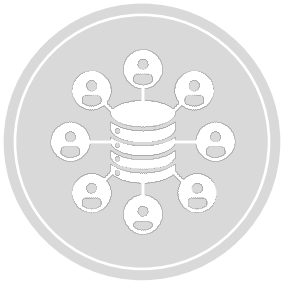
A guide to:
Experimental design

Key concepts
& study plan

Experimental design



Data collection
& processing

Model specification
& estimation

Interpretation
& application

A guide to: Experimental design

Steps in designing choice experiments

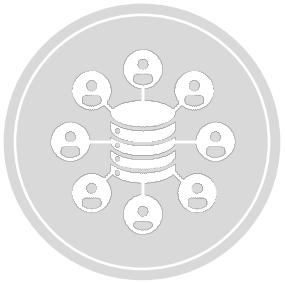
1. Determine design type
2. Determine design size
3. Generate design
4. Block design
5. Create choice tasks in survey instrument

Each study is different, some steps may not apply or in a different order

Design considerations

Key concepts & study plan

Experimental design



Data collection & processing

Model specification & estimation

Interpretation & application

Design considerations

Experimental design

- Matrix with attribute level combinations, where
  - Each column represents an attribute of an alternative
  - Each row represents a choice task

Bike	Bus A		Bus B	
dist	$time_A$	$cost_A$	$time_B$	$cost_B$
10	10	1	20	3
2	20	2	30	1
5	30	3	10	2
5	10	2	20	1
5	20	3	30	2
5	30	1	10	3
5	10	2	20	1
5	20	1	10	2
10	30	3	30	3

Design considerations

Realism

- A good design contains realistic attribute level combinations
- Non-sensible attribute level combinations should be avoided

Bike	Bus A		Bus B	
dist	$time_A$	$cost_A$	$time_B$	$cost_B$
10	10	1	20	3
2	20	2	30	1
5	30	3	10	2
5	10	2	20	1
5	20	3	30	2
5	30	1	10	3
5	10	2	20	1
5	20	1	10	2
10	30	3	30	3

Unrealistic attribute level combinations

Design considerations

Balance

- A good design contains a high degree of attribute level balance
- Highly unbalanced designs should be avoided

Bike	Bus A		Bus B	
dist	$time_A$	$cost_A$	$time_B$	$cost_B$
10	10	1	20	3
2	20	2	30	1
5	30	3	10	2
5	10	2	20	1
5	20	3	30	2
5	30	1	10	3
5	10	2	20	1
5	20	1	10	2
10	30	3	30	3

Unbalanced attribute levels

Design considerations

Variety

- A good design contains a variety of attribute level combinations
- Repeated or similar choice tasks should be avoided

Bike	Bus A		Bus B	
dist	$time_A$	$cost_A$	$time_B$	$cost_B$
10	10	1	20	3
2	20	2	30	1
5	30	3	10	2
5	10	2	20	1
5	20	3	30	2
5	30	1	10	3
5	10	2	20	1
5	20	1	10	2
10	30	3	30	3

Repeated choice tasks

Design considerations

Trade-offs

- A good design allows trade-offs between attributes
- Choice tasks with dominant alternatives should be avoided

Bike	Bus A		Bus B	
dist	$time_A$	$cost_A$	$time_B$	$cost_B$
10	10	1	20	3
2	20	2	30	1
5	30	3	10	2
5	10	2	20	1
5	20	3	30	2
5	30	1	10	3
5	10	2	20	1
5	20	1	10	2
10	30	3	30	3

No trade-off

Design considerations

Trade-offs

- A good design allows trade-offs between attributes
- Choice tasks with dominant alternatives should be avoided

Bike	Bus A		Bus B	
dist	$time_A$	$cost_A$	$time_B$	$cost_B$
10	10	1	20	3
2	20	2	30	1
5	30	3	10	2
5	10	2	20	1
5	20	3	30	2
5	30	1	10	3
5	10	2	20	1
5	20	1	10	2
10	30	3	30	3

Bus B is dominant

Design considerations

Dominant alternative

- Alternative that is best across all attributes
- Often present in unlabelled experiments
- Choice task with dominant alternative provides no information, and should be detected and avoided

