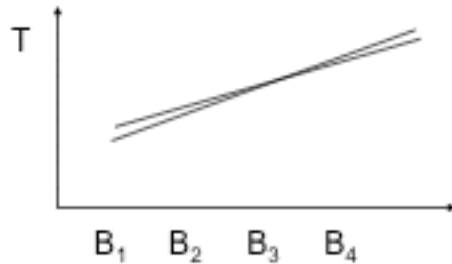
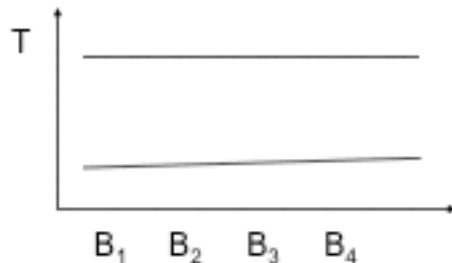


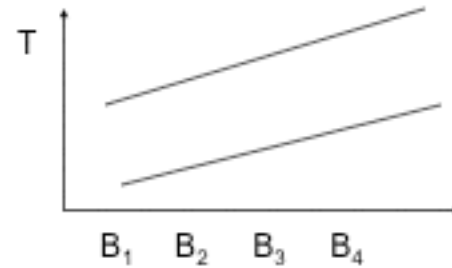
Pas d'effet significatif



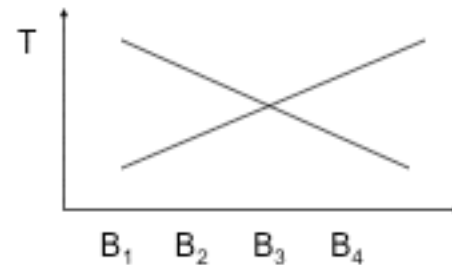
Effet significatif de Bloc. Pas d'autre effet



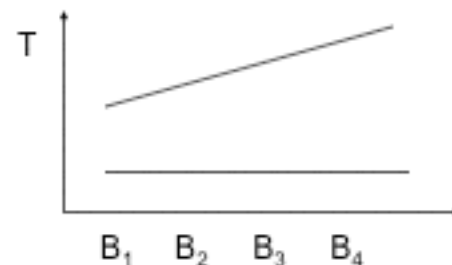
Effet significatif du périphérique.
Pas d'autre effet



Effet significatif des périphériques
et des blocs. Pas d'interaction



Interaction significative. Pas
d'autre effet



Effet significatif des périphériques.
Interaction

NIHM - ÉVALUATION

Si vous avez besoin de...

2

- Comparer différentes solutions technologiques
 - e.g. prouver la supériorité de votre logiciel sur l'existant
- Choisir le meilleur réglage pour un paramètre
 - e.g. la vitesse du curseur dans un jeu vidéo
- Modéliser un phénomène
 - pour le comprendre, le prédire, le causer ou l'éviter
- Comprendre les causes d'un problème, les effets d'un phénomène, etc.

Il vous faut...

3

- Des données
 - Lesquelles ?
 - Comment les obtenir ?
 - Comment les représenter ?
 - Sont-elles utilisables en l'état ?

- Les analyser
 - Comment ?
 - Comparer, trier, modéliser, ...
 - Quel type de différence m'intéresse ?
 - Quel "quantité" de différence m'intéresse ?
 - Quel degré de certitude m'intéresse ?

Nature des données à analyser

4

	Quantitatives Données représentables sous la forme de valeurs, de quantités ou de catégories	Qualitatives Tout ce qui n'est pas quantitatif : descriptions, caractéristiques, propriétés...

Nature des données à analyser

5

	Quantitatives Données représentables sous la forme de valeurs, de quantités ou de catégories	Qualitatives Tout ce qui n'est pas quantitatif : descriptions, caractéristiques, propriétés...
Objectives Données issues de mesures fiables et reproductibles, sans biais possible		
Subjectives Données exprimant une opinion, une perspective, une expérience personnelle		

Nature des données à analyser

6

	Quantitatives Données représentables sous la forme de valeurs, de quantités ou de catégories	Qualitatives Tout ce qui n'est pas quantitatif : descriptions, caractéristiques, propriétés...
Objectives Données issues de mesures fiables et reproductibles, sans biais possible		
Subjectives Données exprimant une opinion, une perspective, une expérience personnelle		<ul style="list-style-type: none">- Interview audio- Question ouverte sur l'opinion d'un utilisateur- Suggestion d'amélioration d'un système...

Nature des données à analyser

7

	Quantitatives Données représentables sous la forme de valeurs, de quantités ou de catégories	Qualitatives Tout ce qui n'est pas quantitatif : descriptions, caractéristiques, propriétés...
Objectives Données issues de mesures fiables et reproductibles, sans biais possible	<ul style="list-style-type: none">- Intervalle de température- Taille- Durée d'exécution- Taux d'erreur...	
Subjectives Données exprimant une opinion, une perspective, une expérience personnelle		<ul style="list-style-type: none">- Interview audio- Question ouverte sur l'opinion d'un utilisateur- Suggestion d'amélioration d'un système...

Nature des données à analyser

8

	Quantitatives Données représentables sous la forme de valeurs, de quantités ou de catégories	Qualitatives Tout ce qui n'est pas quantitatif : descriptions, caractéristiques, propriétés...
Objectives Données issues de mesures fiables et reproductibles, sans biais possible	<ul style="list-style-type: none">- Intervalle de température- Taille- Durée d'exécution- Taux d'erreur...	<ul style="list-style-type: none">- Description factuelle- Fait avéré (format texte)
Subjectives Données exprimant une opinion, une perspective, une expérience personnelle	<ul style="list-style-type: none">- Score d'utilisabilité- Classement de préférence- Oui / Non...	<ul style="list-style-type: none">- Interview audio- Question ouverte sur l'opinion d'un utilisateur- Suggestion d'amélioration d'un système...

Nature des données à analyser

9

	Quantitatives Données représentables sous la forme de valeurs, de quantités ou de catégories	Qualitatives Tout ce qui n'est pas quantitatif : descriptions, caractéristiques, propriétés...
Objectives Données issues de mesures fiables et reproductibles, sans biais possible	<ul style="list-style-type: none">- Expé. contrôlées- Recueil de données "in the wild"etc.	<ul style="list-style-type: none">- Interviews- Questionnaires- Panels d'utilisateurs
Subjectives Données exprimant une opinion, une perspective, une expérience personnelle	<ul style="list-style-type: none">- Questionnaires- Classements- Échelles de Likertetc.	<ul style="list-style-type: none">- Observationsetc.

Nature des données à analyser

10

- Quels types de données ?
 - ▣ Les utilisateurs préfèrent-ils le site web A ou B ?
 - Pourquoi ?
 - ▣ Est-on plus précis au trackpad ou à la souris ?
 - Se **considère-t-on** plus précis au trackpad ou à la souris ?
 - ▣ Dans quelles situations les gens utilisent-ils la lampe de leur téléphone ?
 - **Combien de fois par jour** les gens s'en servent-ils ?
 - ▣ La taille de la cible affecte-t-elle le temps de pointage ?
 - Comment ?

Nature des données à analyser

11

	Quantitatives Données représentables sous la forme de valeurs, de quantités ou de catégories	Qualitatives Tout ce qui n'est pas quantitatif : descriptions, caractéristiques, propriétés...
Objectives Données issues de mesures fiables et reproductibles, sans biais possible	<ul style="list-style-type: none">- Expé. contrôlées- Recueil de données "in the wild"etc.	<ul style="list-style-type: none">- Interviews- Questionnaires- Panels d'utilisateurs
Subjectives Données exprimant une opinion, une perspective, une expérience personnelle	<ul style="list-style-type: none">- Questionnaires- Classements- Échelles de Likertetc.	<ul style="list-style-type: none">- Observationsetc.

Il vous faut...

12

- Des données
 - Lesquelles ?
 - Comment les obtenir ?
 - Comment les représenter ?
 - Sont-elles utilisables en l'état ?

- Les analyser
 - Comment ?
 - Comparer, trier, modéliser, ...
 - Quel type de différence m'intéresse ?
 - Quel "quantité" de différence m'intéresse ?
 - Quel degré de certitude m'intéresse ?

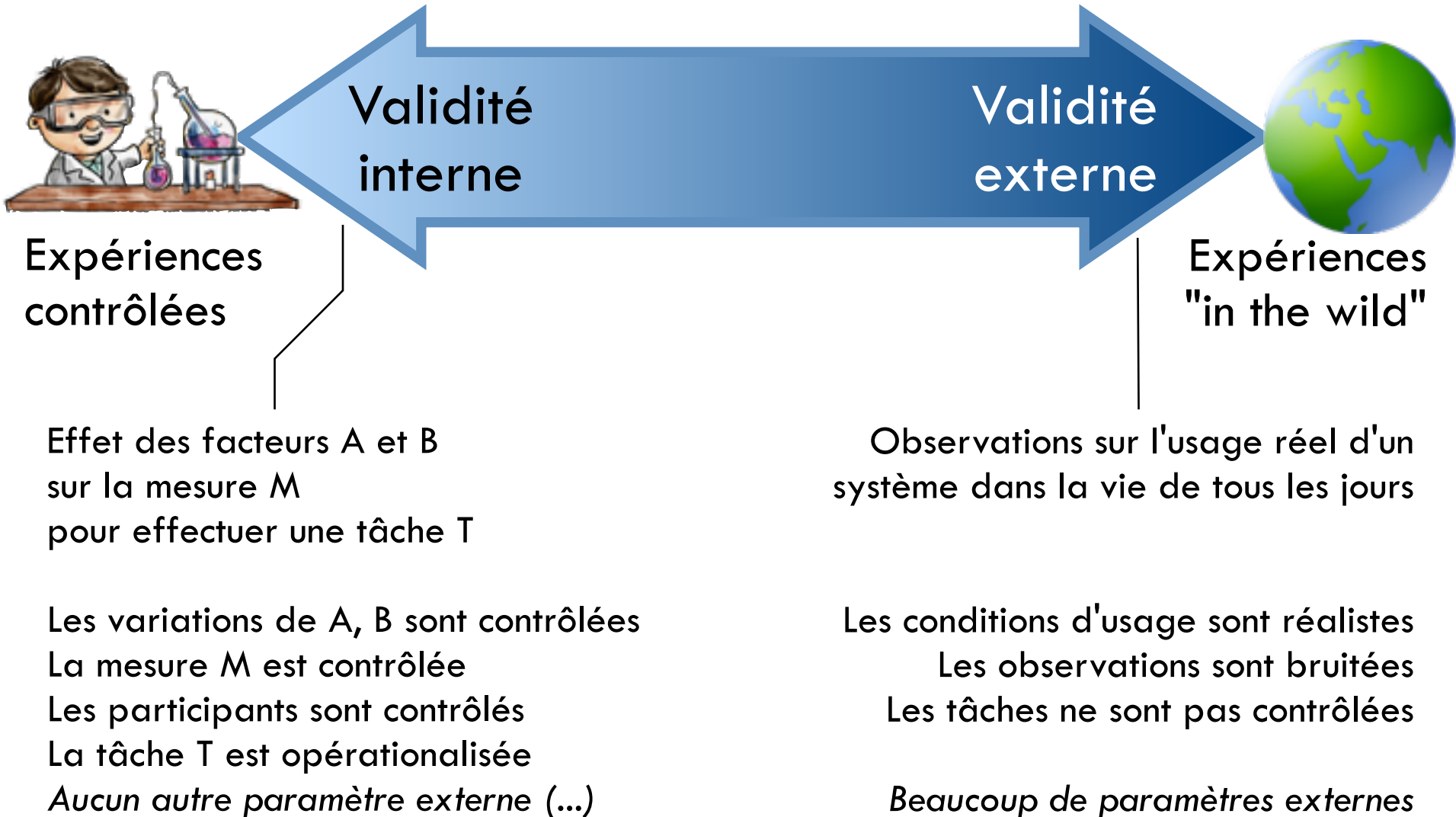
Obtention des données

13



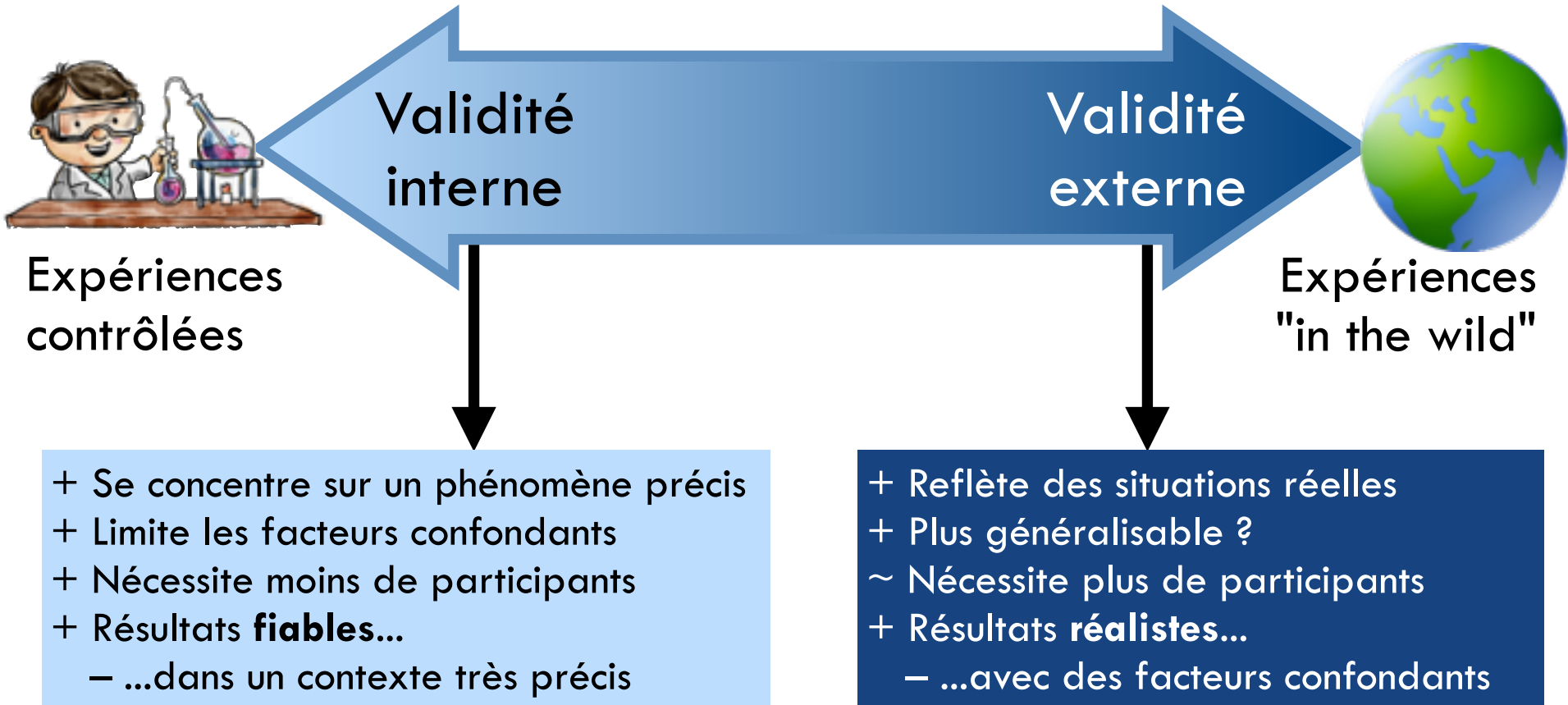
Obtention des données

14



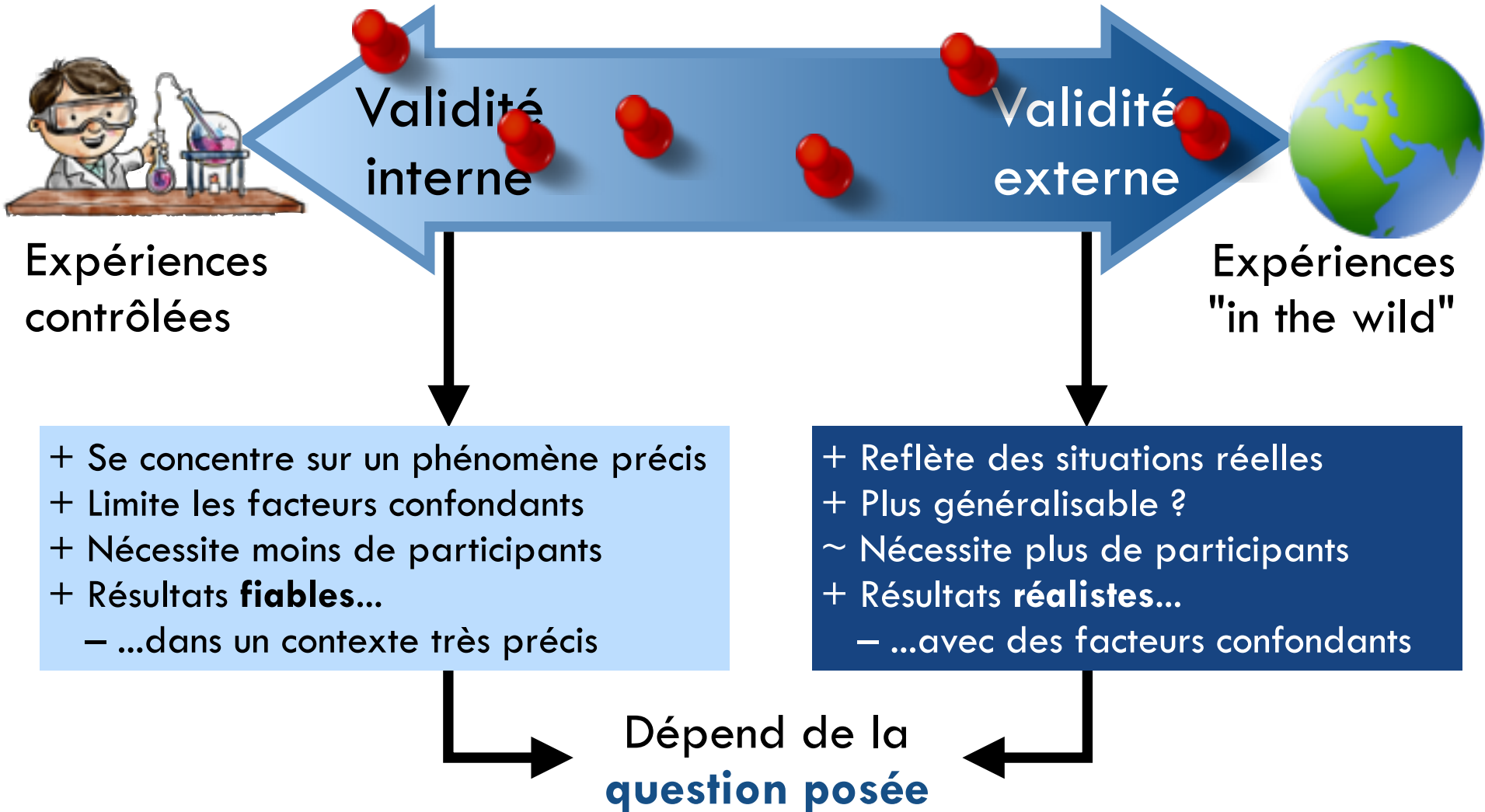
Obtention des données

15



Obtention des données

16



Il vous faut...

