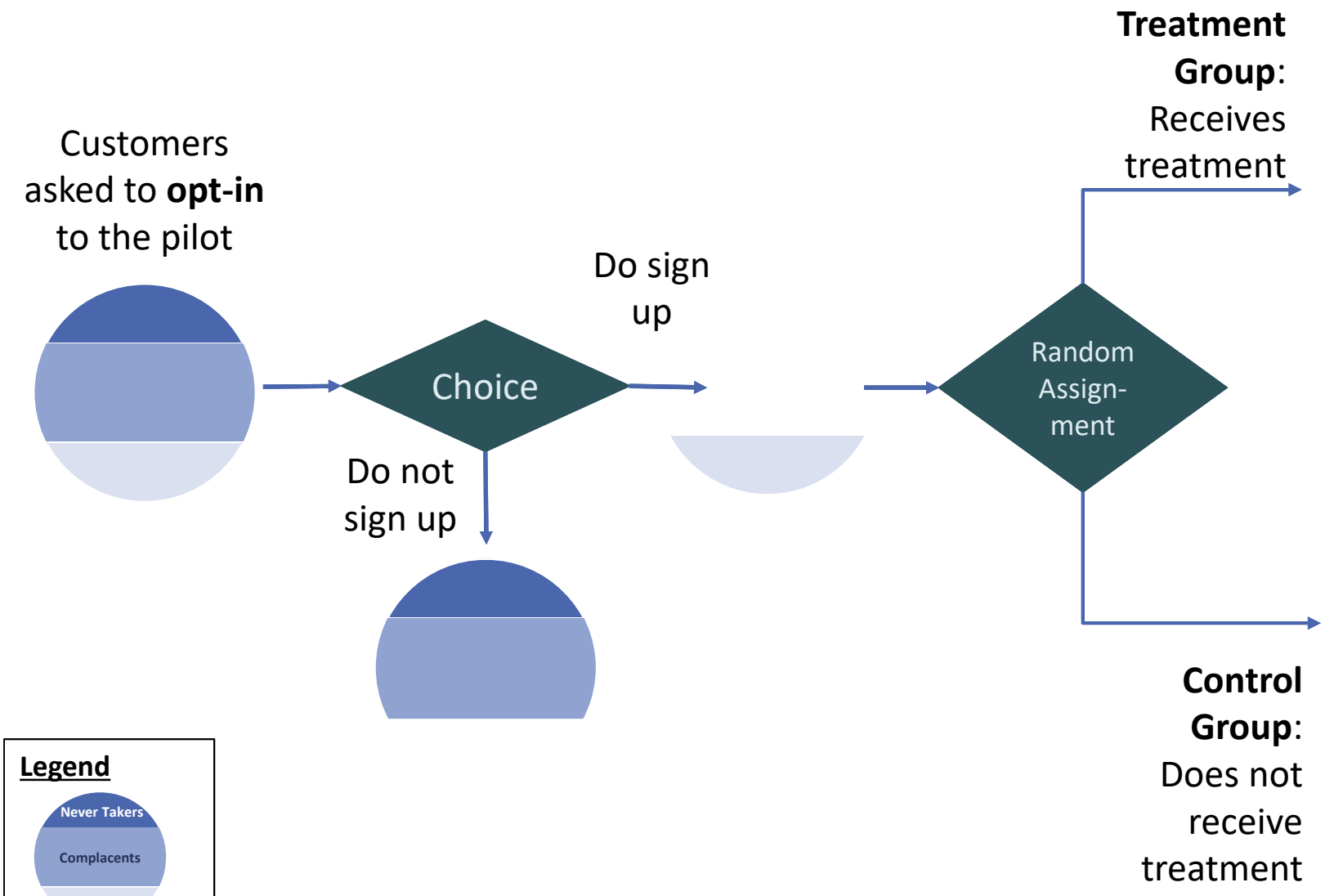




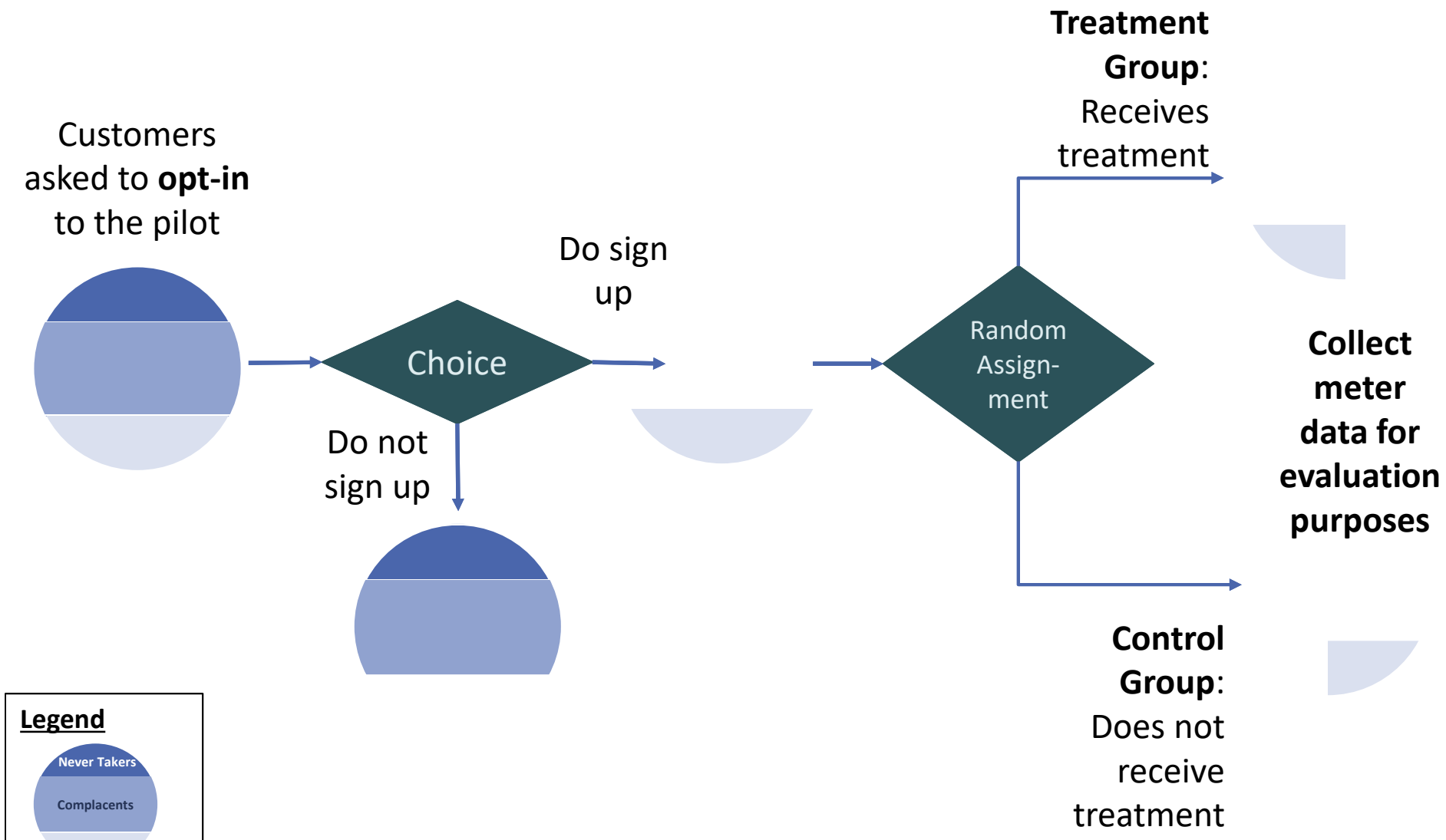
Randomized Control Trial: Voluntary Enrollment