Design considerations

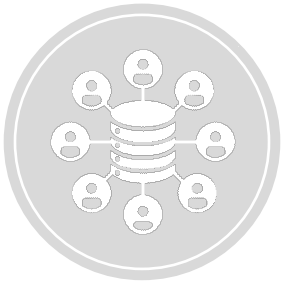
Dominant alternative

- Results in silly choice tasks

Step 1 – Determine design type

Key concepts & study plan

Experimental design



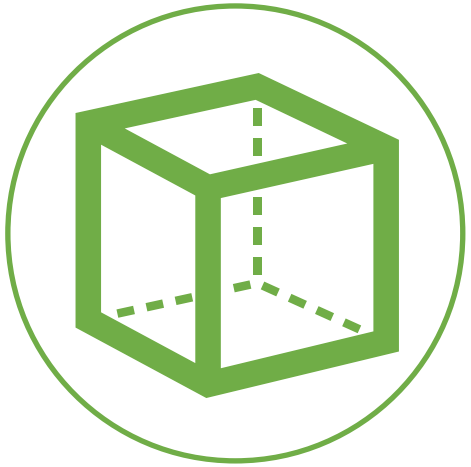
Data collection & processing

Model specification & estimation

Interpretation & application

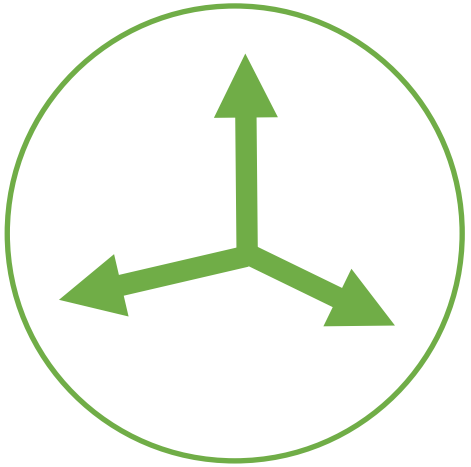
Step 1 – Determine design type

Design types

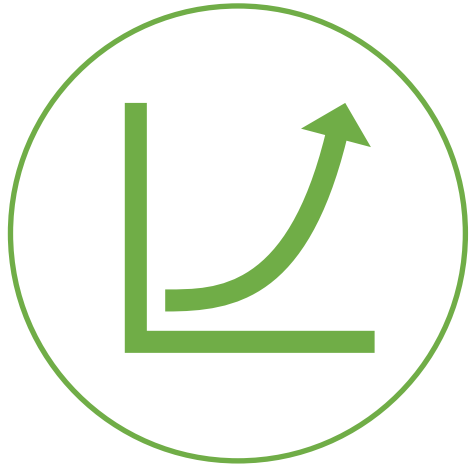


Full factorial design

Random design



Orthogonal design

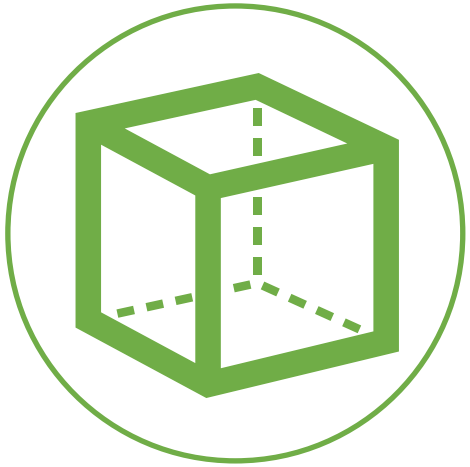


Efficient design

Step 1 – Determine design type

Design types

Fractional factorial designs



Full factorial design

all possible
attribute level combinations

Random design

random
attribute level combinations

Orthogonal design

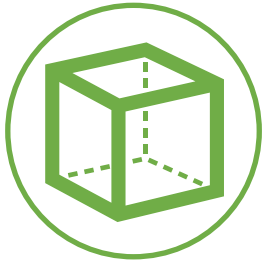
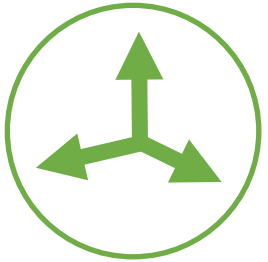
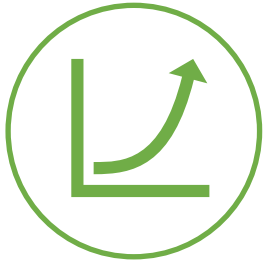
balanced
attribute level combinations

Efficient design

trade-off maximising
attribute level combinations

Step 1 – Determine design type

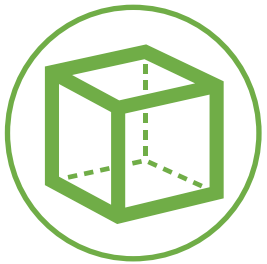
Each design type has different strengths and weaknesses

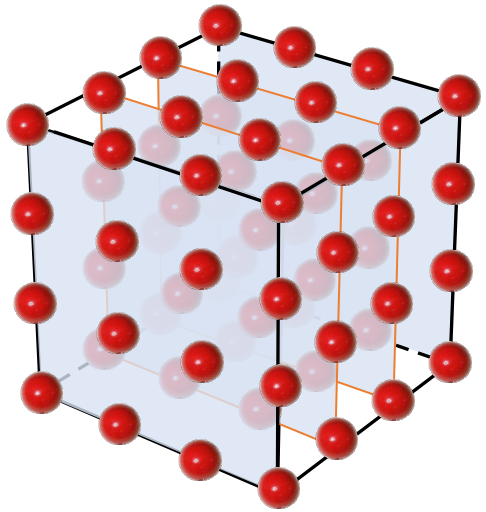
	Full factorial	Random	Orthogonal	Efficient
				
Realism	●○○	●●○	●○○	●●●
Trade-offs	●○○	●○○	●●○	●●●
Balance	●●●	●○○	●●●	●●○
Variety	●●●	●●●	●●○	●●○

Step 1 – Determine design type

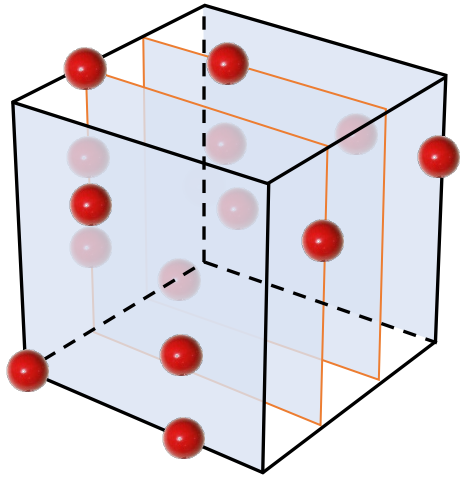
Each design type results in different attribute level combinations