17

- Des données
 - Lesquelles ?
 - Comment les obtenir ?
 - Comment les représenter ?
 - Sont-elles utilisables en l'état ?

- Les analyser
 - Comment ?
 - Comparer, trier, modéliser, ...
 - Quel type de différence m'intéresse ?
 - Quel "quantité" de différence m'intéresse ?
 - Quel degré de certitude m'intéresse ?

Préparer une étude

18

Préparation

Nature des données recherchées
Type de données
Type d'étude
Participants recherchés
Facteurs à contrôler...

Recueil des données

Recrutement et planning
Confidentialité & vie privée
Instructions...

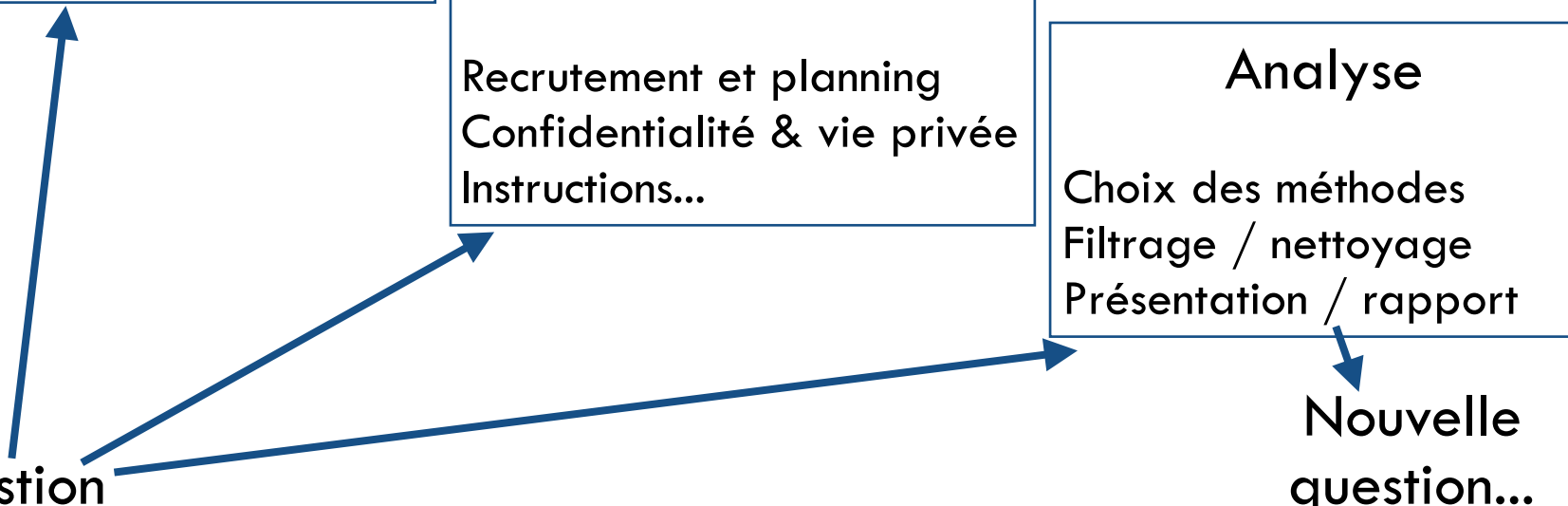
Analyse

Choix des méthodes
Filtrage / nettoyage
Présentation / rapport

Question

Nouvelle
question...

Temps



Préparer une expérience contrôlée

19

Préparation

Nature des données cherchées
Type de données
Type d'étude
Participants recherchés
Facteurs à contrôler

Protocole expérimental :

Paramètre(s)
Mesure(s)
Tâche(s)
Ordonnancement
Entraînement, pauses, ...

Recueil des données

Recrutement et planning
Confidentialité & vie privée
Instructions
Environnement contrôlé
Débrief...

Analyse

Choix des méthodes
Filtrage / nettoyage
Présentation / rapport

Question

Nouvelle
question...

Temps



Nature des données (encore)

20

Préparation

Nature des données

Type de données

Type d'étude

Participants recherchés

Facteurs à contrôler

Protocole expérimental :

Paramètre(s)

Mesure(s)

Tâche(s)

Ordonnancement

Entraînement, pauses, ...

Objectives

Subjectives

Quantitatives

Qualitatives

Question

	Quantitatives	Qualitatives
Objectives		
Subjectives		

Type des données

21

Préparation

Nature des données recherchées

Type de données →

Type d'étude

Participants recherchés

Facteurs à contrôler

Protocole expérimental :

Paramètre(s)

Mesure(s)

Tâche(s)

Ordonnancement

Entraînement, pauses, ...

Quantitatives

Valeurs continues

Entiers : 3 erreurs

Flottants : 1.37 s

Ratios : 41%

etc.

Intervalles : $[1, 4]x/jr$

Séries (e.g. temporelle)

Catégories

Booléens : oui / non

etc.

Qualitatives

Texte

Audio

Vidéo

etc.

Question

Type des données

22

Préparation

Nature des données recherchées

Type de données →

Type d'étude

Participants recherchés

Facteurs à contrôler

Protocole expérimental :

Paramètre(s)

Mesure(s)

Tâche(s)

Ordonnancement

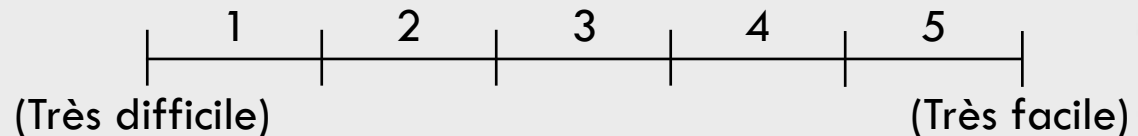
Entraînement, pauses, ...

Question ↑

Quantitatives	Qualitatives
<u>Valeurs continues</u> Entiers : 3 erreurs Flottants : 1.37 s Ratios : 41% etc. Intervalles : [1, 4]x/jr Séries (e.g. temporelle)	Texte Audio Vidéo etc.
<u>Catégories</u> Booléens : oui / non etc.	

! Une valeur peut être numérique et non continue

Vous avez trouvé la tâche :



Non continu : "4" n'est pas deux fois plus grand que "2"

Type d'étude (encore)

23

Préparation

Nature des données recherchées

Type de données

Type d'étude

Participants recherchés

Facteurs à contrôler

Protocole expérimental :

Paramètre(s)

Mesure(s)

Tâche(s)

Ordonnancement

Entraînement, pauses, ...



"Quantité de contrôle"

Validité
interne

Validité
externe



Question

Participants

24

Préparation

Nature des données cherchées

Type de données

Type d'étude

Participants recherchés →

Facteurs à contrôler

Protocole expérimental :

Paramètre(s)

Mesure(s)

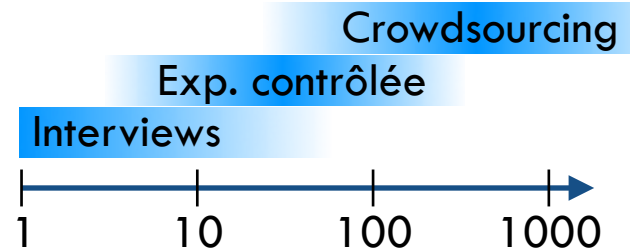
Tâche(s)

Ordonnancement

Entraînement, pauses, ...

↑
Question

□ Nombre



□ Caractéristiques

□ Tranches d'âges

□ Sexe

□ Occupation

□ Expertise (novices, pros, ...)

□ etc.

□ Contrôle

□ Random ?

□ Proportions strictes ?

Facteurs externes

25

Préparation

Nature des données cherchées

Type de données

Type d'étude

Participants recherchés

Facteurs à contrôler →

Protocole expérimental :

Paramètre(s)

Mesure(s)

Tâche(s)

Ordonnancement

Entraînement, pauses, ...

Question ↗

□ Variables externes qui ne doivent **pas** affecter l'étude

□ Équipement

■ Ordinateur

■ Souris

■ Capteur 3D...

□ Environnement

■ Éclairage

■ Bruit ambiant

■ Température...

□ Instructions !

Protocole expérimental

26

Préparation

Nature des données recherchées

Type de données

Type d'étude

Participants recherchés

Facteurs à contrôler

Protocole expérimental :

Paramètre(s)

Mesure(s)

Tâche(s)

Ordonnancement

Entraînement, pauses, ...

Question



□ Un peu de vocabulaire (de stats)

□ Variable Indépendante (VI)

- "Paramètre", "facteur", etc.
- Dont on cherche l'effet
- Qui cause le phénomène étudié
- Peut être continu, discret, qualitatif, ...
- Possiblement plusieurs dans une expé
- **Valeurs possibles = "Niveaux"**

□ Variable Dépendante (VD)

- "Mesure"
- Métrique représentative du phénomène
- (Peut être ce qu'on veut)
- Possiblement plusieurs dans une expé

Protocole expérimental

27

Préparation

Nature des données recherchées

Type de données

Type d'étude

Participants recherchés

Facteurs à contrôler

Protocole expérimental :

Paramètre(s)

Mesure(s)

Tâche(s)

Ordonnancement

Entraînement, pauses, ...

Question



- Dépend de la question posée
 - ▣ Les utilisateurs du site web préfèrent-ils le nouveau site ou l'ancien ?
 - Pourquoi ?
 - ▣ Est-on plus rapide pour détourner une image au trackpad ou à la souris ?
 - Se considère-t-on plus rapide ?
 - ▣ Le niveau d'études et l'âge affectent-ils l'usage des gestes sur mobile ?
 - ▣ Est-on moins attentifs aux notifications selon notre tâche en cours ?
 - ▣ La taille & la distance de la cible affectent-elles le temps de pointage ?
 - Comment ?

Protocole expérimental

28

Préparation

Nature des données cherchées

Type de données

Type d'étude

Participants recherchés

Facteurs à contrôler

Protocole expérimental :

Paramètre(s)

Mesure(s)

Tâche(s)

Ordonnancement

Entraînement, pauses, ...

Question



- Dépend de la question posée
 - Les utilisateurs du site web préfèrent-ils le nouveau site ou l'ancien ?
 - Pourquoi ?
 - Est-on plus rapide pour détourner une image au trackpad ou à la souris ?
 - Se considère-t-on plus rapide ?
 - Le niveau d'études et l'âge affectent-ils l'usage des gestes sur mobile ?
 - Est-on moins attentifs aux notifications selon notre tâche en cours ?
 - La taille & la distance de la cible affectent-elles le temps de pointage ?
 - Comment ?

Protocole expérimental

29

Préparation

Nature des données recherchées

Type de données

Type d'étude

Participants recherchés

Facteurs à contrôler

Protocole expérimental :

Paramètre(s)

Mesure(s)

Tâche(s)

Ordonnancement

Entraînement, pauses, ...

Question



- Dépend de la question posée
 - Les utilisateurs du site web préfèrent-ils le nouveau site ou l'ancien ?
 - Pourquoi ?
 - Est-on plus rapide pour détourner une image au trackpad ou à la souris ?
 - Se considère-t-on plus rapide ?
 - Le niveau d'études et l'âge affectent-ils l'usage des gestes sur mobile ?
 - Est-on moins attentifs aux notifications selon notre tâche en cours ?
 - La taille & la distance de la cible affectent-elles le temps de pointage ?
 - Comment ?

Protocole expérimental

30

Préparation

Nature des données recherchées

Type de données

Type d'étude

Participants recherchés

Facteurs à contrôler

Protocole expérimental :

Paramètre(s)

Mesure(s)

Tâche(s)

Ordonnancement

Entraînement, pauses, ...

Question



- Dépend de la question posée
 - Les utilisateurs du site web préfèrent-ils le nouveau site ou l'ancien ?
 - Pourquoi ?
 - Est-on plus rapide pour détourner une image au trackpad ou à la souris ?
 - Se considère-t-on plus rapide ?
 - Le niveau d'études et l'âge affectent-ils l'usage des gestes sur mobile ?
 - Est-on moins attentifs aux notifications selon notre tâche en cours ?
 - La taille & la distance de la cible affectent-elles le temps de pointage ?
 - Comment ?

Protocole expérimental

31

Préparation

Nature des données recherchées

Type de données

Type d'étude

Participants recherchés

Facteurs à contrôler

Protocole expérimental :

Paramètre(s)

Mesure(s)

Tâche(s)

Ordonnancement

Entraînement, pauses, ...

Question



- Dépend de la question posée
 - Les utilisateurs du site web préfèrent-ils le nouveau site ou l'ancien ?
 - Pourquoi ?
 - Est-on plus rapide pour détourner une image au trackpad ou à la souris ?
 - Se considère-t-on plus rapide ?
 - Le niveau d'études et l'âge affectent-ils l'usage des gestes sur mobile ?
 - Est-on moins attentifs aux notifications selon notre tâche en cours ?
 - La taille & la distance de la cible affectent-elles le temps de pointage ?
 - Comment ?

Protocole expérimental

32

Préparation

Nature des données recherchées

Type de données

Type d'étude

Participants recherchés

Facteurs à contrôler

Protocole expérimental :

Paramètre(s)

Mesure(s)

Tâche(s)

Ordonnancement

Entraînement, pauses, ...

Question



□ Dépend de la question posée

- Les utilisateurs du site web préfèrent-ils le nouveau site ou l'ancien ?
 - Pourquoi ?
- Est-on plus rapide pour détourner une image au trackpad ou à la souris ?
 - Se considère-t-on plus rapide ?
- Le niveau d'études et l'âge affectent-ils l'usage des gestes sur mobile ?
- Est-on moins attentifs aux notifications selon notre tâche en cours ?
- La taille & la distance de la cible affectent-elles le temps de pointage ?
 - Comment ?

Protocole expérimental

33

Préparation

Nature des données recherchées

Type de données

Type d'étude

Participants recherchés

Facteurs à contrôler

Protocole expérimental :

Paramètre(s)

Mesure(s)

Tâche(s)

Ordonnancement

Entraînement, pauses, ...

- Assignation des participants
- Ordre des paramètres & tâches
- Ordre des mesures

Question



Protocole expérimental

34

Préparation

Nature des données recherchées

Type de données

Type d'étude

Participants recherchés

Facteurs à contrôler

Protocole expérimental :

Paramètre(s)

Mesure(s)

Tâche(s)

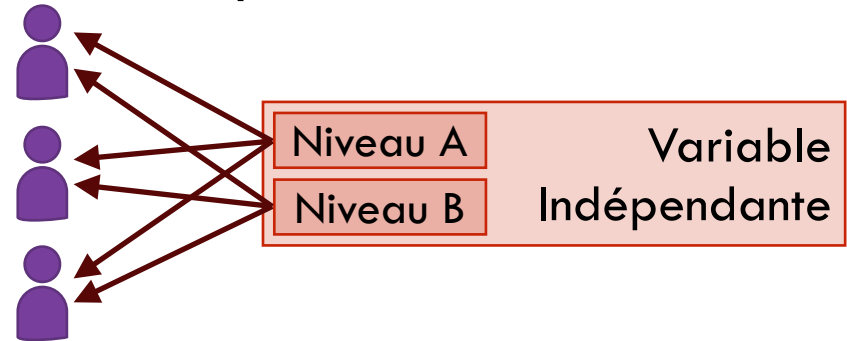
Ordonnancement

Entraînement, pauses, ...