Randomized Control Trial: Voluntary Enrollment

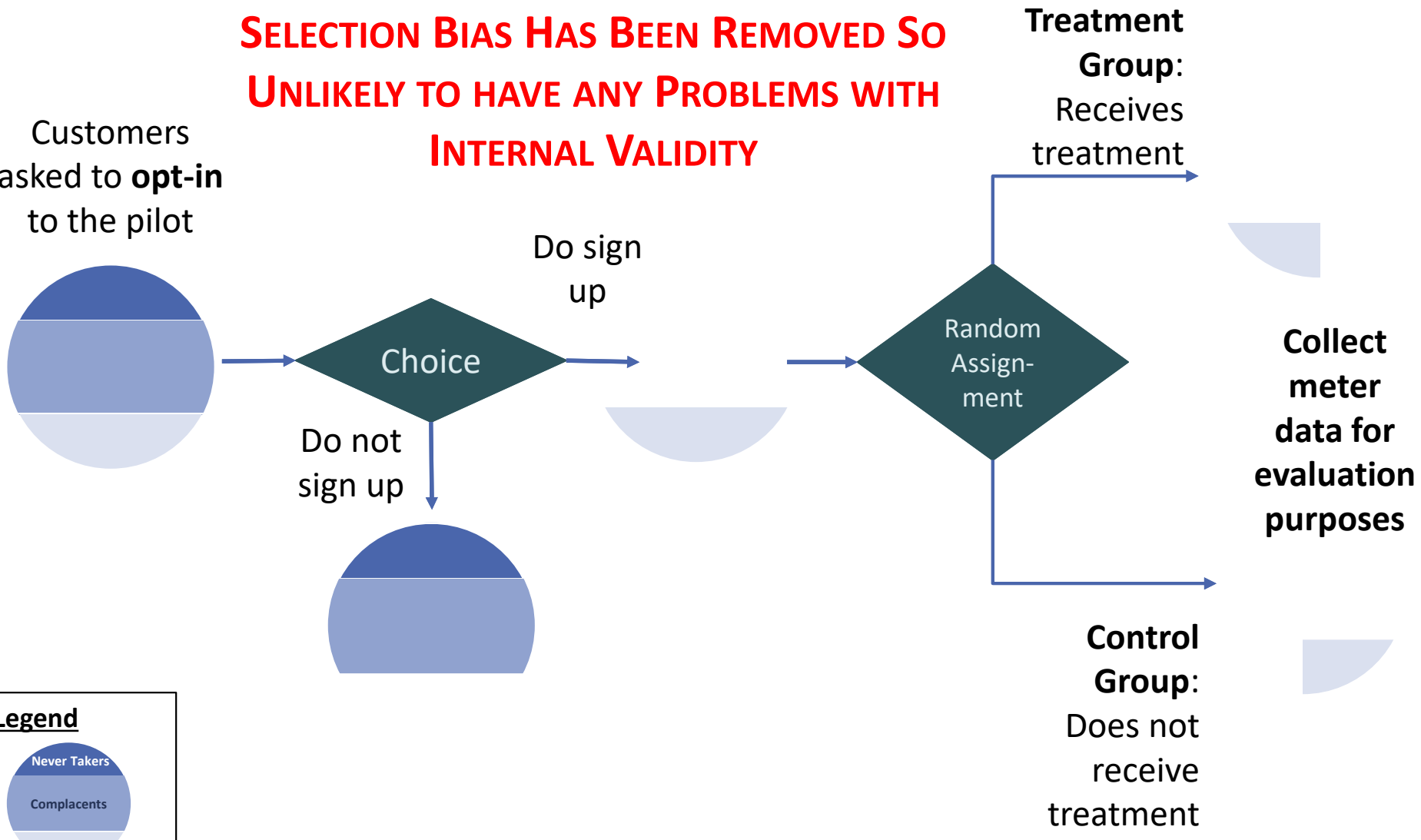


Randomized Control Trial: Voluntary Enrollment



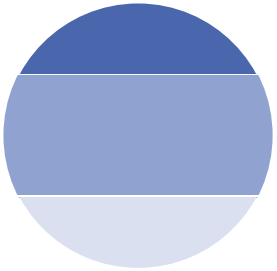
Randomized Control Trial: Voluntary Enrollment