Full factorial

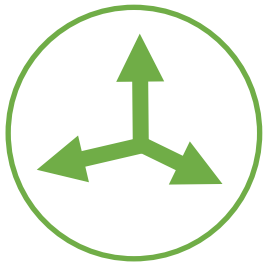


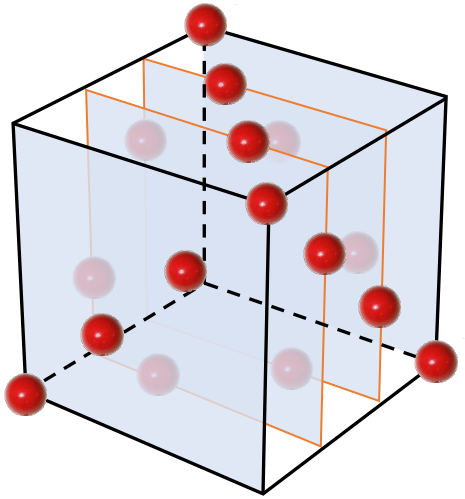


Random

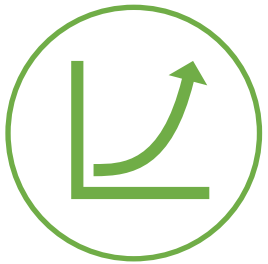


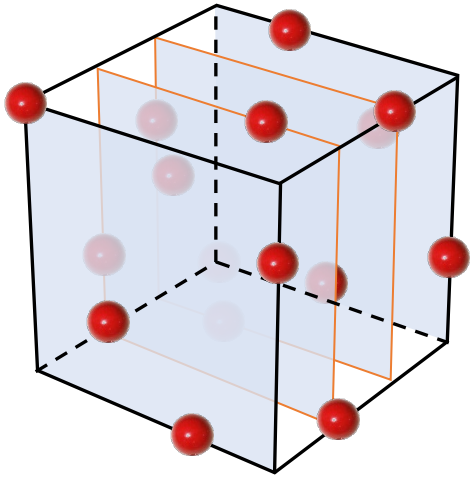
Orthogonal





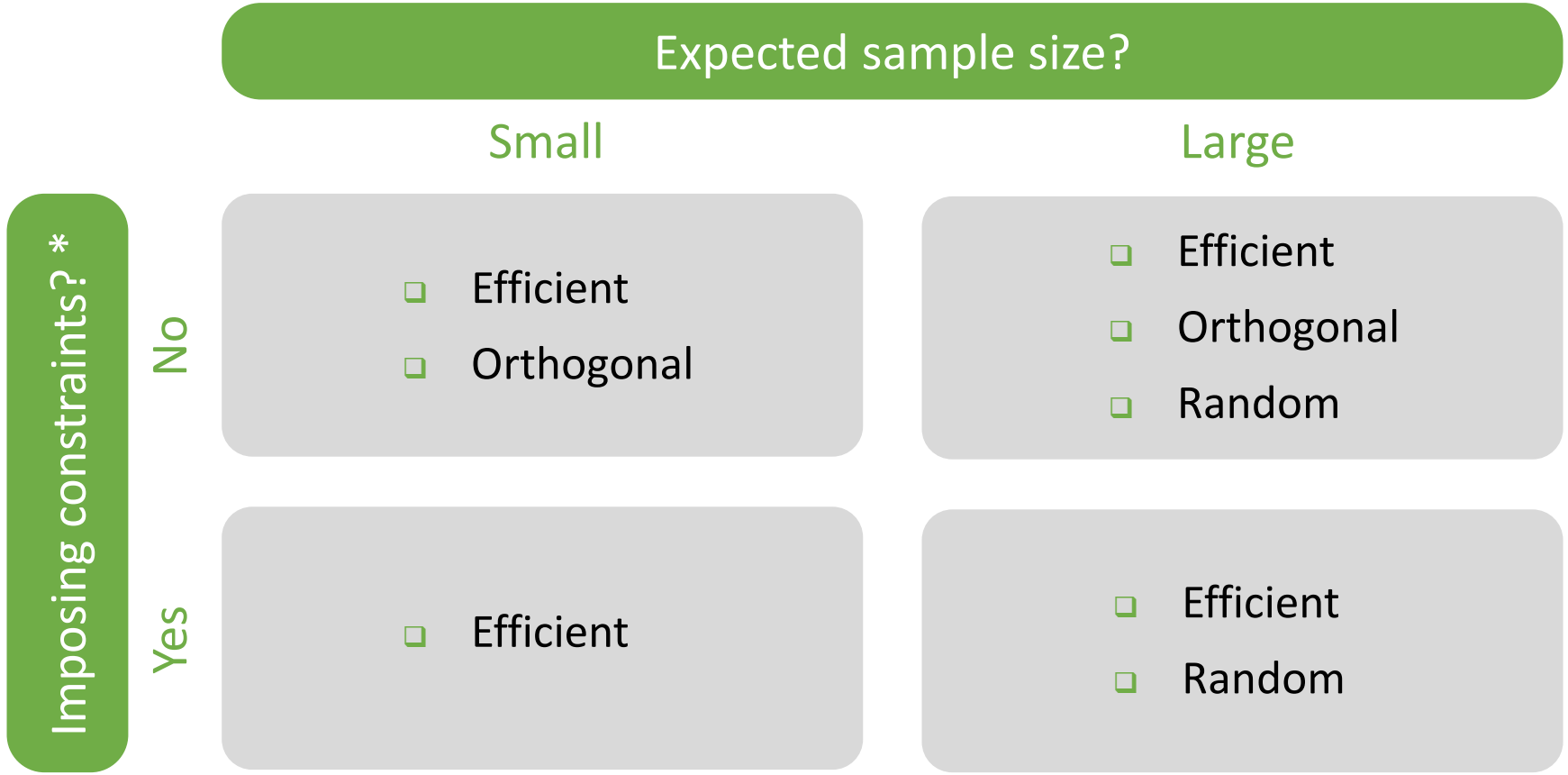
Efficient





Step 1 – Determine design type

Selecting best design type

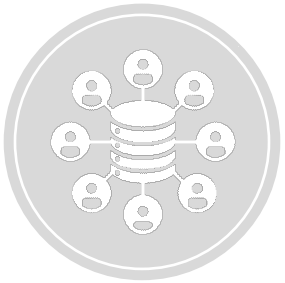


* Prohibitions of certain attribute level combinations to ensure realism, non-dominance, etc.

Step 2 – Determine design size

Key concepts & study plan

Experimental design



Data collection & processing

Model specification & estimation

Interpretation & application

Step 2 – Determine design size

Design size

- Number of choice tasks (rows) in an experimental design

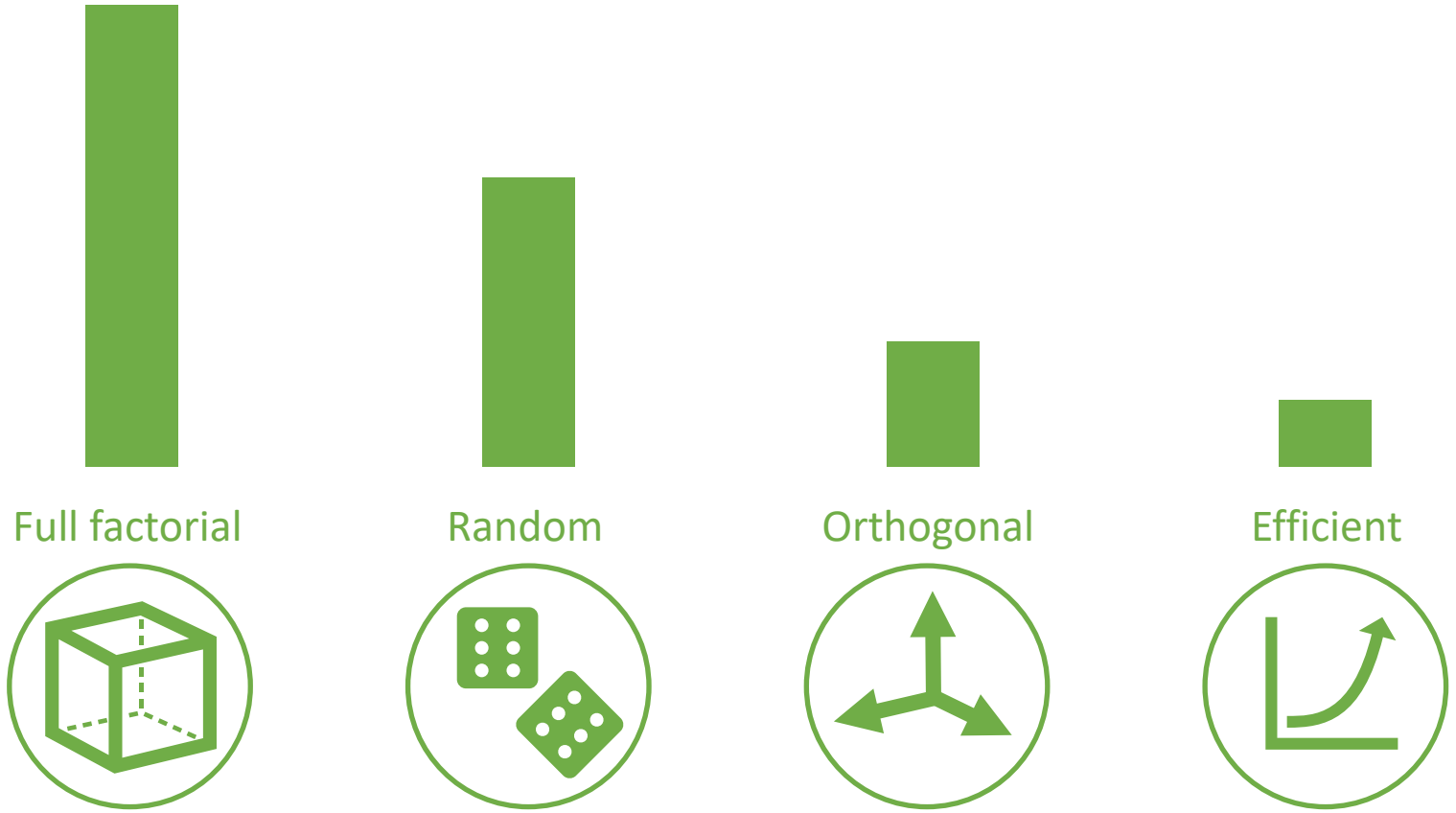
Bike	Bus A		Bus B	
dist	$time_A$	$cost_A$	$time_B$	$cost_B$
2	10	1	20	3
2	20	2	30	1
2	30	3	10	2
5	10	2	20	1
5	20	3	30	2
5	30	1	10	3
10	10	3	20	2
10	20	1	30	3
10	30	2	10	1