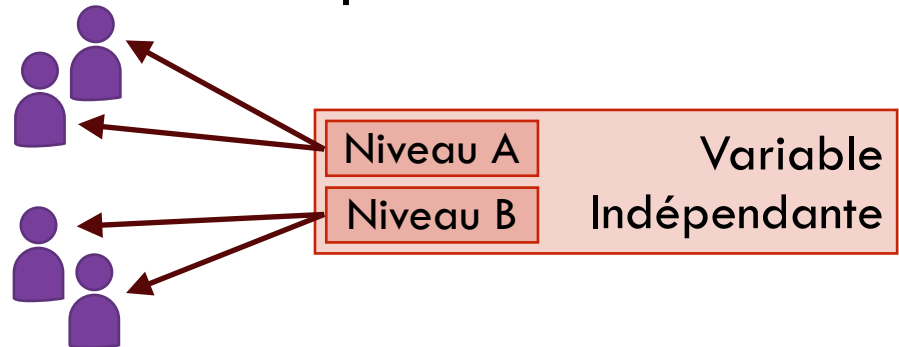
Question

□ Assignation des participants

□ "Within subject"



□ "Between subject"



□ Plans d'expérience mixtes

Protocole expérimental

35

Préparation

Nature des données recherchées

Type de données

Type d'étude

Participants recherchés

Facteurs à contrôler

Protocole expérimental :

Paramètre(s)

Mesure(s)

Tâche(s)

Ordonnancement

Entraînement, pauses, ...

Question



□ Assignation des participants

□ "Within subject"

- + Comparaisons plus faciles
- + Moins de participants
- – Expés plus longues

□ "Between subject"

- + Expés plus courtes
- + Comparer des catégories d'utilisateur
- – Variabilité entre groupes
- – Plus de participants

□ Plans d'expérience mixtes

Protocole expérimental

36

Préparation

Nature des données recherchées

Type de données

Type d'étude

Participants recherchés

Facteurs à contrôler

Protocole expérimental :

Paramètre(s)

Mesure(s)

Tâche(s)

Ordonnancement

Entraînement, pauses, ...

- Assignation des participants
- Ordre des paramètres & tâches
 - ▣ "Order effect"
 - Est-ce que la question impose un ordre ?
 - Ordre fixe

Contre une
technique
"classique"

Technique
de base

Nouvelle
technique

Expérience
d'apprentissage

Performances
de base

Technique
d'entraînement

Performances
altérées ?

Question

Protocole expérimental

37

Préparation

Nature des données recherchées

Type de données

Type d'étude

Participants recherchés

Facteurs à contrôler

Protocole expérimental :

Paramètre(s)

Mesure(s)

Tâche(s)

Ordonnancement

Entraînement, pauses, ...

- Assignation des participants
- Ordre des paramètres & tâches
 - ▣ "Order effect"
 - Est-ce que la question impose un ordre ?
 - Ordre fixe
 - Est-ce qu'une condition risque d'affecter la suivante ?
 - Tous les possibles : $N = (\text{nb conditions})!$
 - Carré Latin

Carré latin 4 X 4

A	B	C	D
B	C	D	A
C	D	A	B
D	A	B	C

Question



Protocole expérimental

38

Préparation

Nature des données cherchées

Type de données

Type d'étude

Participants recherchés

Facteurs à contrôler

Protocole expérimental :

Paramètre(s)

Mesure(s)

Tâche(s)

Ordonnancement

Entraînement, pauses, ...

- Assignation des participants
- Ordre des paramètres & tâches
 - ▣ "Order effect"
 - Est-ce que la question impose un ordre ?
 - Ordre fixe
 - Est-ce qu'une condition risque d'affecter la suivante ?
 - Tous les possibles : $N = (\text{nb conditions})!$
 - Carré Latin

Carré latin 4 X 4

A	B	C	D
B	C	D	A
C	D	A	B
D	A	B	C

Question

Protocole expérimental

39

Préparation

Nature des données recherchées

Type de données

Type d'étude

Participants recherchés

Facteurs à contrôler

Protocole expérimental :

Paramètre(s)

Mesure(s)

Tâche(s)

Ordonnancement

Entraînement, pauses, ...

- Assignation des participants
- Ordre des paramètres & tâches
 - ▣ "Order effect"
 - Est-ce que la question impose un ordre ?
 - Ordre fixe
 - Est-ce qu'une condition risque d'affecter la suivante ?
 - Tous les possibles : $N = (\text{nb conditions})!$
 - Carré Latin

Carré latin 4 X 4

A	B	C	D
B	C	D	A
C	D	A	B
D	A	B	C

Carré latin 4 X 4 balancé

A	B	C	D
B	D	A	C
D	C	B	A
C	A	D	B

Question

Protocole expérimental

40

Préparation

Nature des données cherchées

Type de données

Type d'étude

Participants recherchés

Facteurs à contrôler

Protocole expérimental :

Paramètre(s)

Mesure(s)

Tâche(s)

Ordonnancement

Entraînement, pauses, ...

- Assignation des participants
- Ordre des paramètres & tâches
 - ▣ "Order effect"
 - Est-ce que la question impose un ordre ?
 - Ordre fixe
 - Est-ce qu'une condition risque d'affecter la suivante ?
 - Tous les possibles : $N = (\text{nb conditions})!$
 - Carré Latin
 - Ordre aléatoire
 - ▣ Répétitions

Protocole expérimental

41

Préparation

Nature des données cherchées

Type de données

Type d'étude

Participants recherchés

Facteurs à contrôler

Protocole expérimental :

Paramètre(s)

Mesure(s)

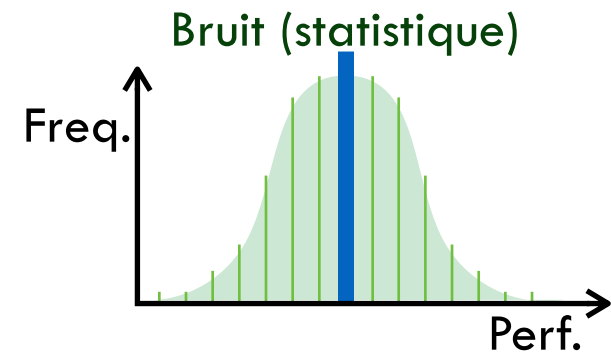
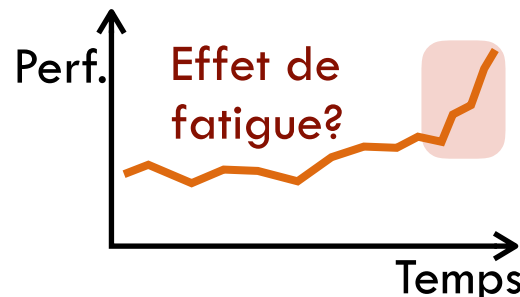
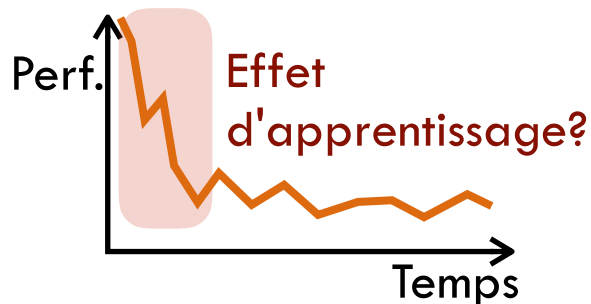
Tâche(s)

Ordonnancement

Entraînement, pauses, ...

- Assignation des participants
- Ordre des paramètres & tâches
 - "Order effect"
 - Est-ce que la question impose un ordre ?
 - Ordre fixe
 - Est-ce qu'une condition risque d'affecter la suivante ?
 - Tous les possibles : $N = (\text{nb conditions})!$
 - Carré Latin
 - Ordre aléatoire

□ Répétitions



Protocole expérimental

42

Préparation

Nature des données recherchées

Type de données

Type d'étude

Participants recherchés

Facteurs à contrôler

Protocole expérimental :

Paramètre(s)

Mesure(s)

Tâche(s)

Ordonnancement

Entraînement, pauses, ...

- Assignation des participants
- Ordre des paramètres & tâches
- Ordre des mesures

Typiquement :

- ▣ Pré-questionnaire
 - Démographie etc.
- ▣ Tests
 - Suivant l'ordonnancement établi
- ▣ Post-questionnaire
 - Opinions, suggestions, commentaires, questions ouvertes, etc.

Question



Préparer une expérience contrôlée

43

Préparation

Nature des données cherchées
Type de données
Type d'étude
Participants recherchés
Facteurs à contrôler

Protocole expérimental :

Paramètre(s)
Mesure(s)
Tâche(s)
Ordonnancement
Entraînement, pauses, ...

Recueil des données

Recrutement et planning
Confidentialité & vie privée
Instructions
Environnement contrôlé
Débrief...

Analyse

Choix des méthodes
Filtrage / nettoyage
Présentation / rapport

Question

Nouvelle question...

Temps



Statistiques

Quelques tests...

45

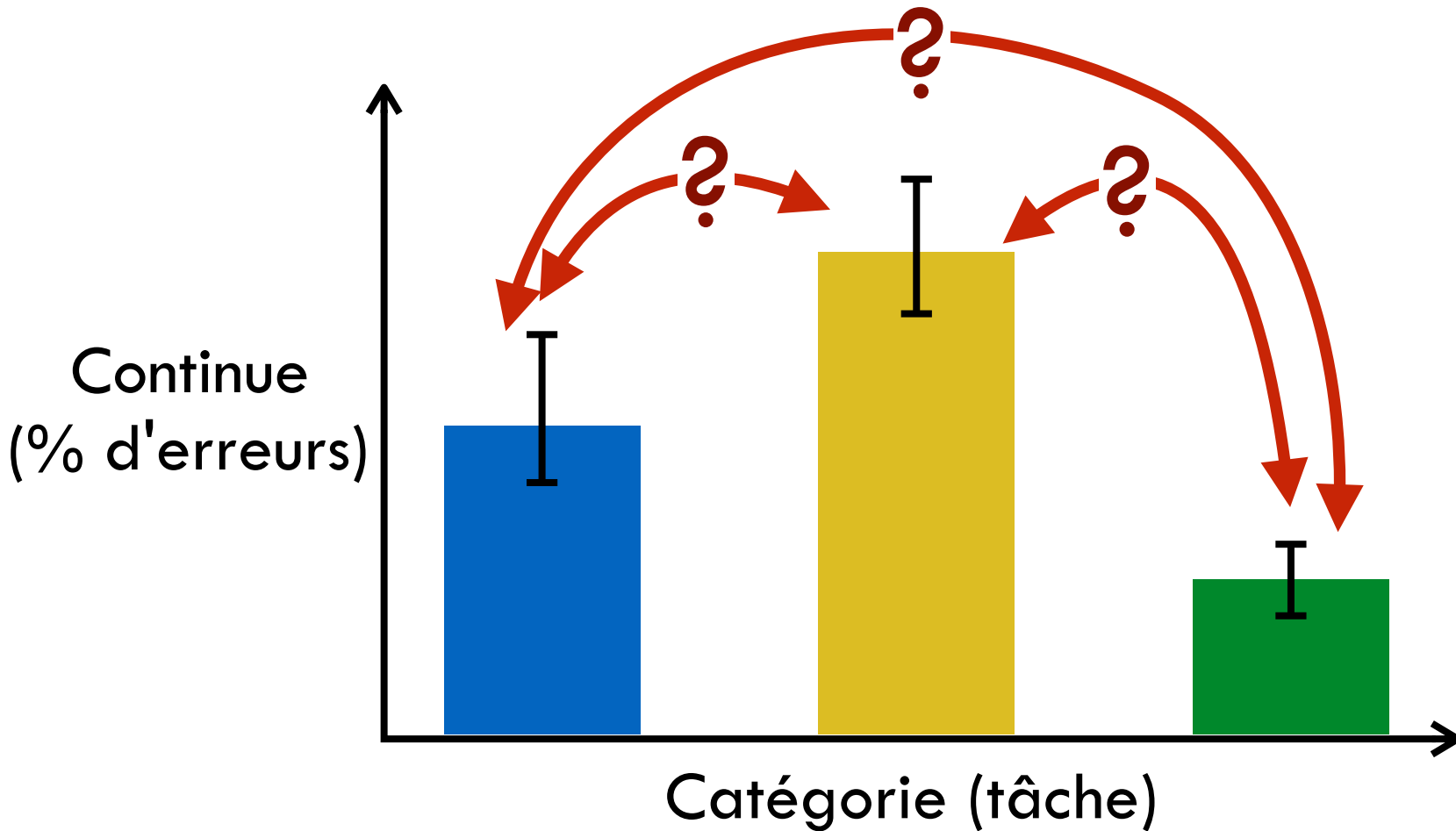
Nb Variables Dépendantes	Nature des Variables Indépendantes	Nature des Variables Dépendantes	Test(s)
1	1 VI avec 2 niveaux ou plus	continues & normales	ANOVA
		ordinales ou continues	Kruskal Wallis Friedmann Test
		catégories	Chi-square test
	2 VIs ou plus	continues & normales	factorial ANOVA
		ordinales ou continues	ordered logistic regression
		catégories (2)	factorial logistic regression
	1 VI continue	continues & normales	correlation
		continues & normales	simple linear regression
		ordinales ou continues	non-parametric correlation
		catégories	simple logistic regression
	1 ou plus VI continue et/ou 1 ou plus VI catégorielle	continues & normales	multiple regression analysis of covariance
		catégories	multiple logistic regression discriminant analysis
2+	1 VI avec 2 niveaux ou plus	continues & normales	one-way MANOVA

Source (simplifiée) : <https://stats.idre.ucla.edu/other/mult-pkg/whatstat/>

Quelques tests...

46

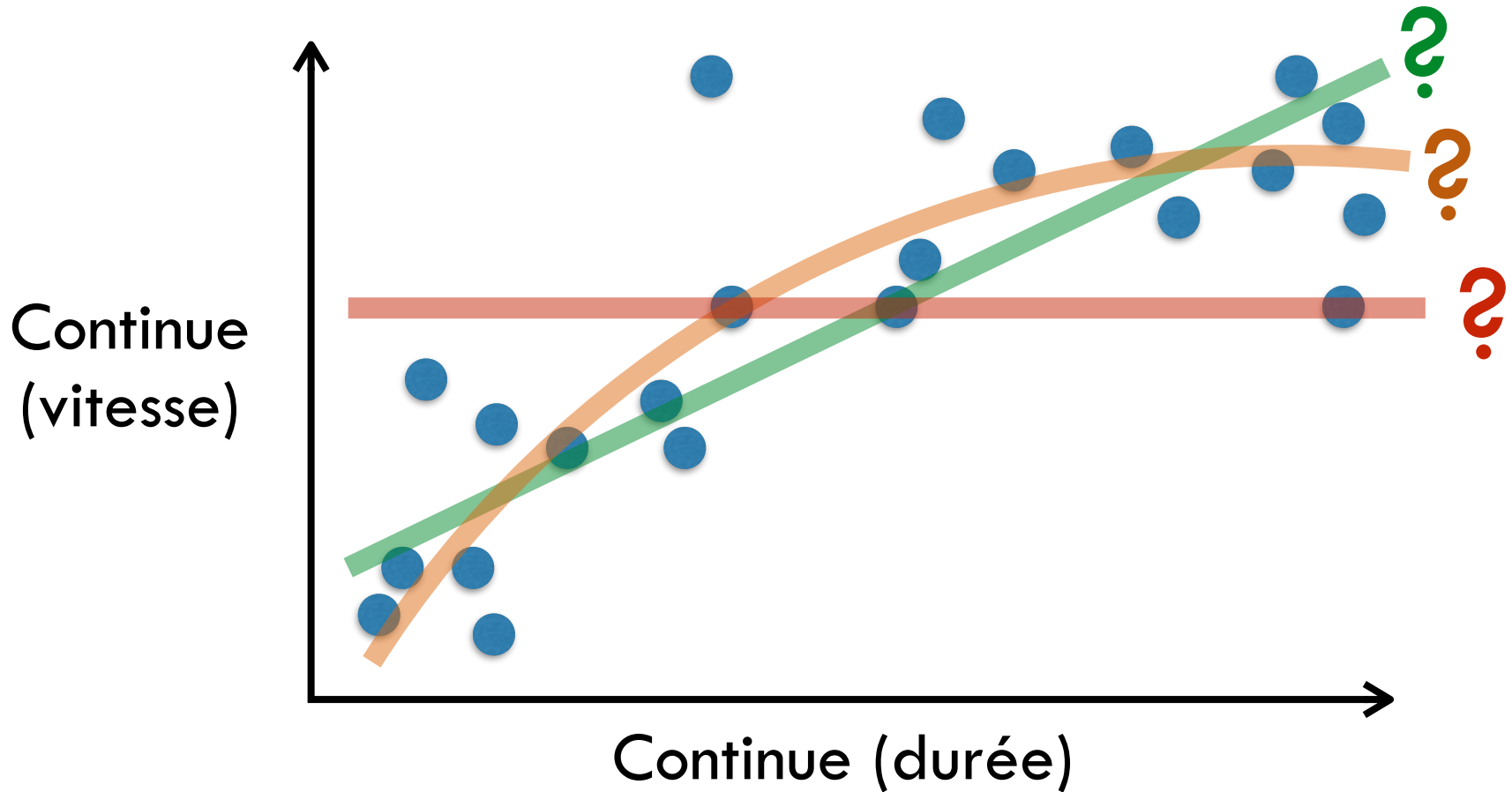
- Cas général : catégorie \rightarrow mesure continue
 - "Le taux d'erreur est-il différent selon la tâche ?"



Quelques tests...