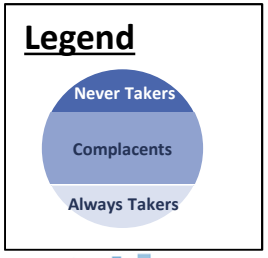
SELECTION BIAS HAS BEEN REMOVED SO UNLIKELY TO HAVE ANY PROBLEMS WITH INTERNAL VALIDITY



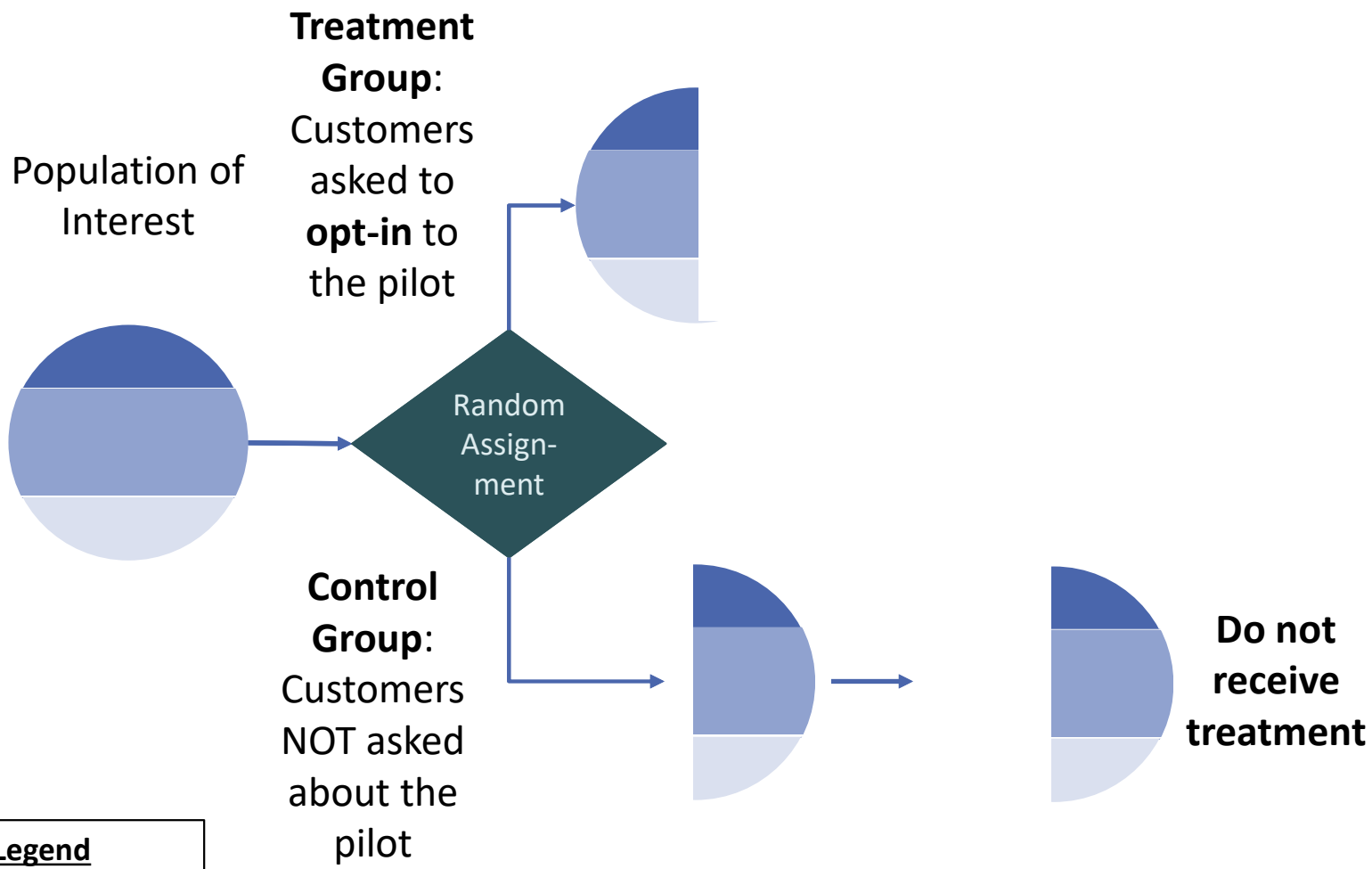
Randomized Encouragement Design: Voluntary Enrollment