design size = 9

Step 2 – Determine design size

Appropriate design size varies by design type

- Typical design size



Step 2 – Determine design size

Full factorial design is typically too large

- The full factorial design for this experiment contains 41,472 choice tasks!

Bike	Bus			Train			Tram		
dist	time	cost	wait	time	cost	wait	time	cost	wait
2	20	1	10	10	3	5	15	2	8
5	30	2	20	20	4	10	25	3	15
10	40	3		30	5		35	4	
		4			6			5	

levels

$$3 \times 3 \times 4 \times 2 \times 3 \times 4 \times 2 \times 3 \times 4 \times 2 = 41{,}472$$

Step 2 – Determine design size

Minimum design size

- Theoretical minimum
  - Statistical requirement for any design

$$\text{design size} > \frac{\text{number of parameters}}{\text{number of alternatives} - 1}$$

- Recommended
  - Multiplier M > 1
  - My rule of thumb:
    - Random design: M = 10
    - Efficient design: M = 3

$$\text{design size} > M \cdot \frac{\text{number of parameters}}{\text{number of alternatives} - 1}$$

Step 2 – Determine design size

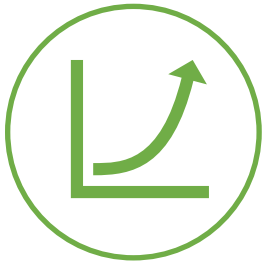
Minimum design size – Example efficient design

- Theoretical minimum
  - Statistical requirement for any design

$$\text{design size} > \frac{10}{3-1}$$

- Recommended
  - Multiplier M = 3

$$\text{design size} > 3 \cdot \frac{10}{3-1}$$



$$\text{Utility(Car)} = \beta_1 \cdot \text{TravelTime}_{car} + \beta_2 \cdot \text{FuelCost} + \beta_3 \cdot \text{TollCost}$$

$$\text{Utility(Bus)} = \beta_4 + \beta_5 \cdot \text{TravelTime}_{bus} + \beta_6 \cdot \text{WalkTime} + \beta_7 \cdot \text{NrTransfers} + \beta_2 \cdot \text{Fare}$$

$$\text{Utility(Bike)} = \beta_8 + \beta_9 \cdot \text{TravelTime}_{bike} + \beta_{10} \cdot \text{Rain}$$

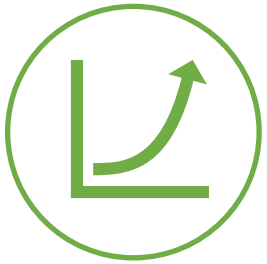
Step 2 – Determine design size

Minimum design size – Example efficient design

- Theoretical minimum
  - Statistical requirement for any design
- Recommended
  - Multiplier M = 3

$$\text{design size} > 5$$

$$\text{design size} > 15$$



$$\text{Utility(Car)} = \beta_1 \cdot \text{TravelTime}_{car} + \beta_2 \cdot \text{FuelCost} + \beta_3 \cdot \text{TollCost}$$

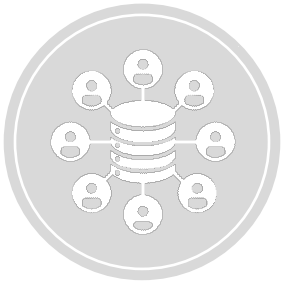
$$\text{Utility(Bus)} = \beta_4 + \beta_5 \cdot \text{TravelTime}_{bus} + \beta_6 \cdot \text{WalkTime} + \beta_7 \cdot \text{NrTransfers} + \beta_2 \cdot \text{Fare}$$

$$\text{Utility(Bike)} = \beta_8 + \beta_9 \cdot \text{TravelTime}_{bike} + \beta_{10} \cdot \text{Rain}$$

Step 3 – Generate design

Key concepts & study plan

Experimental design



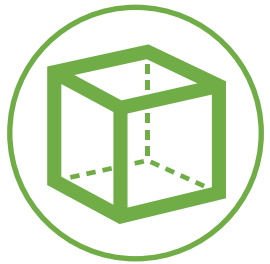
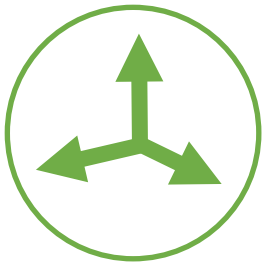
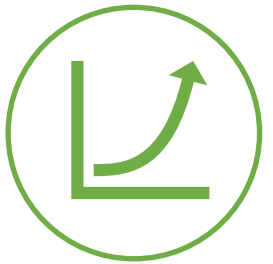
Data collection & processing

Model specification & estimation

Interpretation & application

Step 3 – Generate design

Design generation tools

	Full factorial	Random	Orthogonal	Efficient
				
❑ Spreadsheet	■	■		□
❑ Design library			■	
❑ Design software	■	■	■	■

Step 3 – Generate design

Software for generating designs for choice experiments

- Ngene
- SAS
- JMP
- Stata
- SPSS - orthogonal designs only
- Sawtooth
- Idefix in R - efficient designs only, free