47

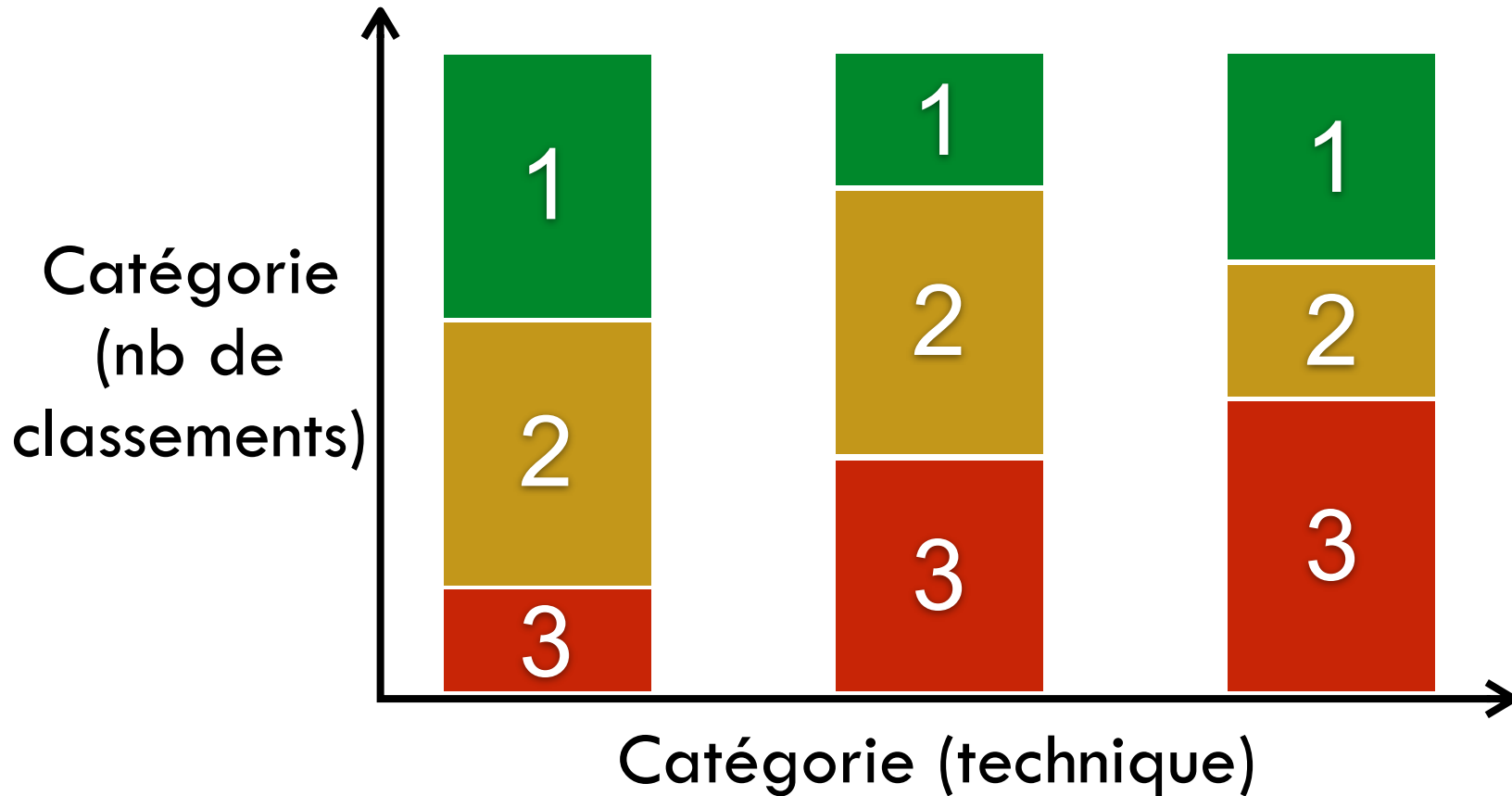
- Facteur continu → mesure continue
 - "La durée d'entrainement est-elle corrélée à la vitesse ?"



Quelques tests...

48

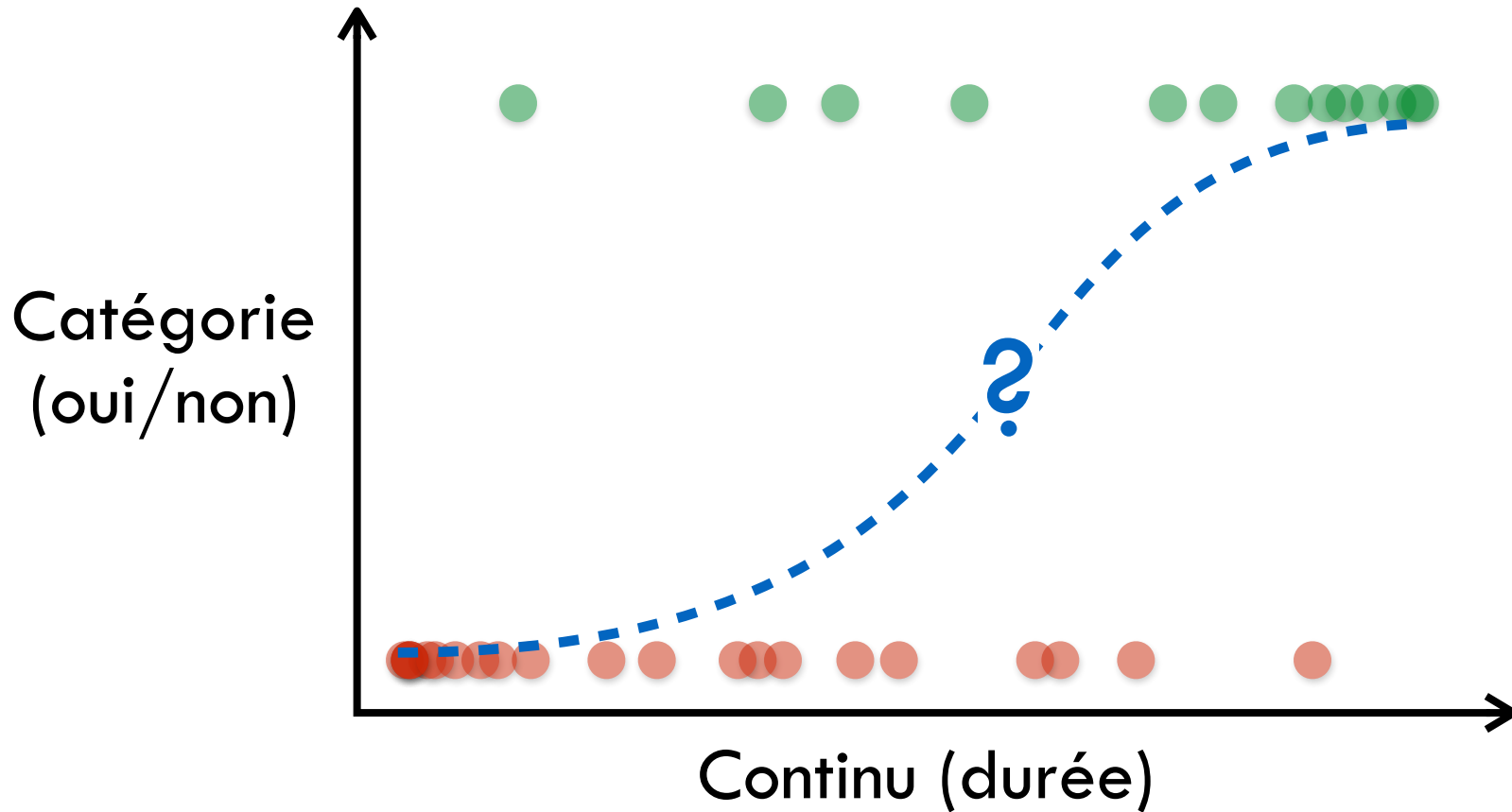
- Catégorie → mesure catégorielle
 - "Quelle technique a été préférée ?"



Quelques tests...

49

- Facteur continu → mesure catégorielle
 - "La durée de révision affecte-t-elle le succès à un exam ?"



Un exemple rapide

50

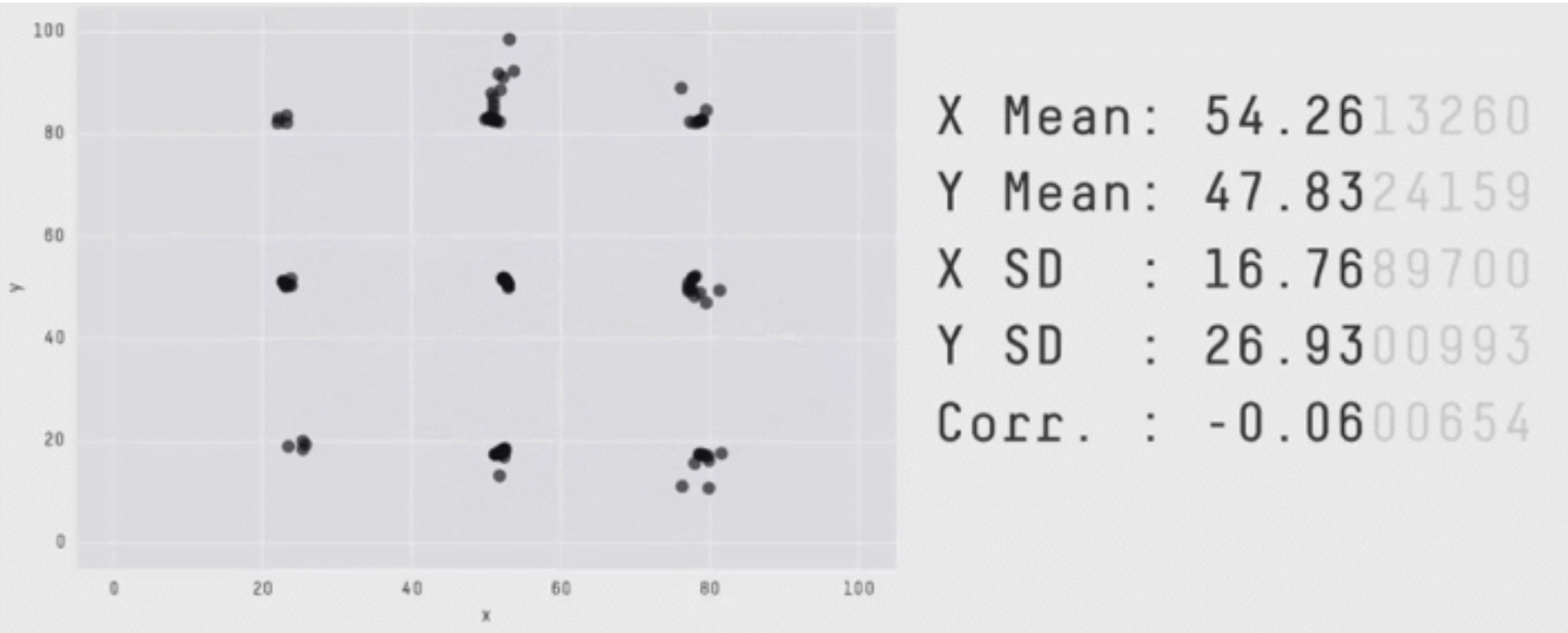
- On veut savoir si la sélection dans un menu s'effectue plus efficacement à la souris ou au trackpad.
 - ▣ 1 variable indépendante : {Souris, Trackpad} (catégorie)
 - ▣ 2 variables dépendantes : {Temps, Précision} (continues)
 - ▣ 1 tâche : sélection d'items dans un menu

- On demande à 50 personnes d'effectuer cette tâche de manière répétée :
 - ▣ 10 fois avec une souris
 - ▣ 10 fois avec un trackpad

Un exemple rapide

51

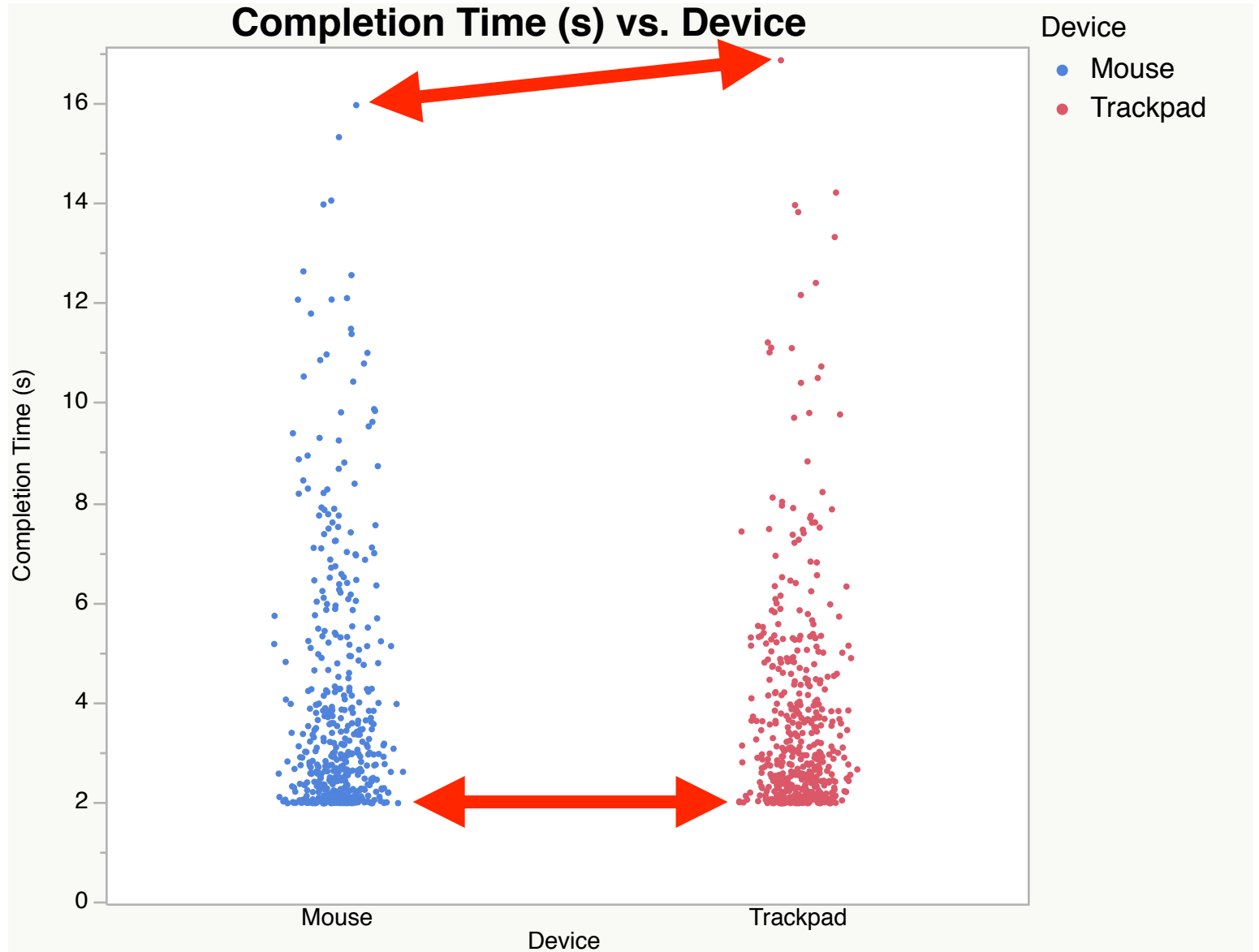
- Toujours visualiser ses données !