Population of Interest

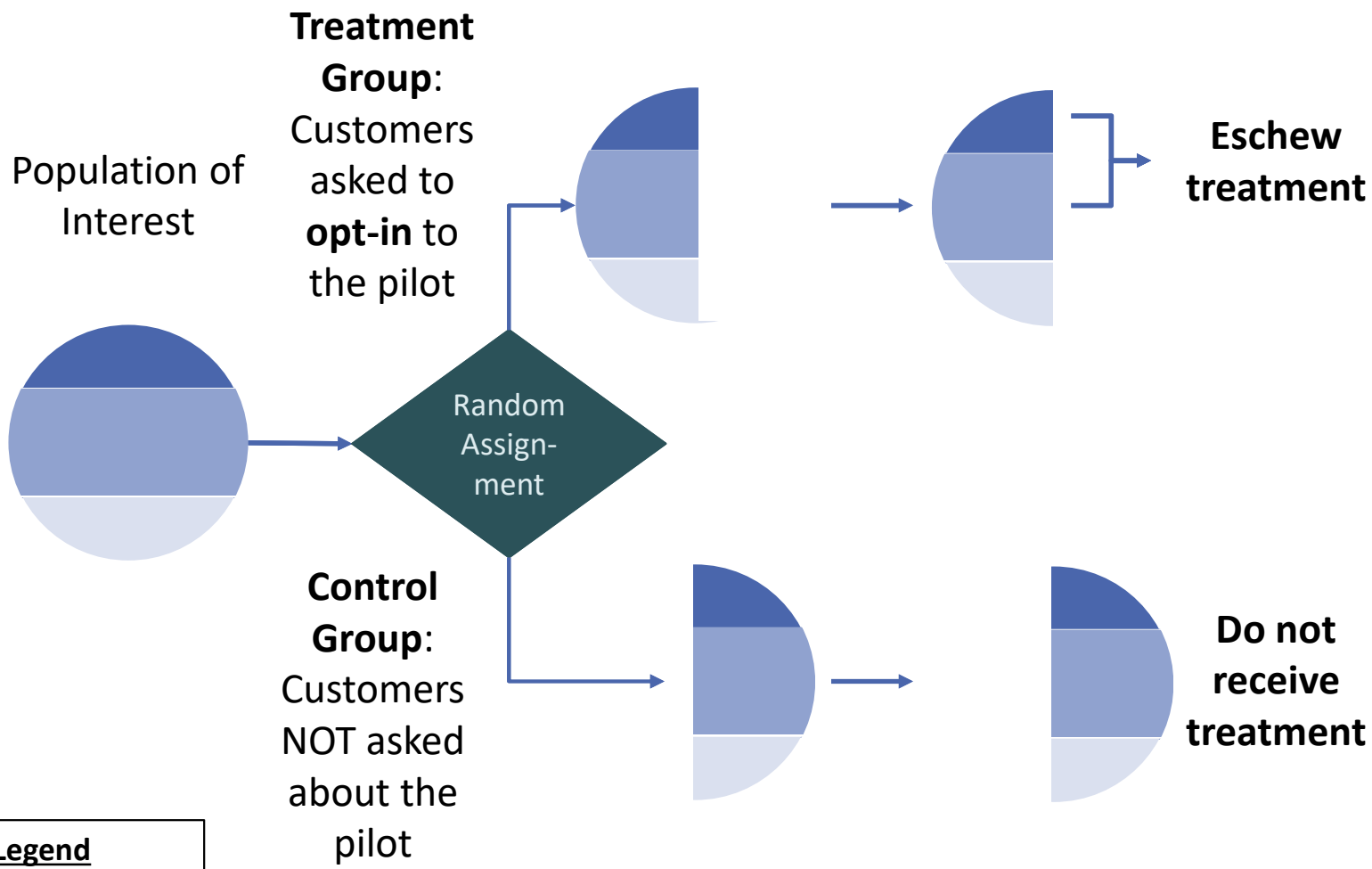