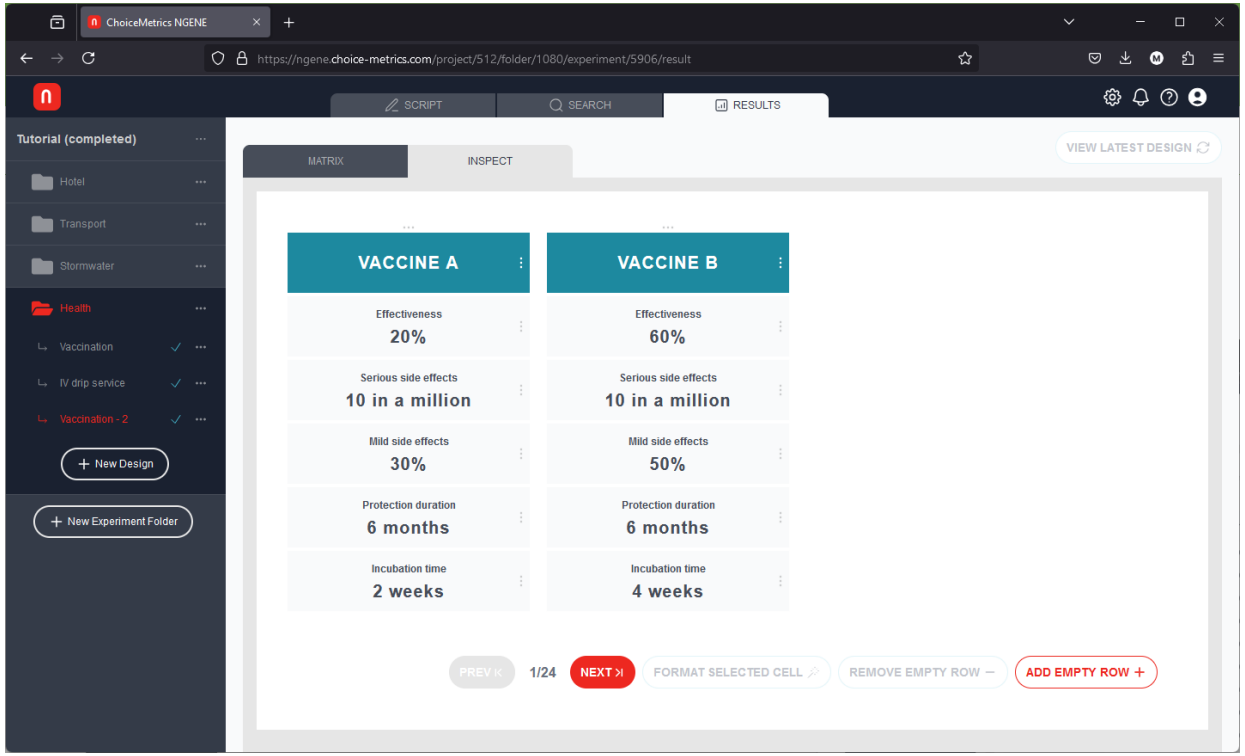
Step 3 – Generate design

Ngene

- Used in this course

www.choice-metrics.com

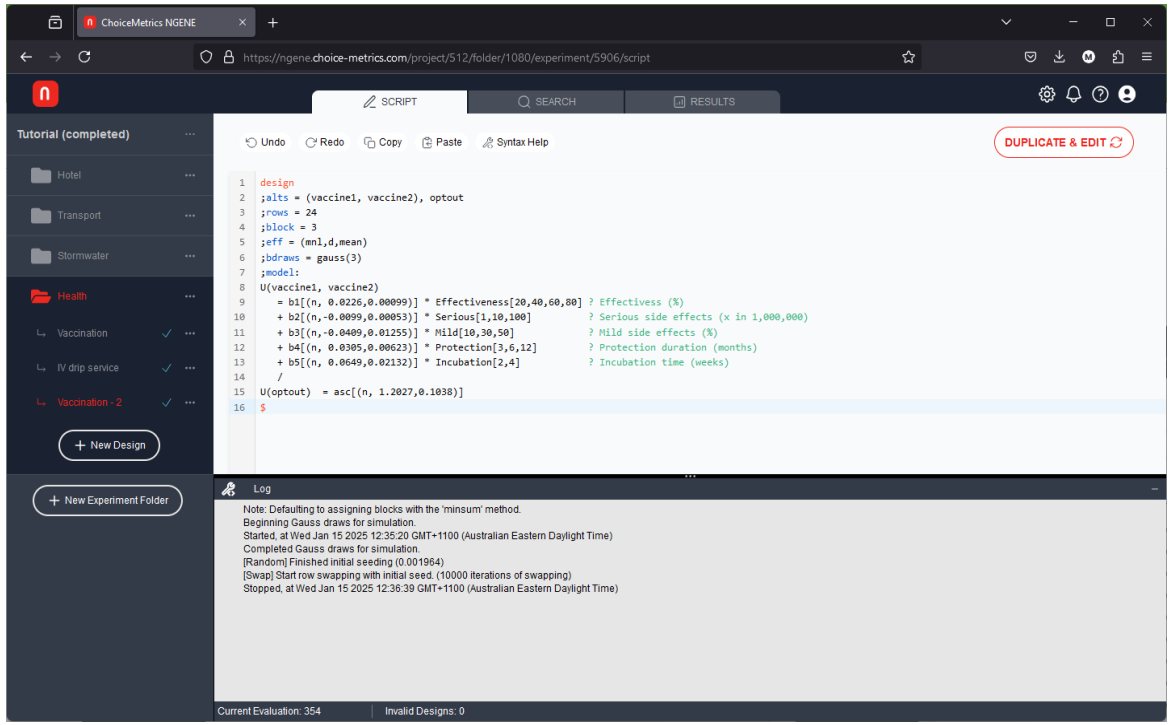
- Full factorial designs
- Random designs
- Orthogonal designs
- Efficient designs