Justin Matejka and George Fitzmaurice. 2017. Same Stats, Different Graphs: Generating Datasets with Varied Appearance and Identical Statistics through Simulated Annealing. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 1290-1294. DOI: <https://doi.org/10.1145/3025453.3025912>

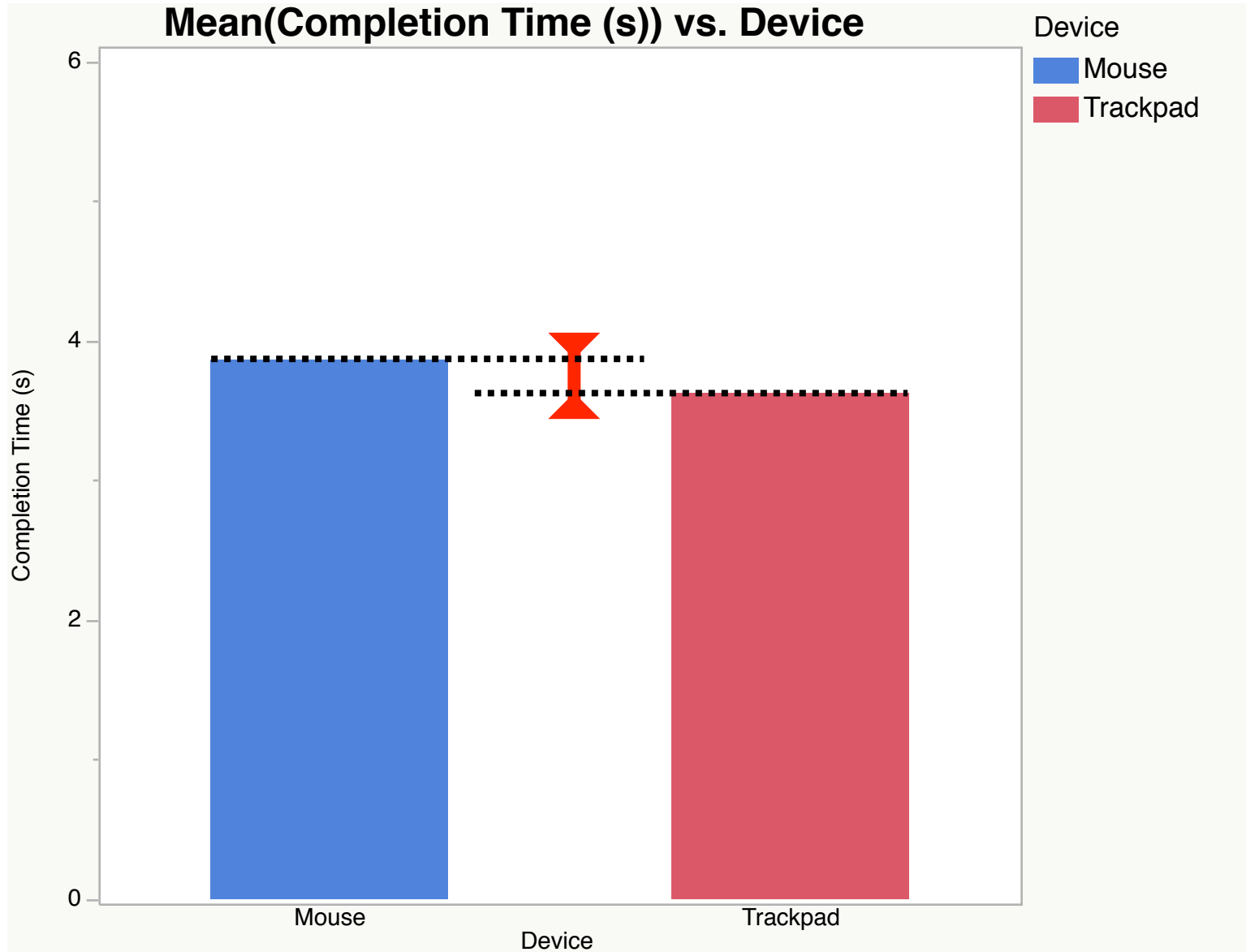
Un exemple rapide

52



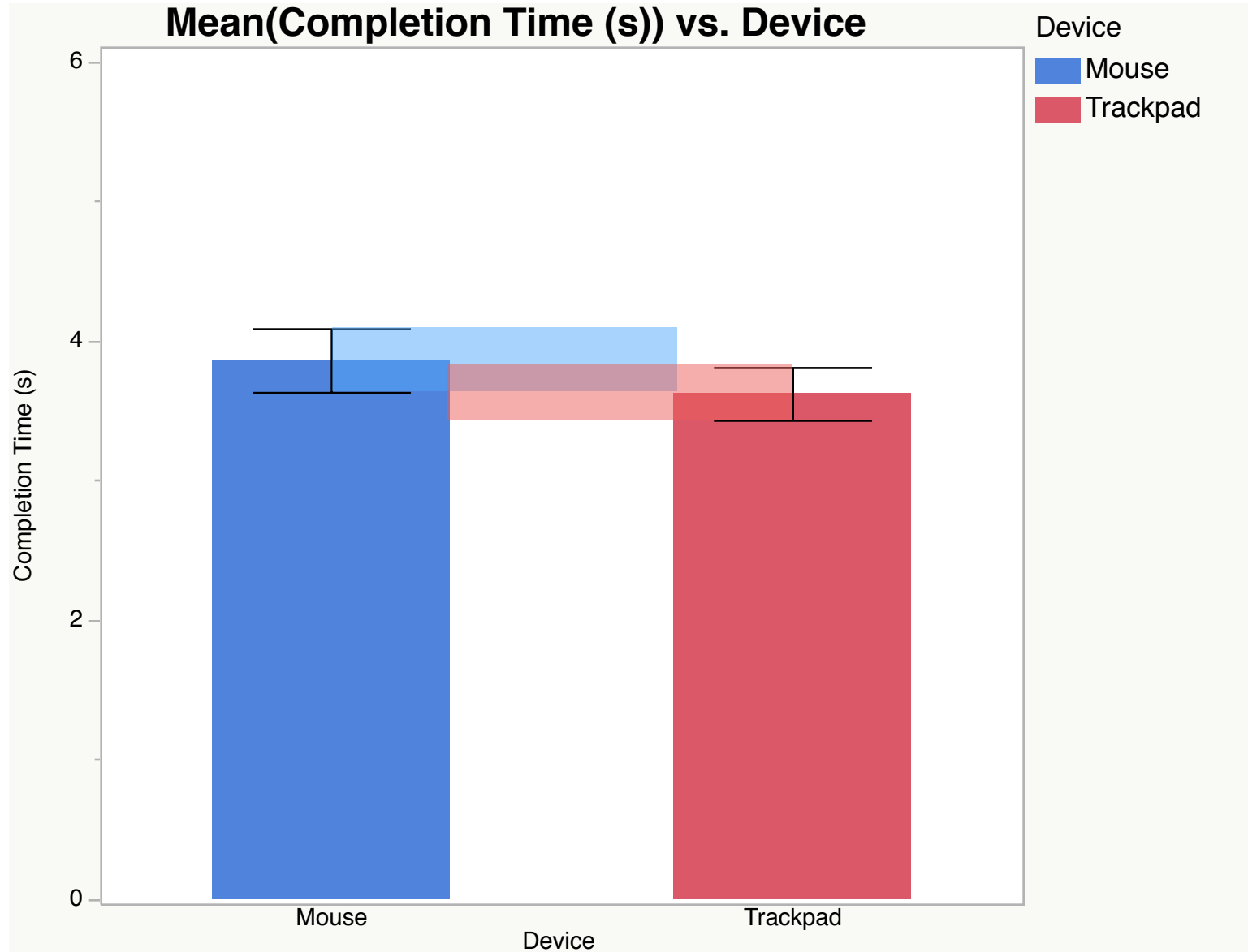
Un exemple rapide

53



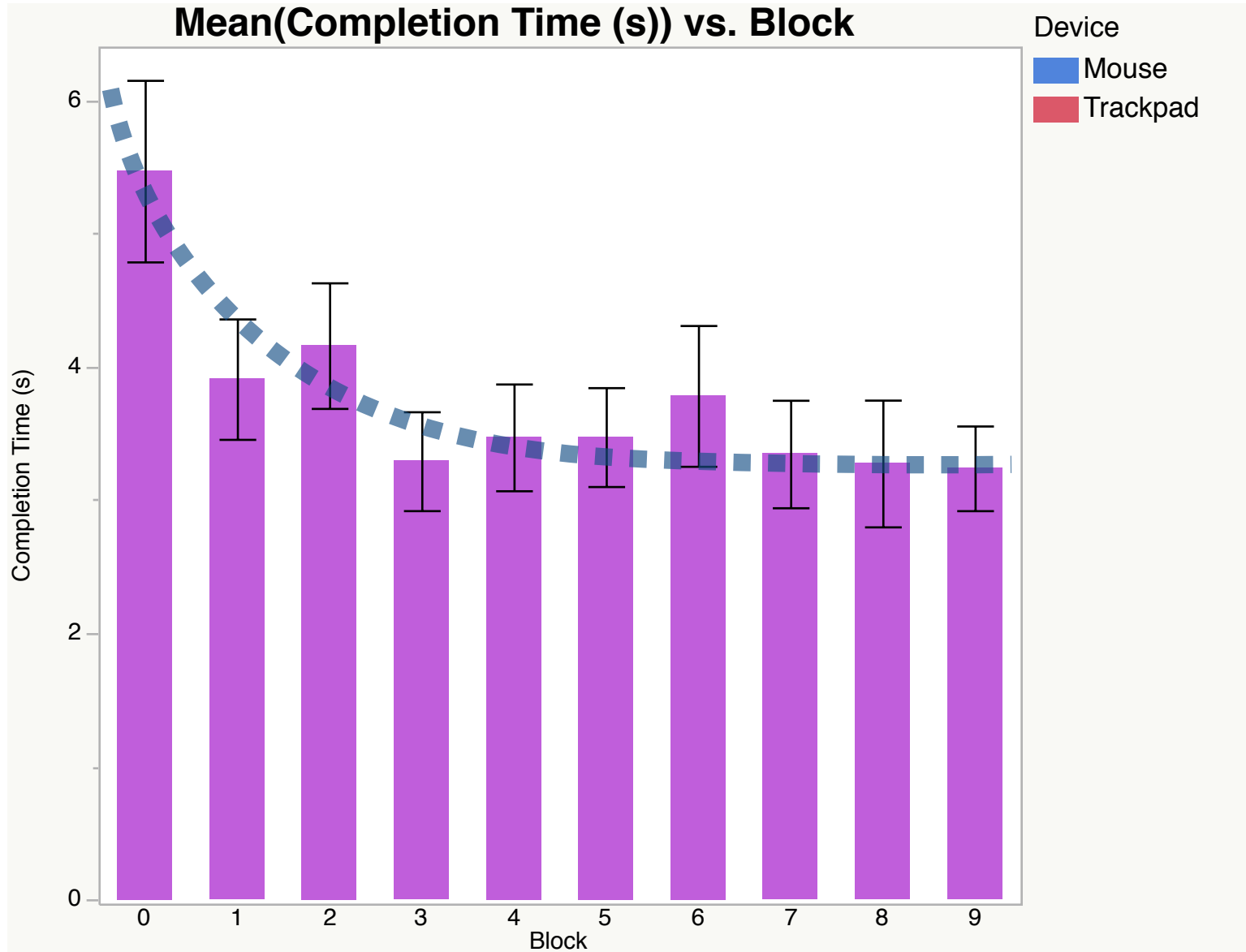
Un exemple rapide

54



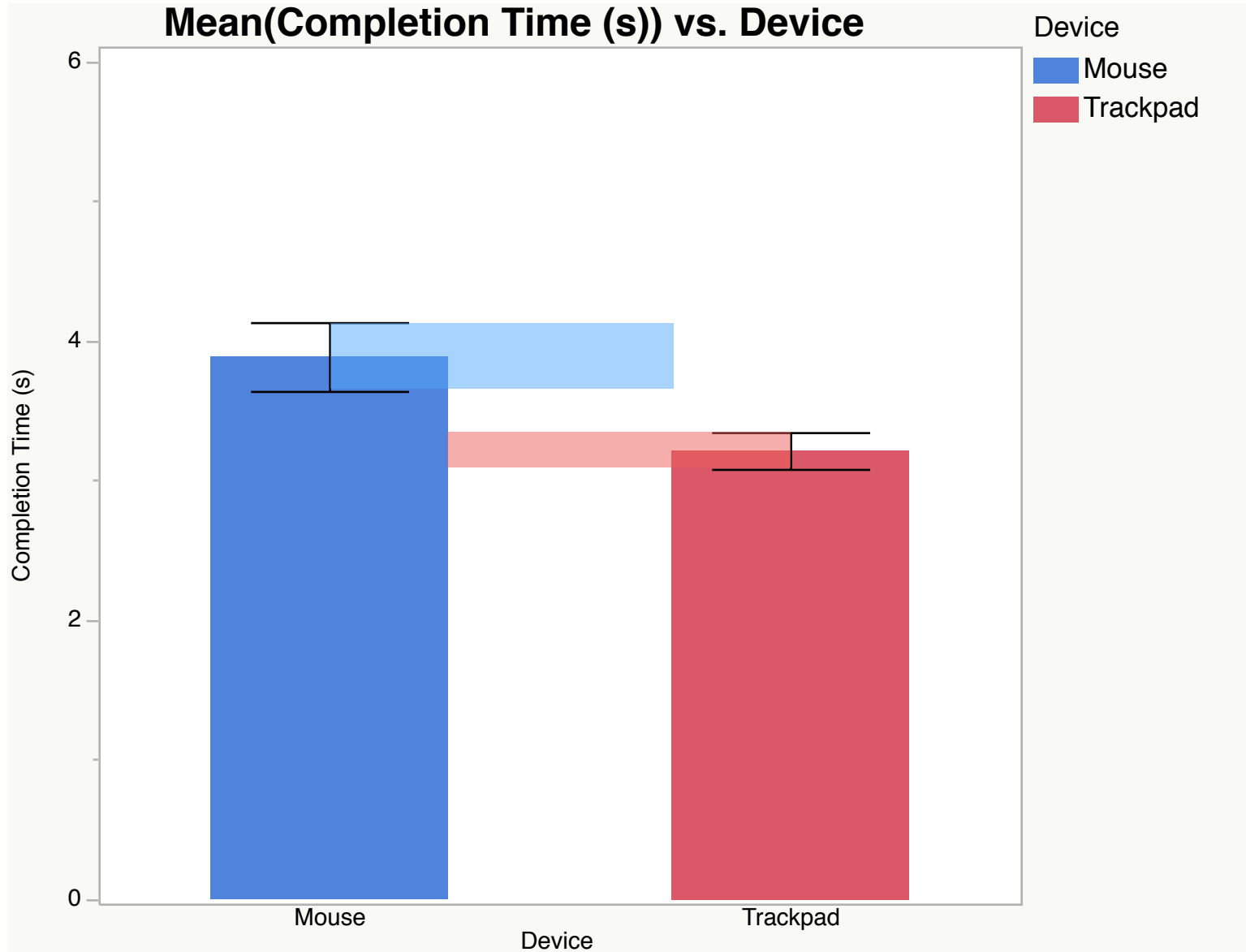
Un exemple rapide

55



Un exemple rapide

56



Dans tous les cas :

57

- Observe-t-on un effet, un phénomène ?
 - ▣ Test statistique et/ou interprétation humaine

- Quel confiance a-t-on dans l'effet observé ?
 - ▣ L'observation est-elle due au hasard ou au traitement ?
 - ▣ Tests statistiques

- Quelle est l'amplitude de l'effet observé ?
 - ▣ Est-elle suffisante ? Négligeable ?
 - ▣ Interprétation humaine

Un outil courant en IHM : ANOVA

58

- ANalysis Of VAriance
- Variable indépendante catégorielle
 - ▣ Technique, traitement, dispositif, ...
- Variable dépendante continue
 - ▣ Vitesse d'exécution, nombre d'erreur, mémorisation, ...
- Teste si des groupes de valeurs (populations) sont significativement distinctes les unes des autres.
- On s'en sert pour tester l'existence d'un effet significatif des différents niveaux de la variable indépendante sur les valeurs de la variable dépendante.

Un peu de stats (pour la culture)

59

- ANOVA ∈ NHST (Null Hypothesis Significance Tests)
 - Pas un test de différence (il y en aura toujours une)
 - Teste l'égalité entre les moyennes de mesures continues
 - But : réfuter cette égalité

- Hypothèse nulle :
 - Mon hypothèse H :
Les niveaux de F (f_1, f_2, f_3, \dots) affectent la valeur de M
 - Hypothèse nulle correspondante H_0 :
Les valeurs de M restent constantes quand F varie

Un peu de stats (pour la culture)

60

□ "Effet significatif" : Rejet de l'hypothèse nulle

□ Avec un certain critère de tolérance α

p = "Probabilité que la différence observée soit due à la chance, si on considère que l'hypothèse nulle est vraie"

α = "La probabilité en-dessous de laquelle je considère que ce n'est plus de la chance"

$p < \alpha$: L'hypothèse nulle est rejetée

→ Il y a donc un effet significatif

Un peu de stats (pour la culture)

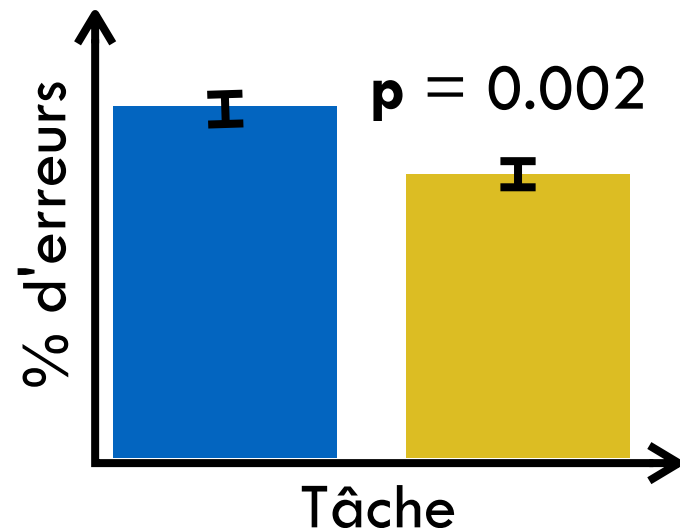
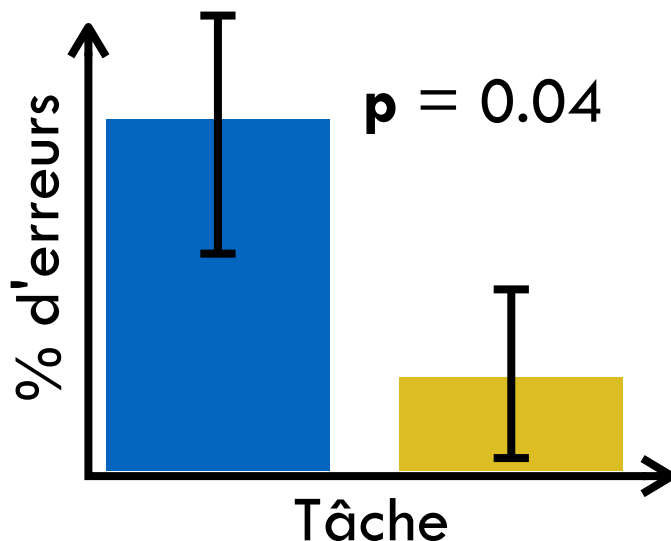
61

- $\alpha = 0.05$:
 - "Partant du principe (H_0) que les valeurs de M ne sont pas affectées par les variations de F , je considèrerai qu'une différence de cette ampleur est due à la chance tant que ça arrive moins d'une fois sur 20."