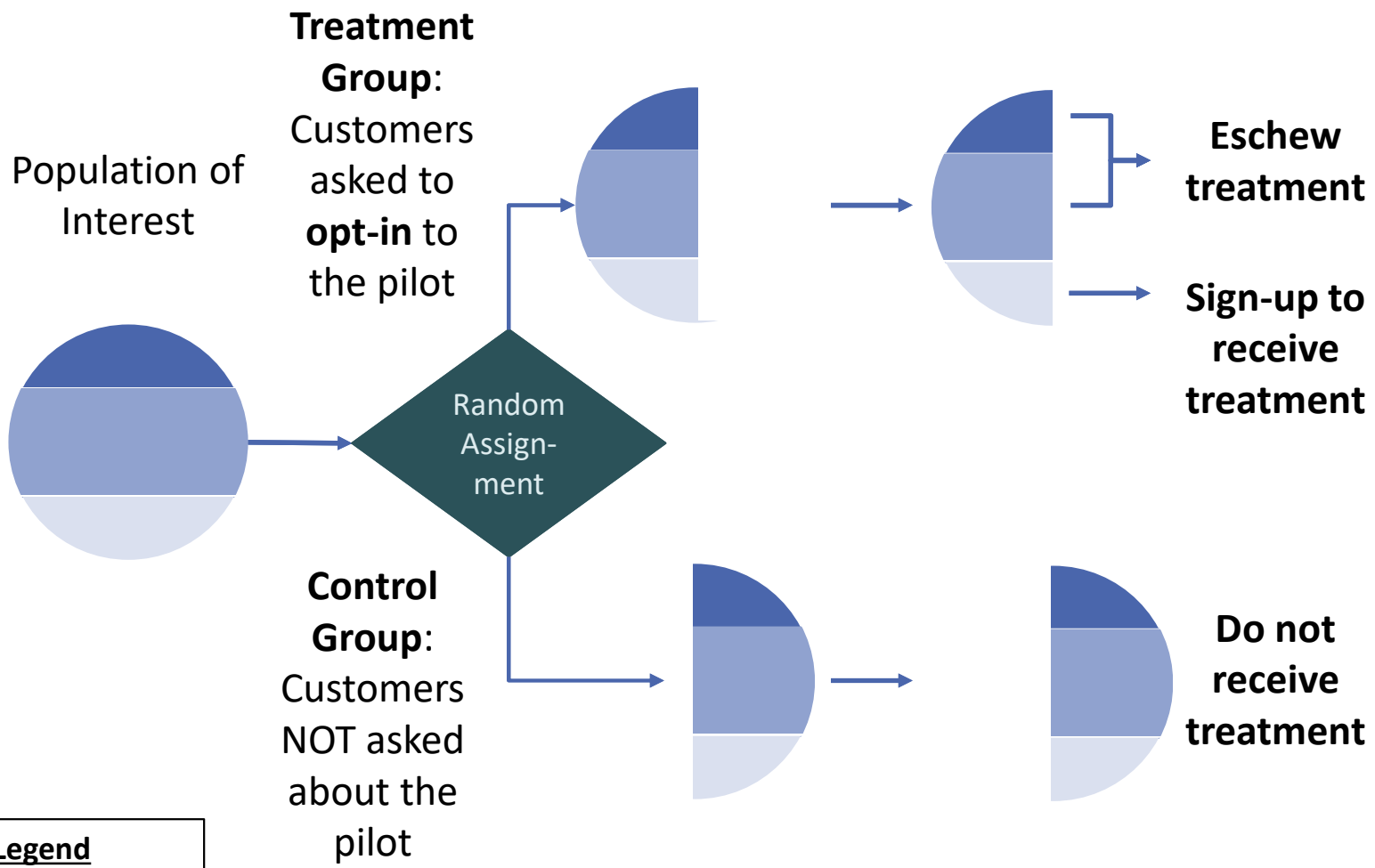
Randomized Encouragement Design: Voluntary Enrollment