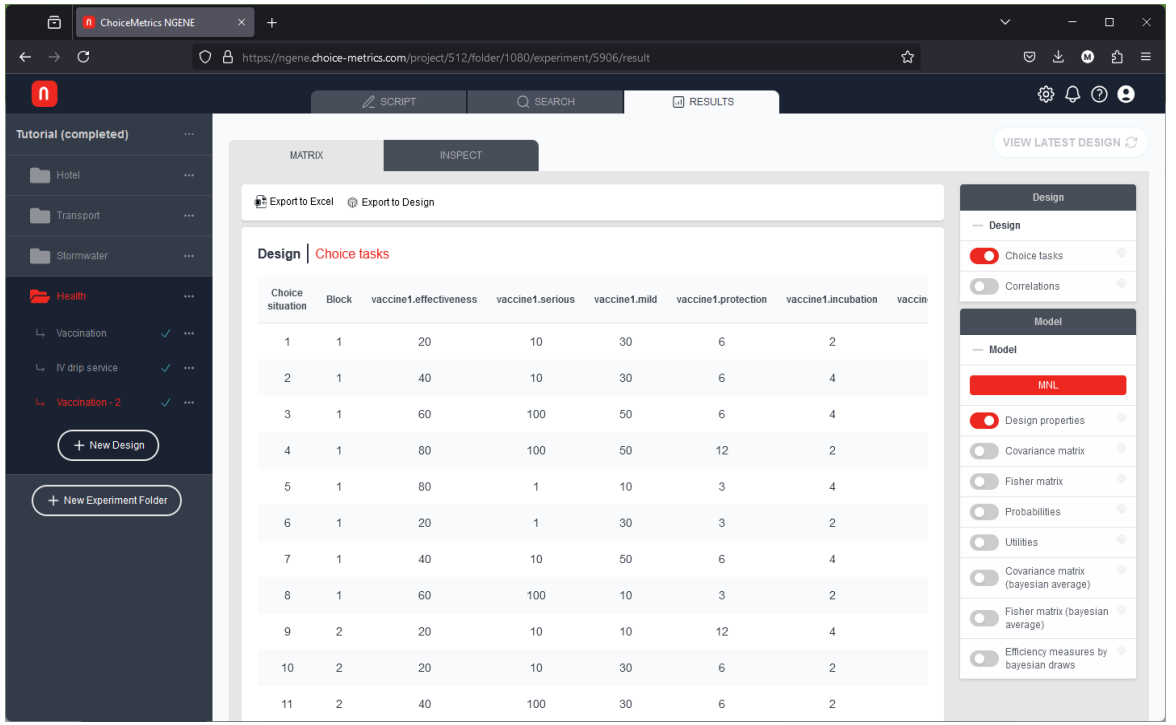


Step 3 – Generate design

Ngene

- Specify choice model and design properties in script
- Run script to generate design

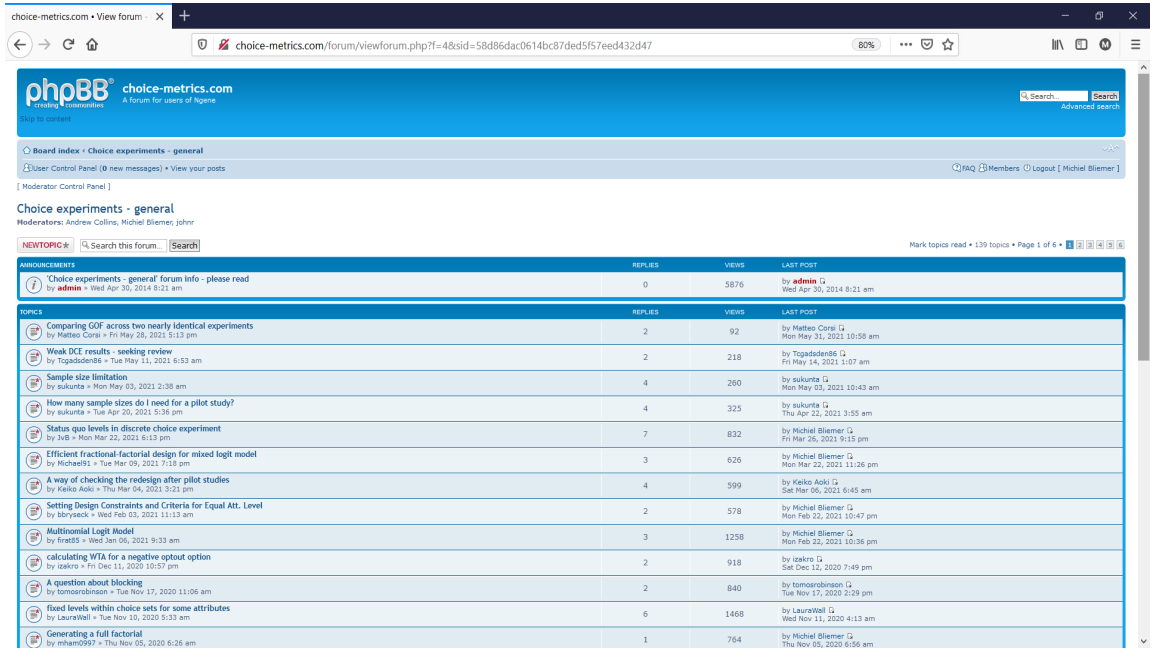


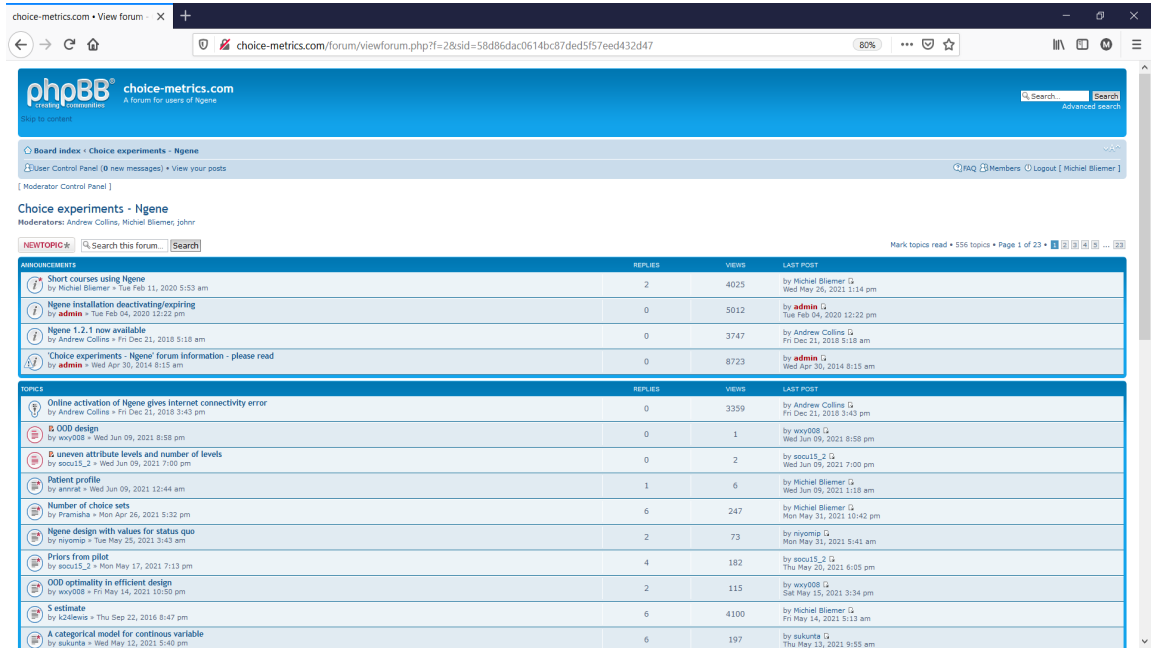


Step 3 – Generate design

Ngene

- Active freely accessible forums for questions about
  - Ngene
  - Experimental design in general