Un peu de stats (pour la culture)

62

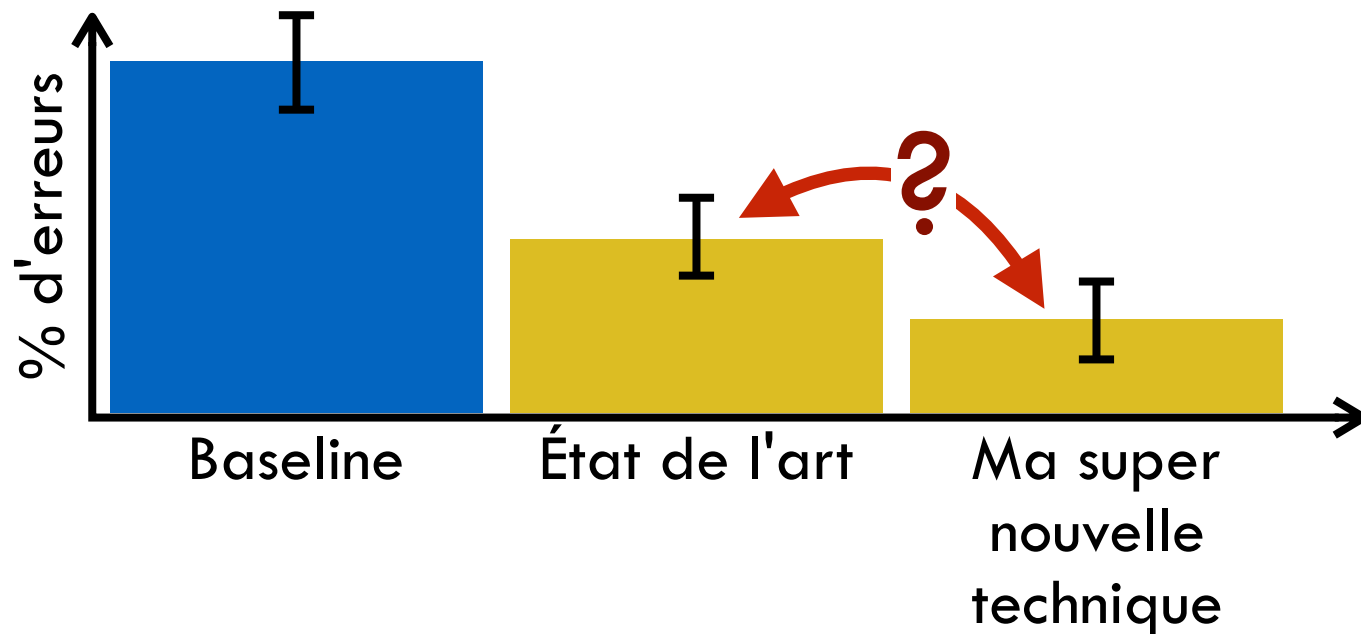
- $\alpha = 0.05$:
 - "Partant du principe (H_0) que les valeurs de B ne sont pas affectées par les variations de A, je considèrerai qu'une différence de cette ampleur est due à la chance tant que ça arrive moins d'une fois sur 20."
- p ne nous dit rien sur la **taille** de l'effet



Pour finir (1) : tests post-hoc

63

- ANOVA peut être utilisé sur des variables indépendantes à plus de deux niveaux

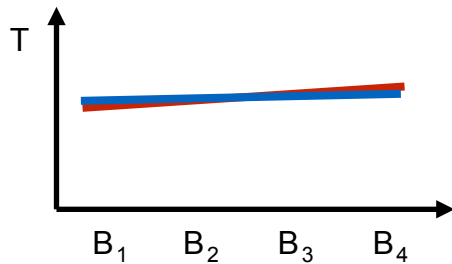


- Tests post-hoc : Comparent les mesures entre chaque niveau de la VI (aussi avec un α)
 - t-test, Tukey, Bonferroni, etc.

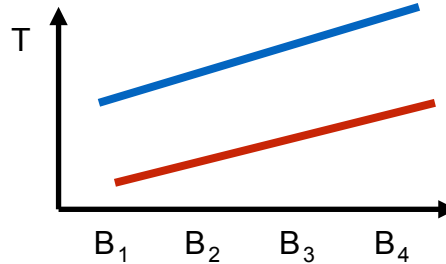
Pour finir (2) : tests sur plusieurs VI

64

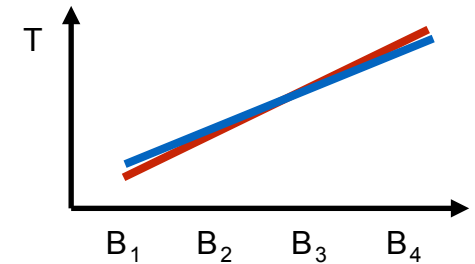
- ANOVA peut être utilisé sur plusieurs VI et sur leurs interactions



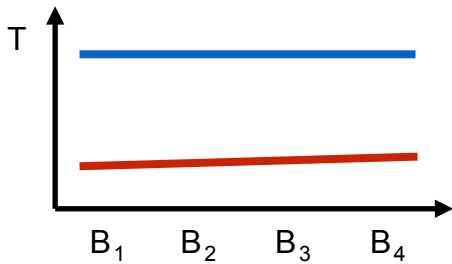
Pas d'effet significatif (E.S.)



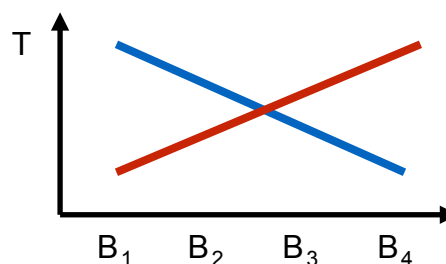
E.S. des périphériques et des blocs. Pas d'interaction.



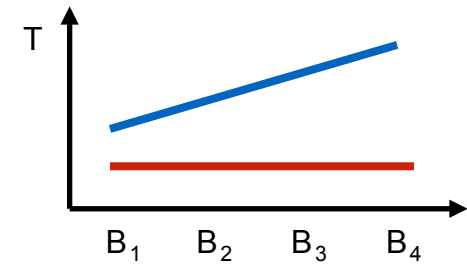
E.S. de Bloc.
Pas d'autre effet



E.S. du périphérique.
Pas d'autre effet



Interaction significative.
Pas d'autre effet



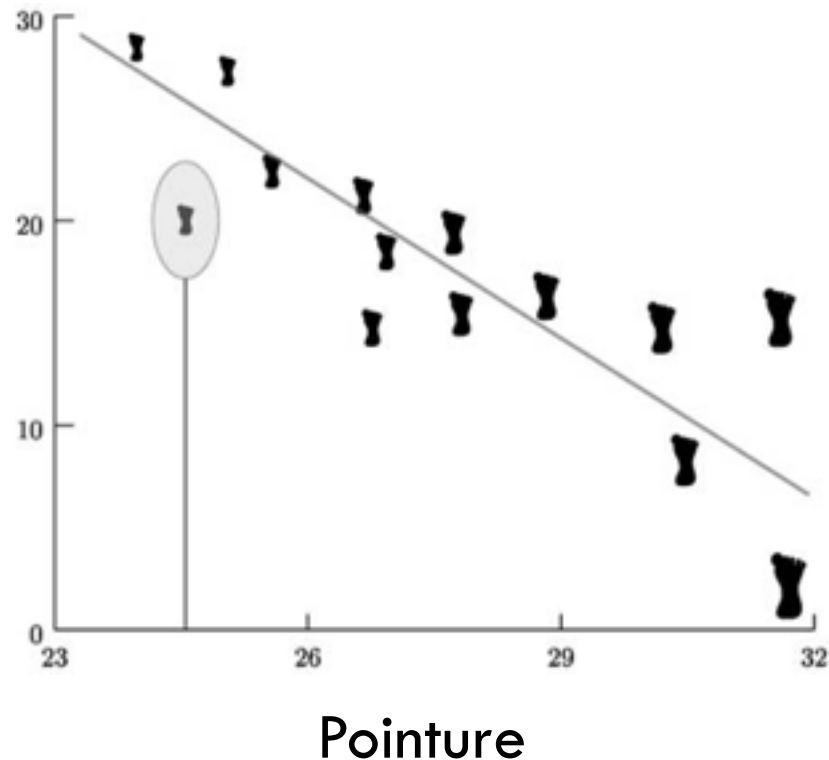
E.S. des périphériques.
Interaction

Corrélation vs. causation

65

- Un effet (statistique) d'un facteur sur une mesure ne signifie pas que la variation du facteur a causé la variation de la mesure.

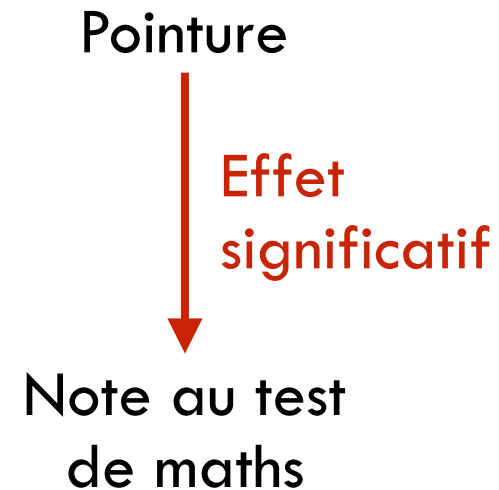
Nb d'erreurs
au test de maths



Corrélation vs. causation

66

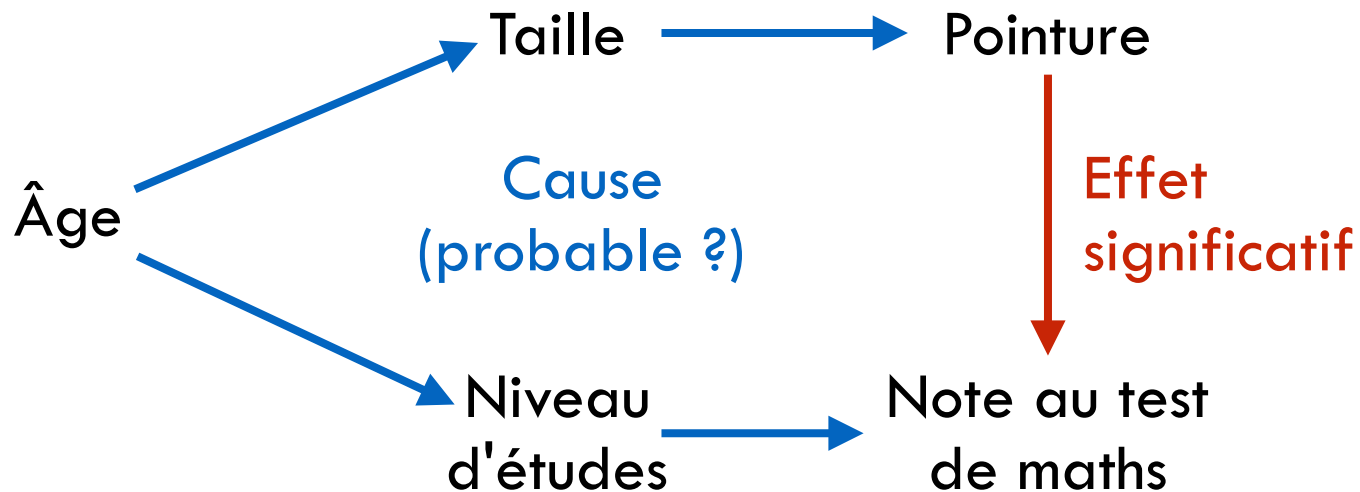
- Un effet (statistique) d'un facteur sur une mesure ne signifie pas que la variation du facteur a causé la variation de la mesure.



Corrélation vs. causation

67

- Un effet (statistique) d'un facteur sur une mesure ne signifie pas que la variation du facteur a causé la variation de la mesure.



Corrélation vs. causation

68

- Un effet (statistique) d'un facteur sur une mesure ne signifie pas que la variation du facteur a causé la variation de la mesure.

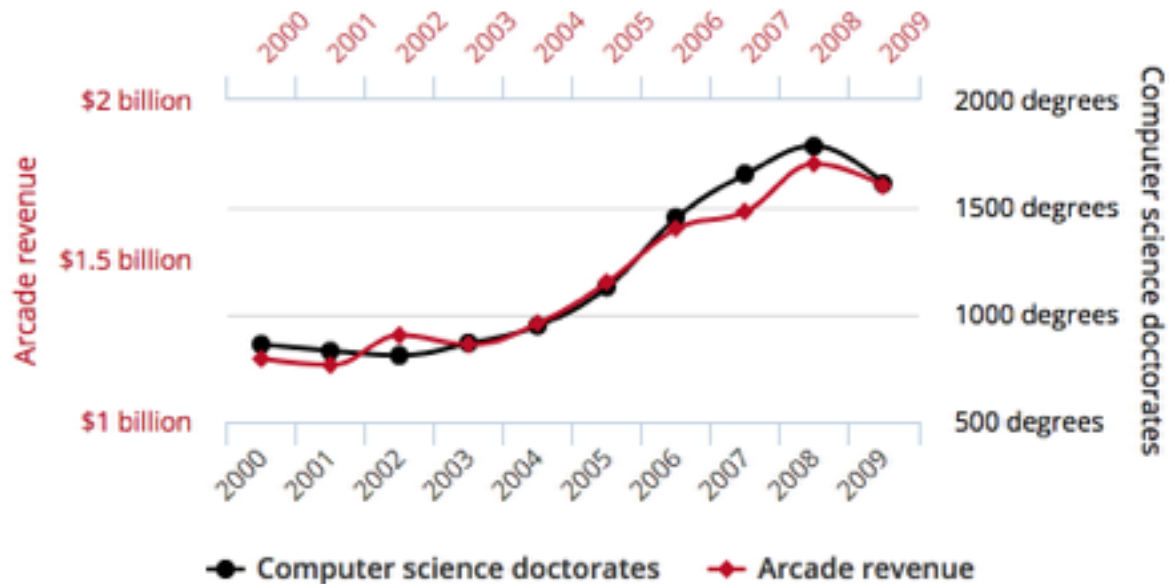


Corrélation vs. causation

69

Total revenue generated by arcades ≡
correlates with
**Computer science doctorates awarded
in the US**

Correlation: 98.51% ($r=0.985065$)



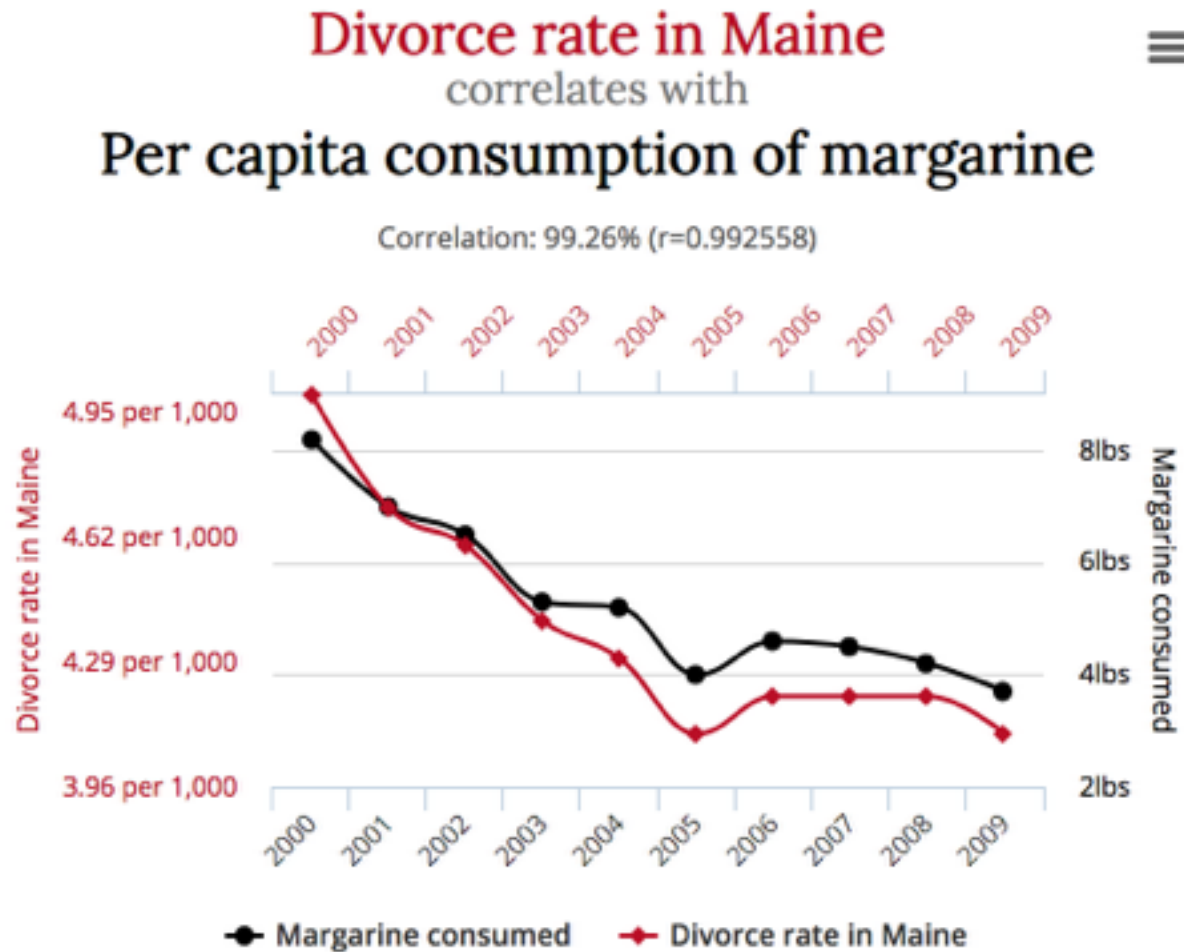
tylervigen.com

Data sources: U.S. Census Bureau and National Science Foundation

Sources : <http://www.tylervigen.com/spurious-correlations>

Corrélation vs. causation

70



tylervigen.com

Data sources: National Vital Statistics Reports and U.S. Department of Agriculture