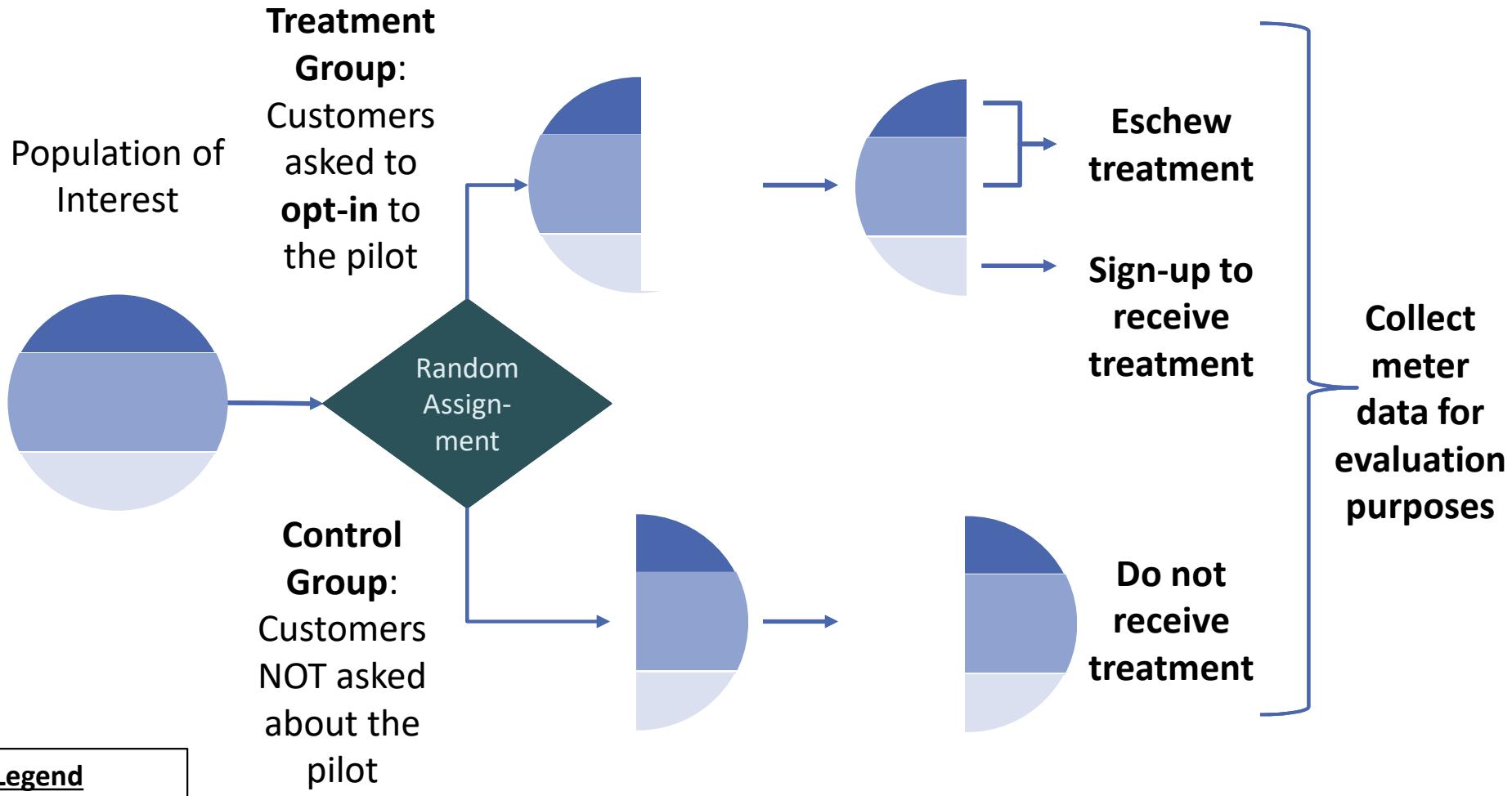
Randomized Encouragement Design: Voluntary Enrollment