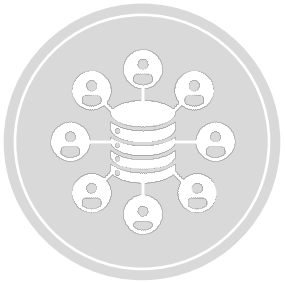




Step 4 – Block design

Key concepts & study plan

Experimental design



Data collection & processing

Model specification & estimation

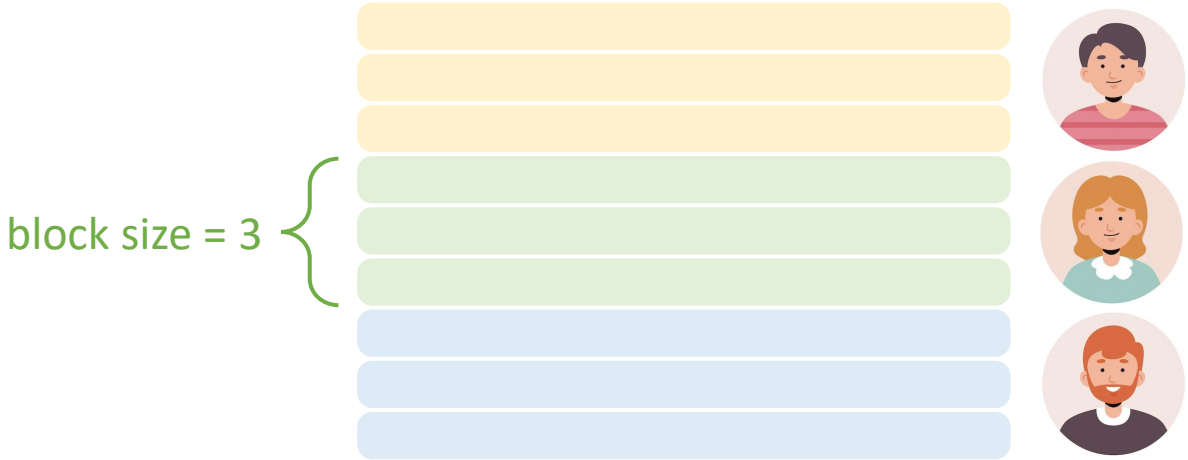
Interpretation & application

Step 4 – Block design

Block

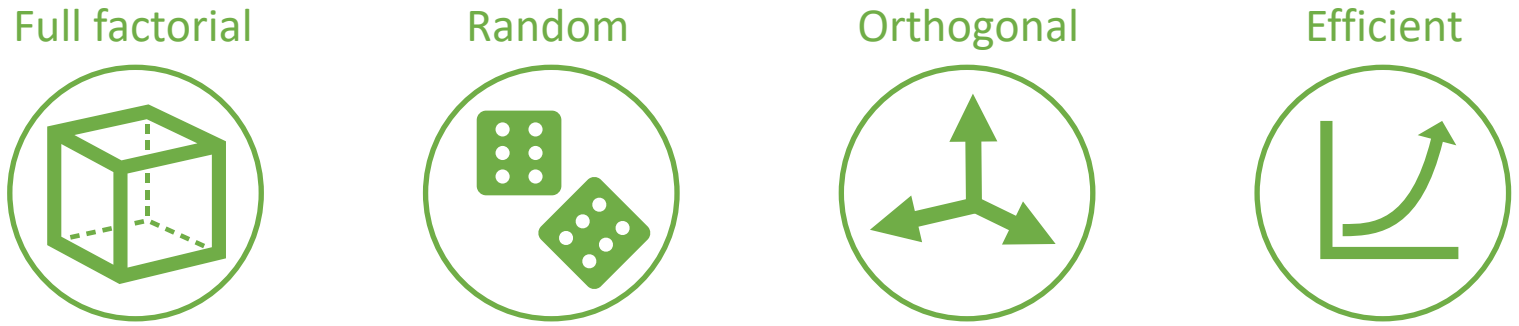
- Subset of choice tasks given to a single decision-maker
- Block size depends on number of choice tasks that decision-makers can handle

Bike	Bus A		Bus B	
dist	$time_A$	$cost_A$	$time_B$	$cost_B$
2	10	1	20	3
2	20	2	30	1
2	30	3	10	2
5	10	2	20	1
5	20	3	30	2
5	30	1	10	3
10	10	3	20	2
10	20	1	30	3
10	30	2	10	1



Step 4 – Block design

Blocking strategy



	Full factorial	Random	Orthogonal	Efficient
❑ Random *	■	■	■	■
❑ Orthogonal			■	
❑ Near-orthogonal	■	■		■

* Often default in survey software

Step 4 – Block design

Orthogonal blocking

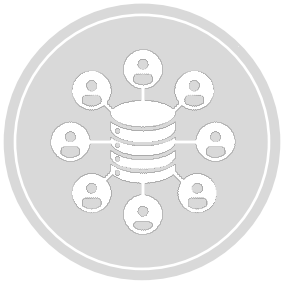
- Attribute level balance within each block

Hotel A			Hotel B		
stars	distance	price	stars	distance	price
1	1	80	3	2	120
3	3	160	5	1	80
5	2	120	1	3	160
1	3	120	3	1	160
3	2	80	5	3	120
5	1	160	1	2	80
1	2	160	3	3	80
5	3	80	1	1	120
3	1	120	5	2	160

Step 5 – Create choice tasks in survey instrument

Key concepts & study plan

Experimental design



Data collection & processing

Model specification & estimation

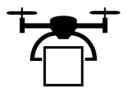
Interpretation & application

Step 5 – Create choice tasks in survey instrument

The easy way

- Import design matrix into dedicated software for choice experiments

www.surveyengine.com

	Drone	Locker	Postie
			
Speed	2 business days	3 business days	5 business days
Delivery method	Leave in a safe place	Secure in locker	Leave at front door
Time window	9am - 5pm (30 minutes)	24/7 (kept for two days)	6pm - 9pm (no choice)
Cost	$2	$6	$8
Which would you choose?	○	○	○

www.qualtrics.com*

Please look over the **benefit packages** below and select the one you like most.