Sources : <http://www.tylervigen.com/spurious-correlations>

Slides de l'an dernier

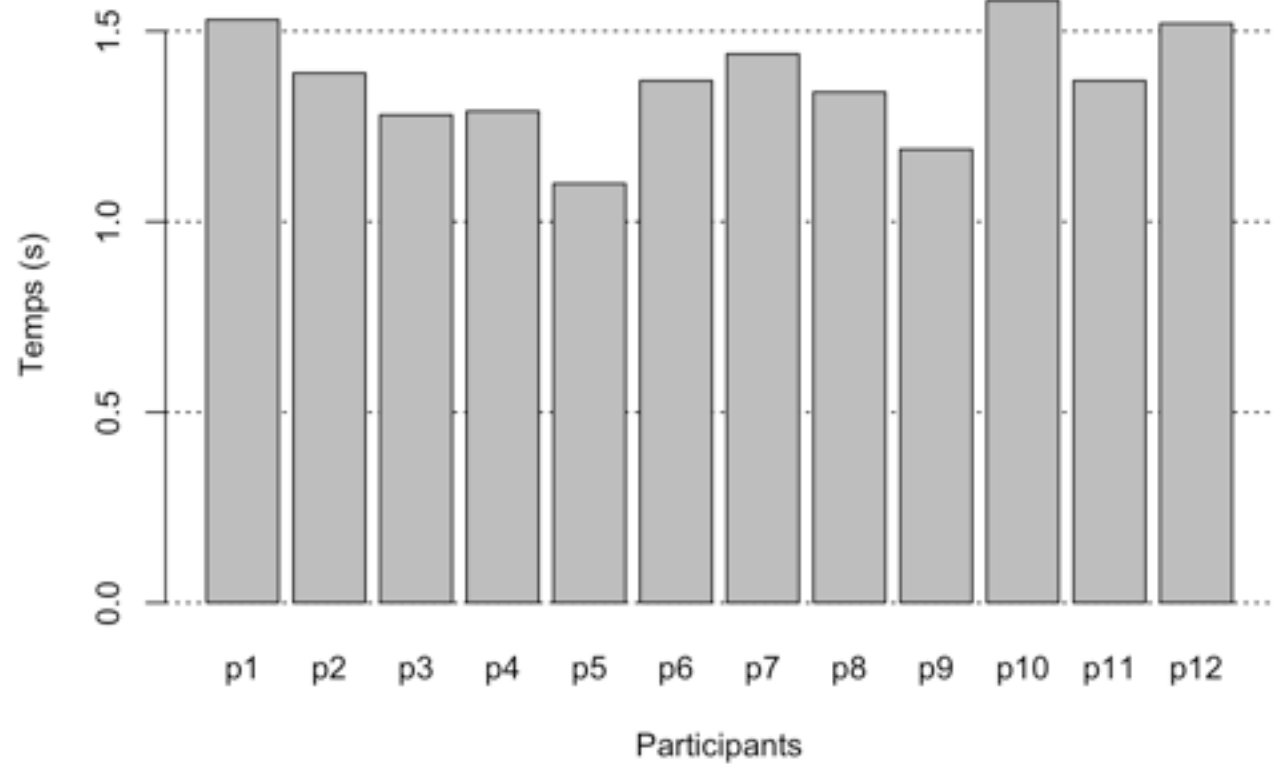
Analyse statistique

- Statistiques descriptives
- Exemple
 - La performance de 12 participants a été mesurée pour 2 périphériques (souris et tablette)
 - Plusieurs mesures de performance ont été réalisées
 - Chaque mesure de performance est une variable dépendante
 - Une des variables est le temps de réalisation (T)
 - La variable indépendante est le périphérique

Statistiques Descriptives

73

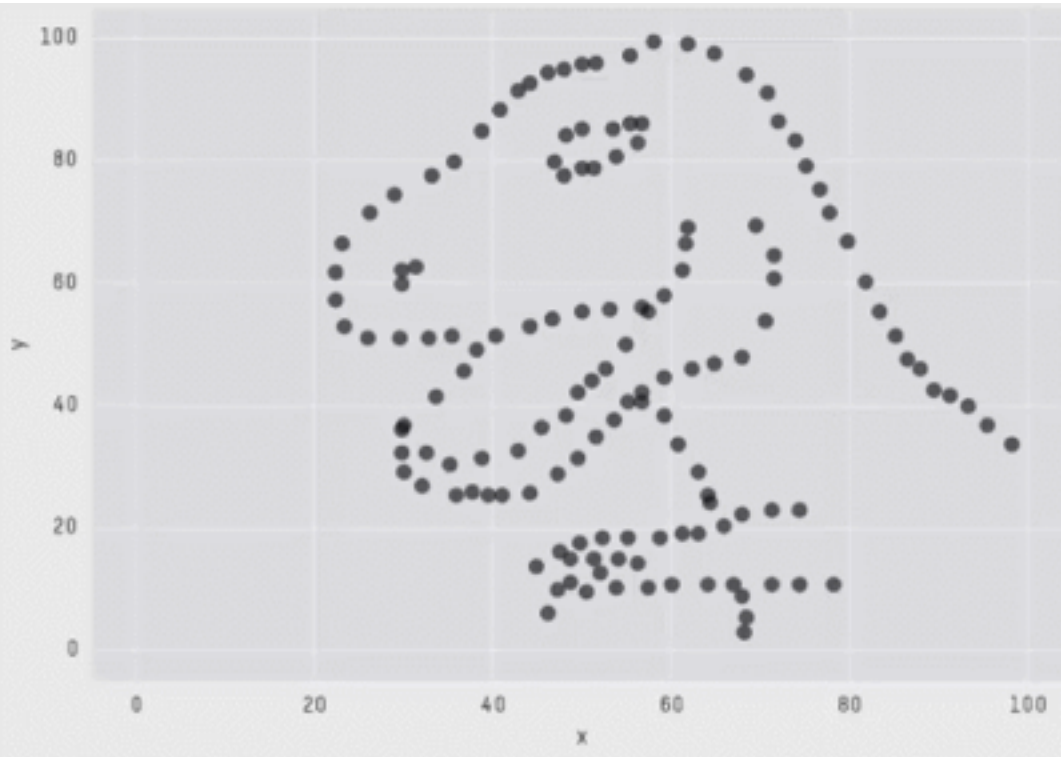
participant	souris
p1	1.53
p2	1.39
p3	1.28
p4	1.29
p5	1.10
p6	1.37
p7	1.44
p8	1.34
p9	1.19
p10	1.58
p11	1.37
p12	1.52



moyenne: 1.37 s
écart-type: 0.14 s

Toujours visualiser des données!

74



```
X Mean: 54.2659224
Y Mean: 47.8313999
X SD   : 16.7649829
Y SD   : 26.9342120
Corr.  : -0.0642526
```

Justin Matejka and George Fitzmaurice. 2017. Same Stats, Different Graphs: Generating Datasets with Varied Appearance and Identical Statistics through Simulated Annealing. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 1290-1294. DOI: <https://doi.org/10.1145/3025453.3025912>

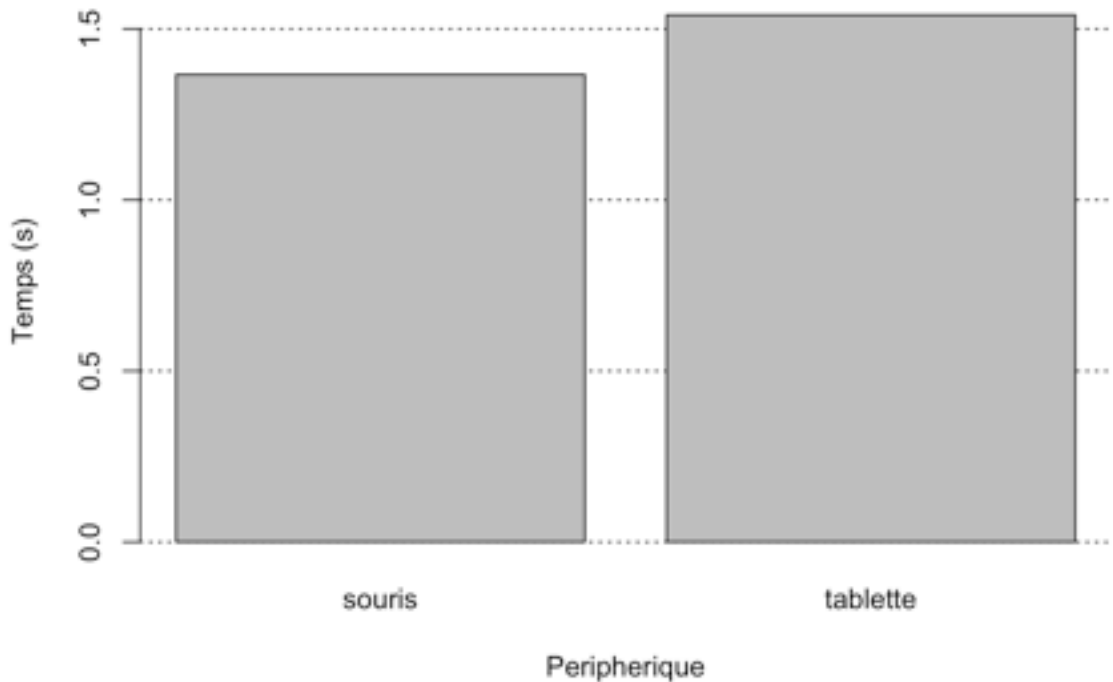
Evaluations comparatives

- Le résultat précédent, seul, n'est pas très intéressant
- L'objectif est souvent de comparer une ou plusieurs conditions
- Les conditions sont les niveaux de la variable indépendante
- Dans l'exemple, la variable indépendante est "Périphérique" et les niveaux sont "Souris" vs. "Tablette"

Evaluations comparatives

76

participant	souris	tablette
p1	1.53	1.35
p2	1.39	1.74
p3	1.28	1.81
p4	1.29	1.16
p5	1.10	1.43
p6	1.37	1.30
p7	1.44	1.46
p8	1.34	1.73
p9	1.19	1.81
p10	1.58	1.68
p11	1.37	1.66
p12	1.52	1.36



Les études comparatives sont plus intéressantes mais est-ce que ce résultat est plus intéressant?

Hypothèse nulle

- Déclarer comme “hypothèse statistiquement nulle” quelque chose qui est logiquement l’opposé de ce que l’on croit.
- Appeler cette hypothèse H_0
- Montrer à partir des données que H_0 est fausse, et doit être rejetée
- En rejetant H_0 , on confirme ce en quoi on croit

Hypothèse nulle

78

- L'hypothèse nulle est rejetée ou non

		Etat du monde	
		H0	H1
Décision	H0	Acceptation correcte	Erreur de type II (β)
	H1	Erreur de type I (α)	Rejet correct

ANOVA

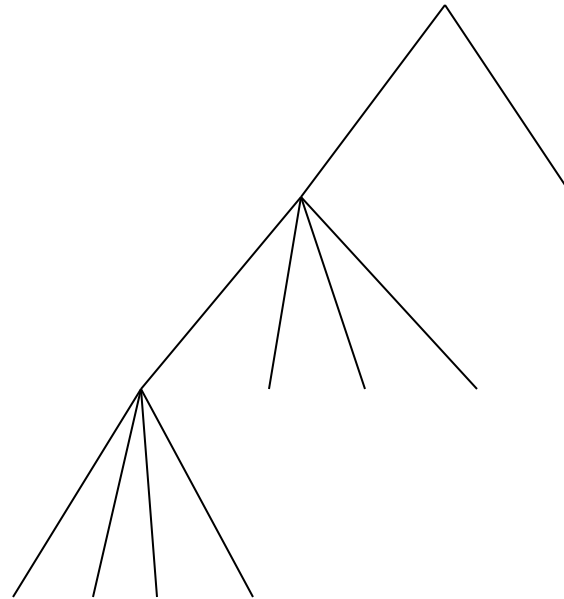
- C'est le principal outil statistique utilisé dans le domaine de l'interaction homme-machine pour évaluer des expériences
- Utilisé pour répondre à des questions du type "Est-ce que le temps pour accomplir telle tâche varie différemment suivant le type de technique d'interaction utilisé? "

Plan expérimental

Facteur 1 (ex: Périphérique): 2 niveaux (e.g. souris and tablette)

Facteur 2 (ex: Bloc): 4 niveaux

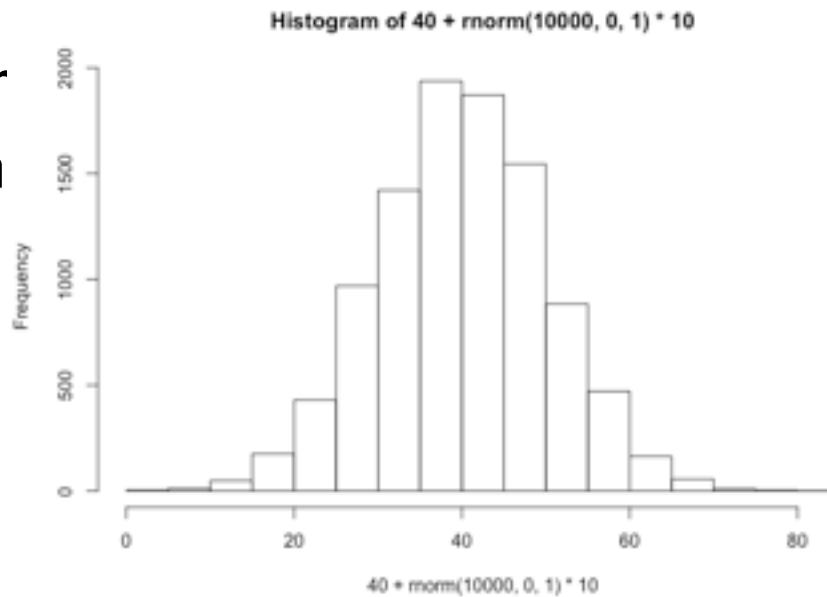
Facteur 3 (ex: ID): 4 niveaux



Conditions d'utilisation

- Indépendance des données
- Distributions normales

□ Hom
(con



noscedasticity)

ANOVA avec un facteur intra-sujets

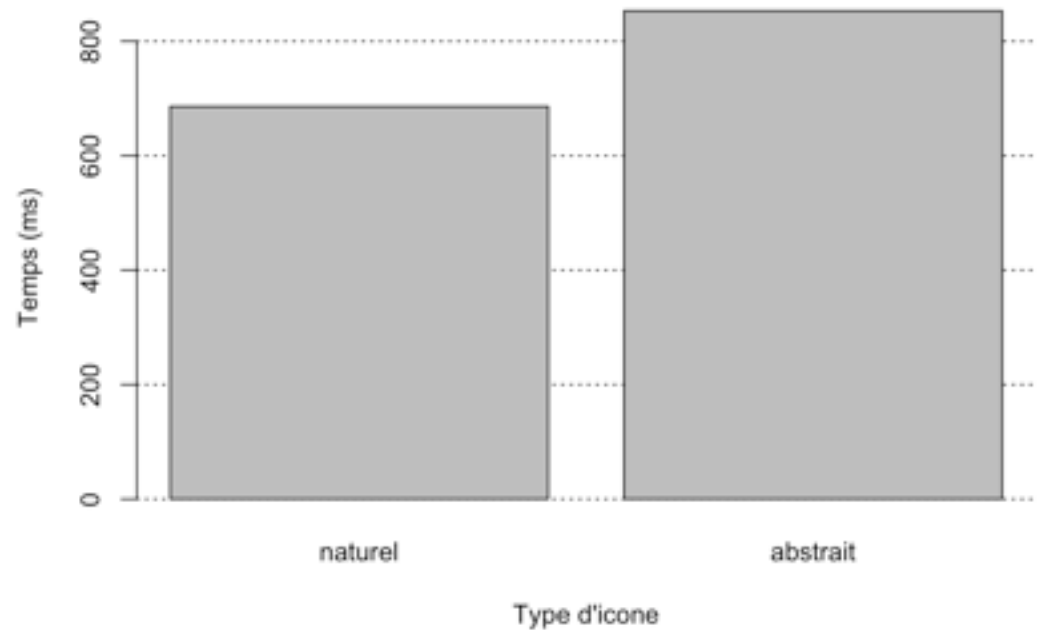
□ Exemple

- Apparence d'une icône avec 2 niveaux: naturel et abstrait
- Mesure du temps de sélection (ms)
- 10 participants

ANOVA avec un facteur intra-sujet

83

participant	naturel	abstrait
p1	741	936
p2	727	890
p3	671	838
p4	681	797
p5	693	839
p6	680	822
p7	721	854
p8	613	867
p9	637	845
p10	697	844



ANOVA avec un facteur intra-sujet

84

Test de la normalité des données

```
shapiro.test(data[, "naturel"])
```

```
##  
## Shapiro-Wilk normality test  
##  
## data: data[, "naturel"]  
## W = 0.95859, p-value = 0.7697
```

```
shapiro.test(data[, "abstrait"])
```

```
##  
## Shapiro-Wilk normality test  
##  
## data: data[, "abstrait"]  
## W = 0.92602, p-value = 0.4099
```

$p > 0.05$ donc données considérées comme suivant une distribution normale

Test d'homogénéité des variances

```
library(car)  
d3$type = factor(d3$type)  
kable(leveneTest(temps ~ type, data=d3))
```

	Df	F value	Pr(>F)
group	1	0.1420698	0.7106387
	18	NA	NA

$p > 0.05$ donc les variances des 2 groupes ne sont pas statistiquement différentes.