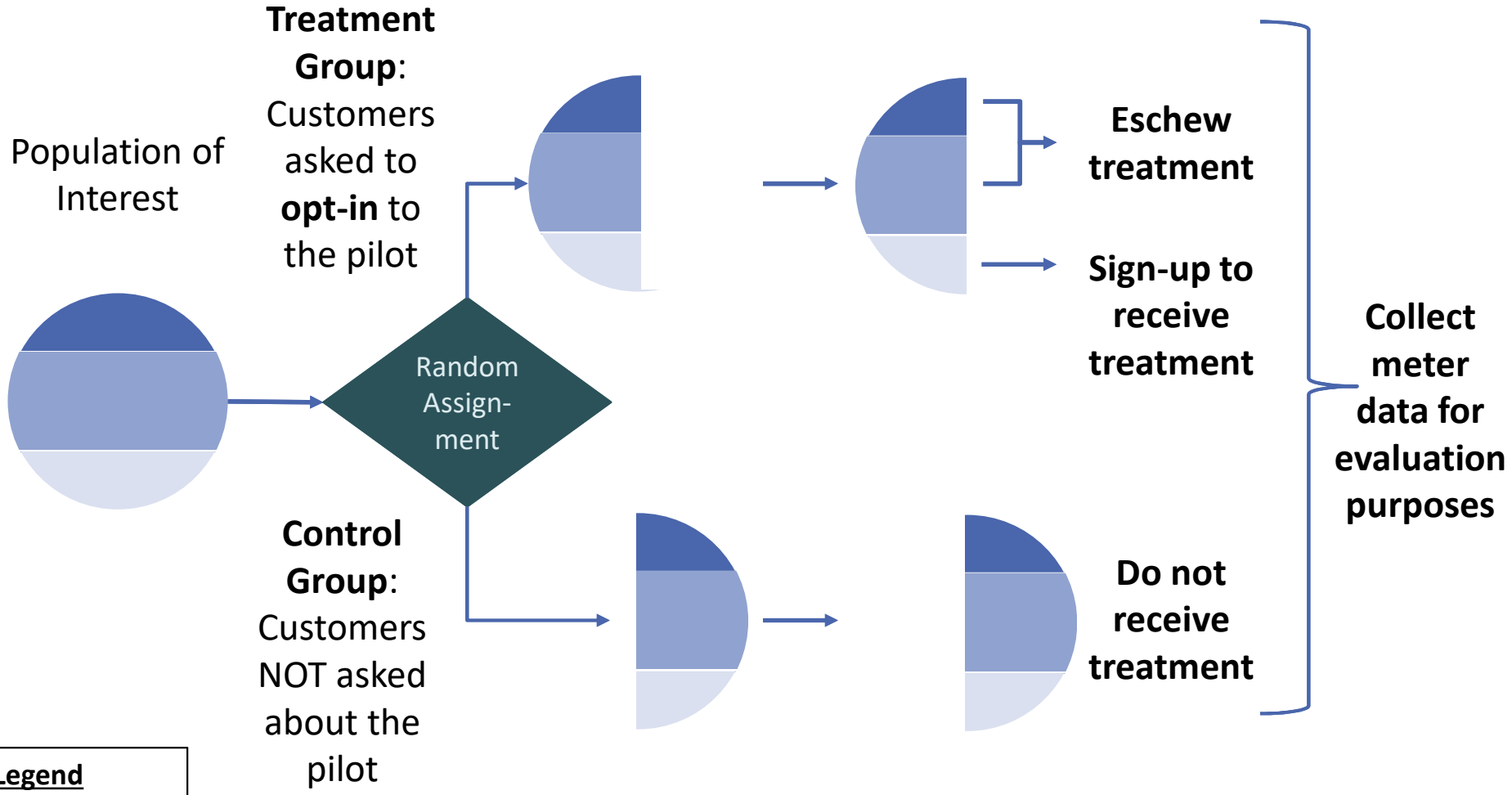
Randomized Encouragement Design: Voluntary Enrollment