	Package 1	Package 2
Compensation Increases	2.5% per year	Based on promotion
Healthcare	50% of premium	100% of premium
401k Matching	50% matching	100% matching
PTO	10 days	20 days
Work from home	No	1 day a week
Free Lunch	Everyday	Never
	○	○

* Requires purchasing additional conjoint analysis module

Step 5 – Create choice tasks in survey instrument

The hard way

- Create each choice task manually
- May require javacript coding
- May require screenshots of choice tasks

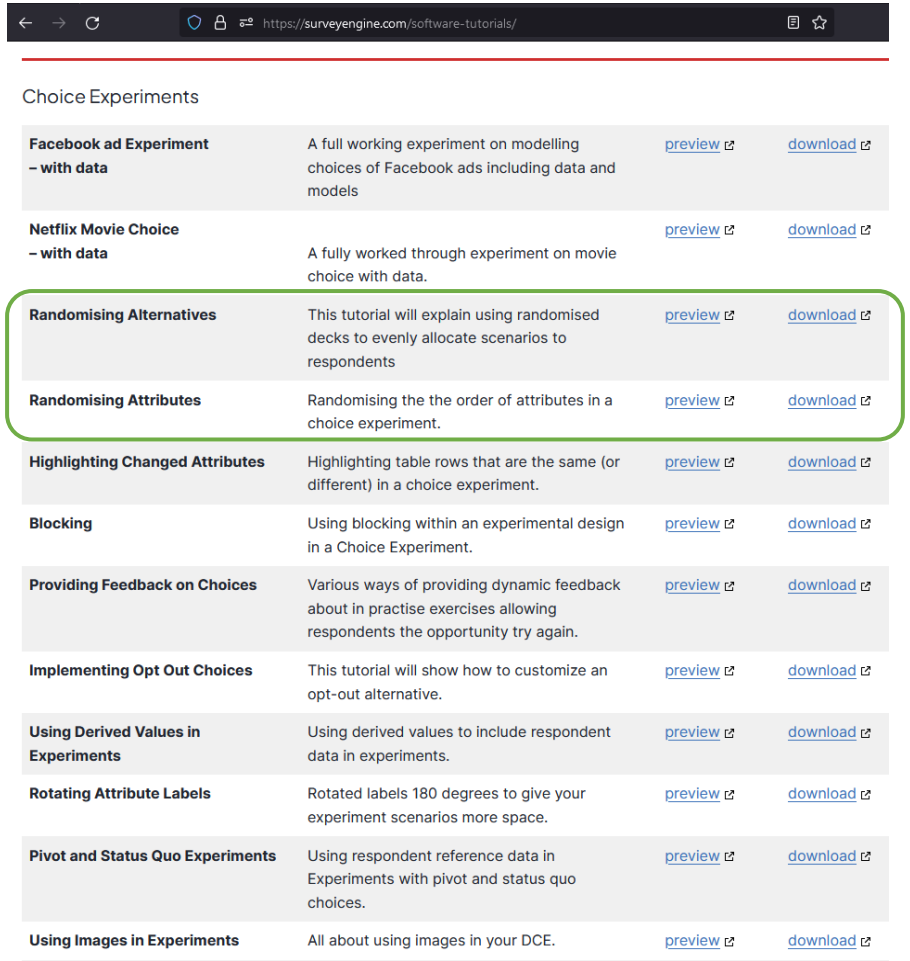
Motorway	Urban road
Speed limit of **90 km/h**, **no traffic lights**.	Speed limit of **50 km/h**, **four traffic lights**.
The travel time is **6 minutes** every day. 6 6 6 6 6	The travel time varies. You will experience one of the following **travel times (in minutes)** with equal probability: 4 6 12 12 12
Toll cost: **$ 1.00**	Toll cost: **$ 0.00**

Step 5 – Create choice tasks in survey instrument

Randomise

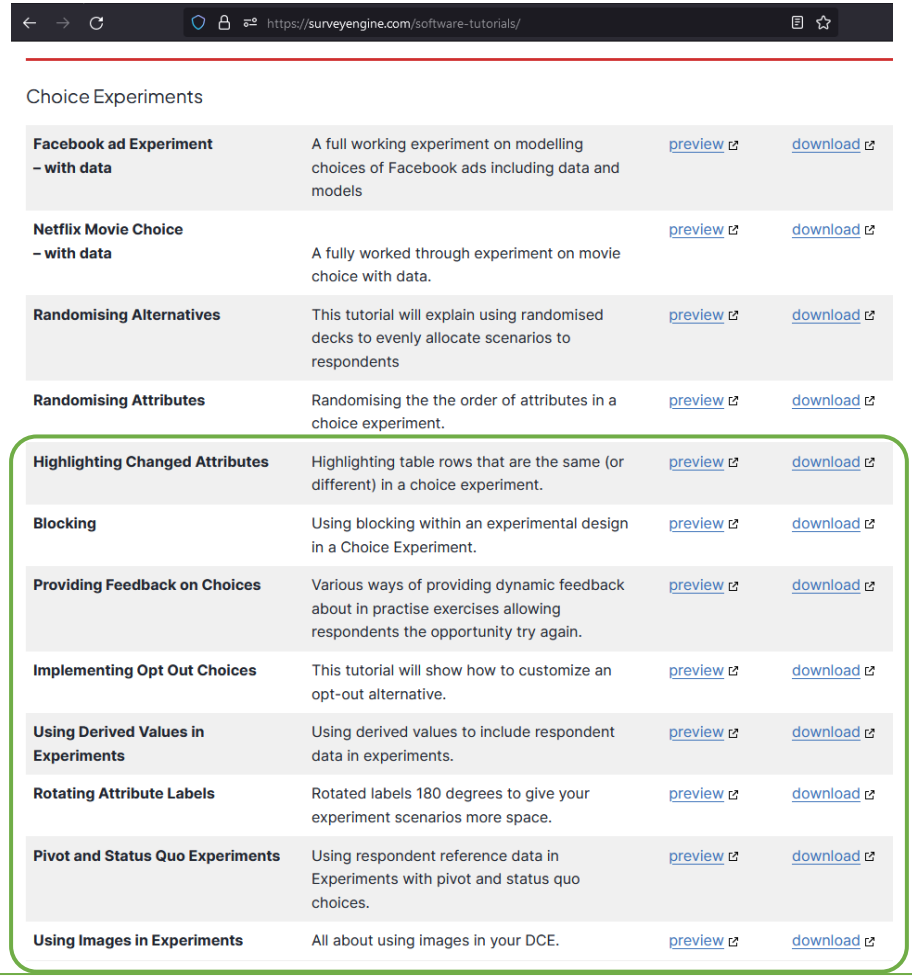
- Order of labelled alternatives across respondents
- Order of attributes across respondents (optional)
- Order of choice tasks across respondents (optional)

- See tutorials on SurveyEngine website



Step 5 – Create choice tasks in survey instrument

Other choice task considerations

- Highlighting attribute level overlap
- Include respondent data in choice tasks
- Status quo and opt-out alternatives
- Using images in experiments
- Etc.

- See tutorials on SurveyEngine website