Randomized Encouragement Design: Voluntary Enrollment

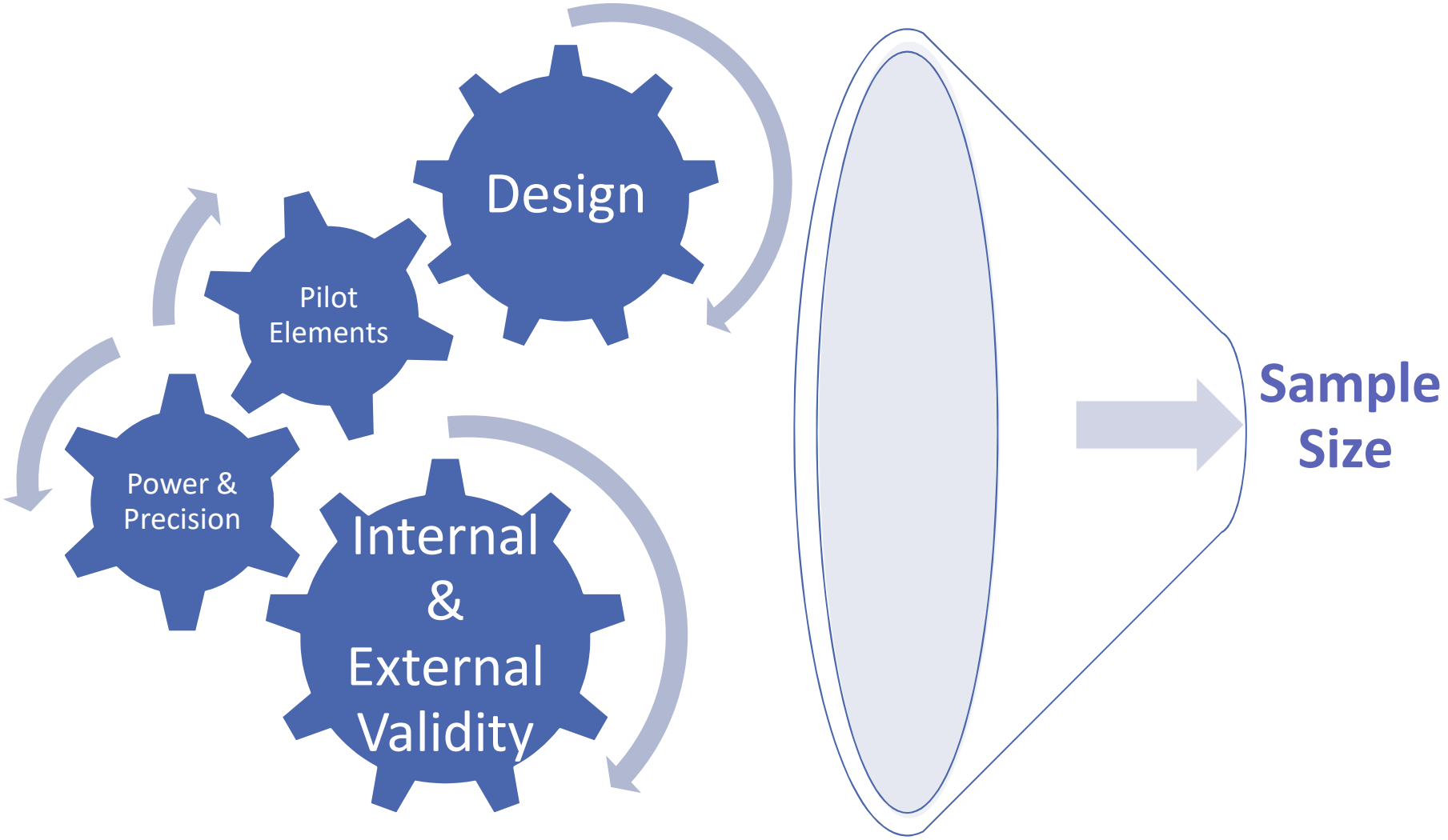


Randomized Encouragement Design: Voluntary Enrollment



SELECTION BIAS HAS BEEN REMOVED SO UNLIKELY TO HAVE ANY PROBLEMS WITH INTERNAL VALIDITY

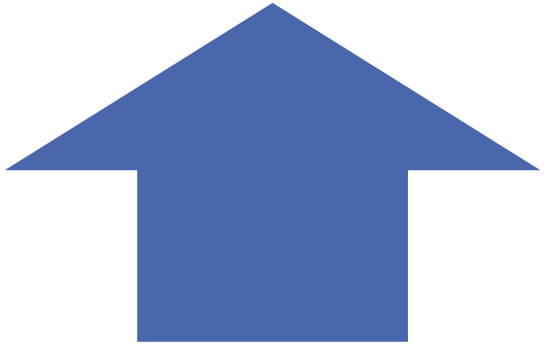
Process for Deriving Sample Sizes



Sample Sizes

Bigger sample sizes → larger pilot budgets

Bigger sample sizes → greater likelihood of statistical power and precision



Power Calculations: RCT

$$N = \frac{(t_{1-\kappa} + t_{\alpha})^2}{p(1-p)} \frac{\sigma^2}{MDE^2}$$

- N=Number of households in the study
- α = Type 1 Error Rate (0.1)
- k = Desired level of statistical power (0.80)
- p = Proportion of the sample receiving the treatment (0.50)
- MDE = Minimum Detectable Effect of the outcome (2% of coincident peak demand)
- σ = Variance of the outcome (based on observed data)

Power Calculations: RED

$$N = \frac{(t_{1-\kappa} + t_{\alpha})^2}{p(1-p)} \frac{\sigma^2}{MDE^2} \frac{1}{c^2}$$

- N=Number of households in the study
- α = Type 1 Error Rate (0.1)
- k = Desired level of statistical power (0.80)
- p = Proportion of the sample receiving the treatment (0.50)
- MDE = Minimum Detectable Effect of the outcome (2% of coincident peak demand)
- σ = Variance of the outcome (based on observed data)
- c = Enrollment rate

Power Calculations: RED

$$N = \frac{(t_{1-\kappa} + t_{\alpha})^2}{p(1-p)} \frac{\sigma^2}{MDE^2} \frac{1}{c^2}$$

Voluntary enrollment rate of 10% results in increase in sample size of 100 times relative to the sample for the RCT.

Voluntary enrollment rate of 20% results in increase in sample size of 25 times relative to the sample for the RCT.

- N=Number of households in the study
- α = Type 1 Error Rate (0.1)
- k = Desired level of statistical power (0.80)
- p = Proportion of the sample receiving the treatment (0.50)
- MDE = Minimum Detectable Effect of the outcome (2% of coincident peak demand)
- σ = Variance of the outcome (based on observed data)
- c = Enrollment rate

Power Calculations: RED

$$N = \frac{(t_{1-\kappa} + t_{\alpha})^2}{p(1-p)} \frac{\sigma^2}{MDE^2} \frac{1}{c^2}$$

Default enrollment rate of 90% results in increase in sample size of **23%**, relative to RCT.

Default enrollment rate of 80% results in increase in sample size of **56%**, relative to RCT

- N=Number of households in the study
- α = Type 1 Error Rate (0.1)
- k = Desired level of statistical power (0.80)
- p = Proportion of the sample receiving the treatment (0.50)
- MDE = Minimum Detectable Effect of the outcome (2% of coincident peak demand)
- σ = Variance of the outcome (based on observed data)
- c = Enrollment rate