ANOVA avec un facteur intra-sujet

85

Effect	DFn	DFd	F	p	p<.05	ges
2 type	1	9	164.5409	4e-07	*	0.8367692

Effet significatif $F_{1,9} = 164.5$ $p < 0.001$ de l'apparence de l'icône sur le temps d'acquisition.

Nombre de degrés de liberté (nb de niveaux – 1),
(nb de niveaux – 1) x (nb de participants – 1)

Si $p < 0.05$, il y a 95% de chances que la différence observée n'est pas due au hasard

ANOVA avec un facteur intra-sujet

Effect	DFn	DFd	F	p	p<.05	ges
2 peripherique	1	11	4.660843	0.0538073		0.1926046

□ Retour à l'exemple des périphériques

$$F_{1,11} = 4.66 \quad p=0.054$$

p-value > 0.05 pas d'effet significatif sur le temps. Peut-on conclure qu'il n'y a pas de différence entre les deux périphériques?

Taille de l'effet (effect size)

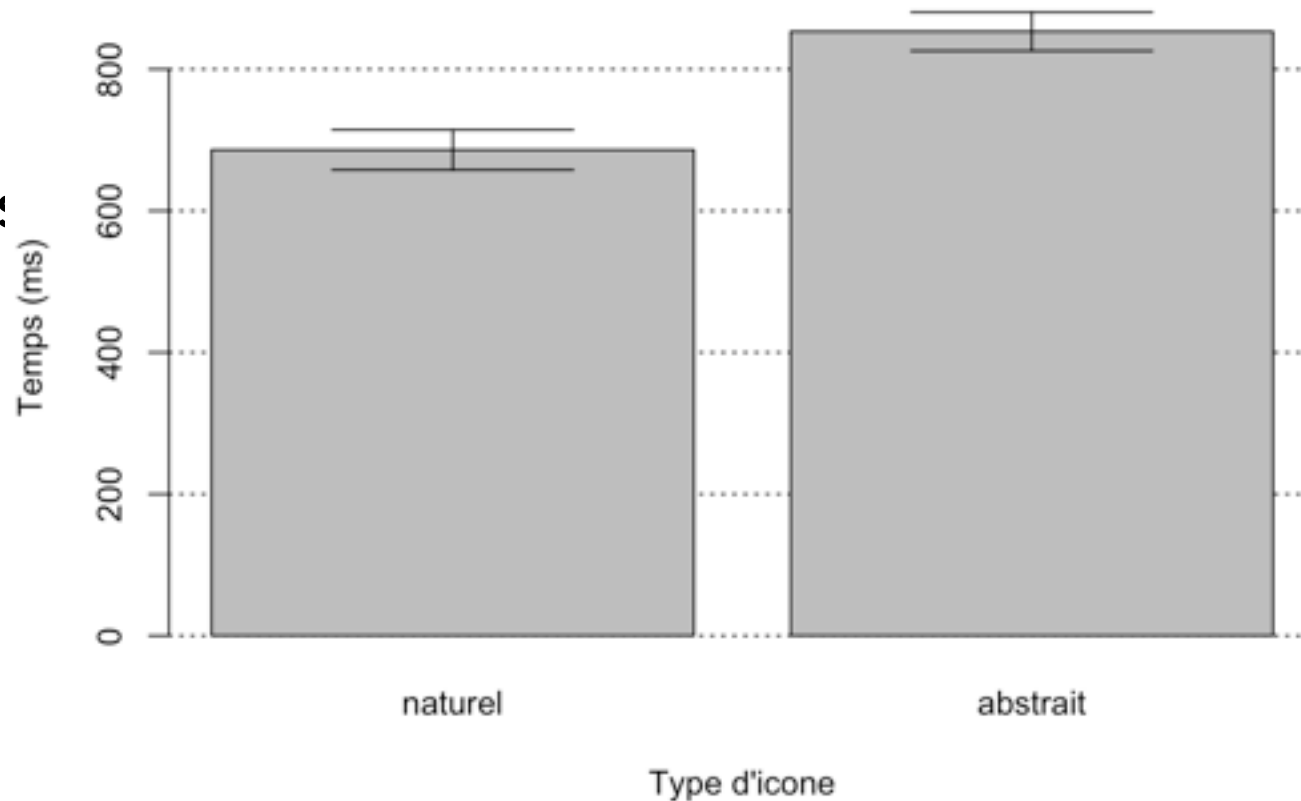
87

	Small	Medium	Large
Cases: →	Difference between the heights of 15 and 16 year old girls in the USA.[1]	M vs. F verbal fluency[2]	Difference between the heights of 13 and 18 year old girls in the USA.[3]
Measures ↓			
Cohen's d	0.2 - 0.3	≈ 0.5	0.8 - or larger
r (correlation coeff.) [7]	0.1 - 0.3	0.3 - 0.5	0.5 or larger
R²: % of variance explained	1%	9%	25%
R²: (≈ ≈ Partial eta-squared)	0.02	0.13	0.26
Partial eta-squared (η^2)	0.01	0.06	0.14
Generalized eta-squared η_G^2 [8]	0.02	0.13	0.26

Intervalle de confiance à 95%

88

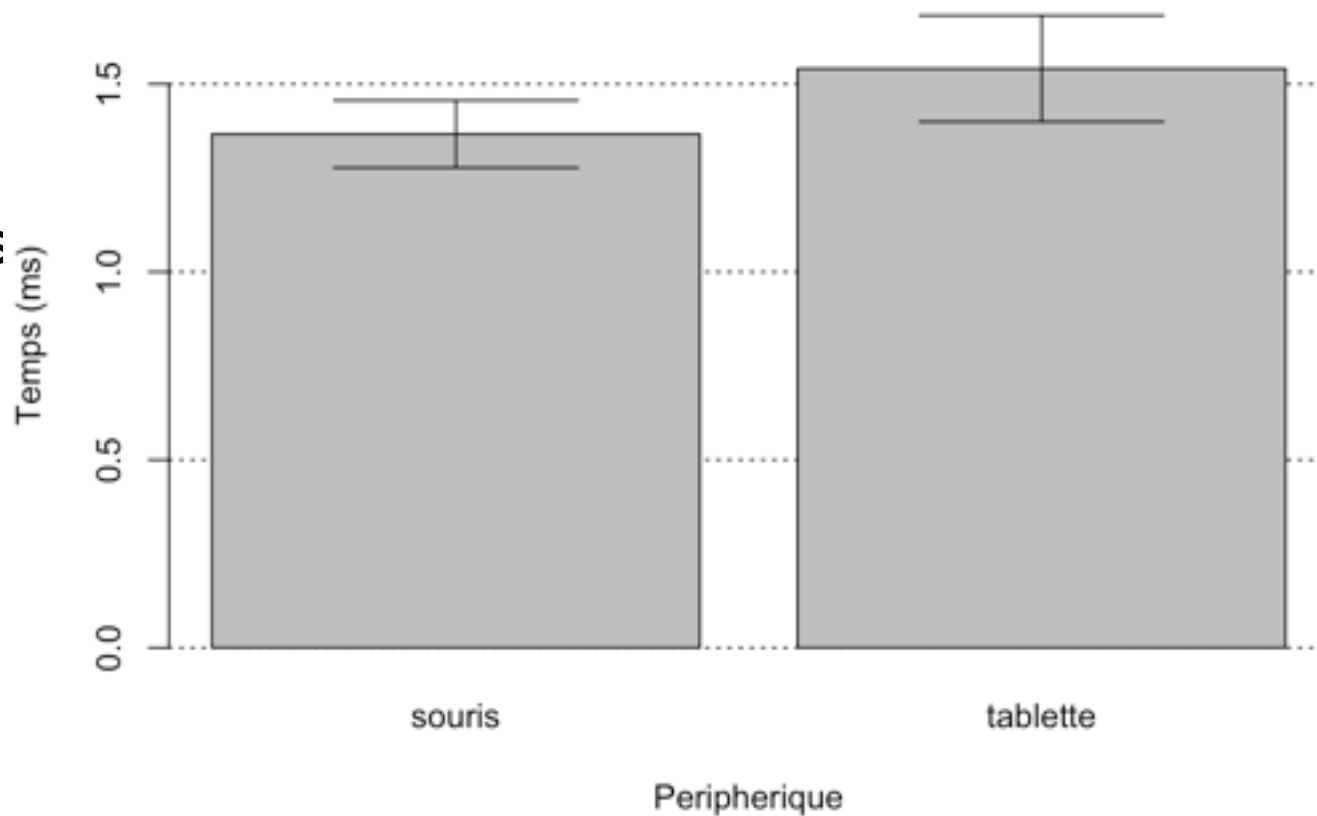
□ Cas



Intervalle de confiance à 95%

89

□ Cas



95% CI

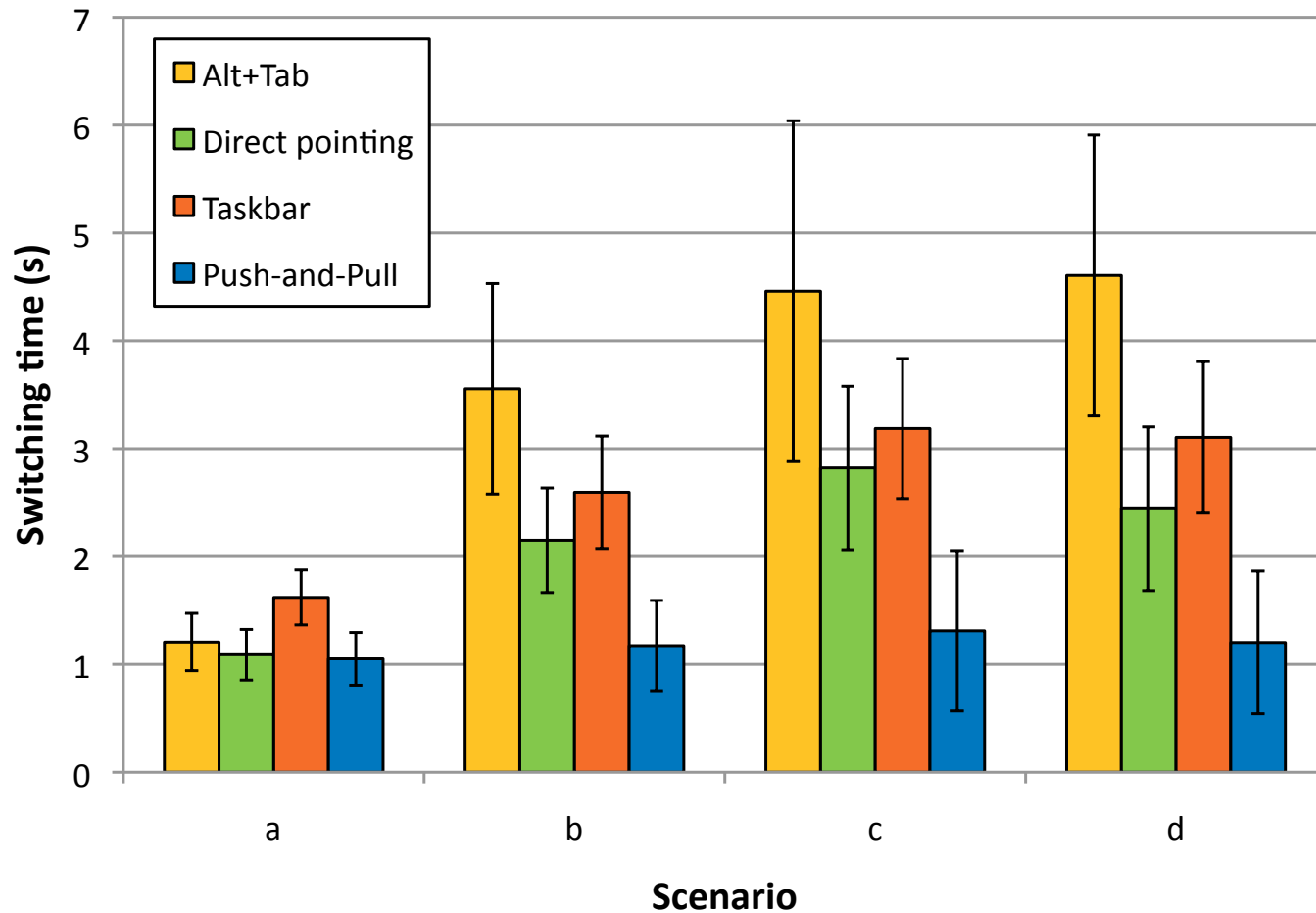


Figure 4. Mean switching time for SWITCHING TECHNIQUE and SCENARIO. Error bars represent 95% confidence interval.

ANOVA avec un facteur inter-sujets

□ Exemple:

- Tester si une interface ou une technique d'interaction fonctionne mieux avec les gaucher ou les droitiers (ou hommes vs femmes)
- 2 groupes de participants sont nécessaires: 5 gauchers (G) et 5 droitiers (D)
- La variable indépendante est la latéralité avec 2 niveaux, Gauche et Droite
- La variable dépendante est le temps (secondes) pour accomplir la tâche.

Conception intra vs. inter sujets

- Intra-sujets: 2 fois plus puissant avec 2 fois moins de participants (si 2 niveaux)...
- ... mais demande 2 fois plus de temps
- Quand c'est possible, la conception intra-sujets est préférée pour les groupes de petite taille

ANOVA avec un facteur intra-sujets et un facteur inter-

- Exemple: Contrôler si le contre-balancement annule les effets d'apprentissage
- Retour sur l'exemple des icônes abstraites et concrètes
- L'ordre a été contrebalancé entre les sujets

ANOVA avec un facteur intra-sujets et un facteur inter-

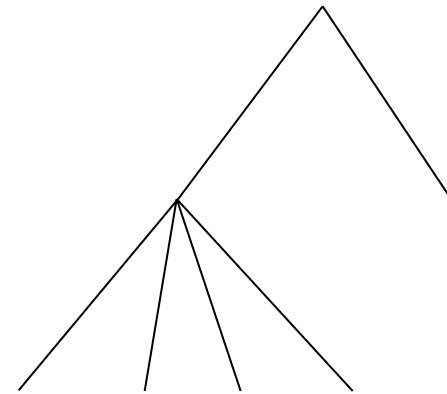
Effect	DFn	DFd	F	p	p<.05	ges
2 groupe	1	8	0.5368515	0.4846522		0.0453139
3 type	1	8	146.9600526	0.0000020	*	0.8431817
4 groupe:type	1	8	0.0383684	0.8495892		0.0014018

- Pas d'effet de groupe significatif ($F_{1,8} = 0.48$, ns)
- Pas d'interaction significative Type d'icône x Groupe ($F_{1,8} = 0.04$, ns)
 - → pas de transfert d'apprentissage asymétrique

ANOVA avec deux facteurs intra-sujets

- Exemple: 2 facteurs
 - Périphérique P1, P2
Facteur 1 (Périphérique): 2 niveaux
 - Bloc B1, B2, B3, B4
Facteur 2 (Bloc): 4 niveaux

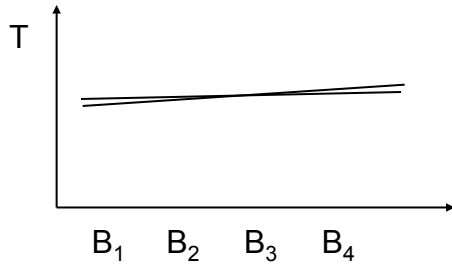
Facteur 2 (Bloc): 4 niveaux



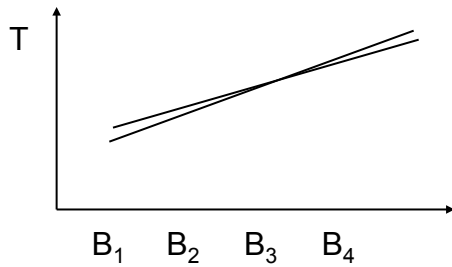
ANOVA avec deux facteurs intra-sujets

- Effet significatif principal: Périphérique ou/et Bloc
 - e.g. Effet significatif entre P1 et P2
 - On peut trouver une différence significative entre les blocs mais sans pouvoir conclure sur l'effet d'apprentissage
 - Pour connaître l'histoire complète: Etudier l'interaction Périphérique x Bloc

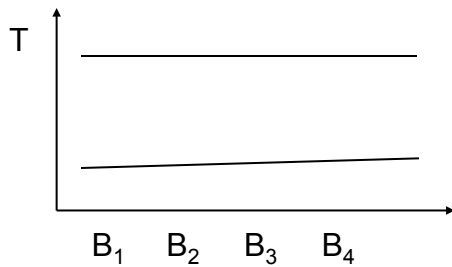
Interaction



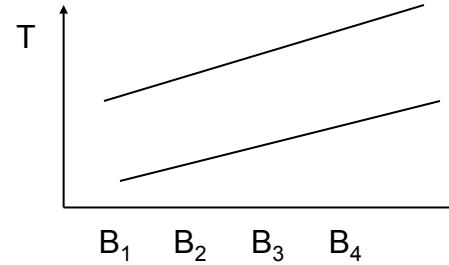
Pas d'effet significatif



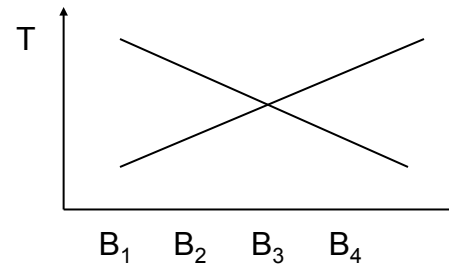
Effet significatif de Bloc. Pas d'autre effet



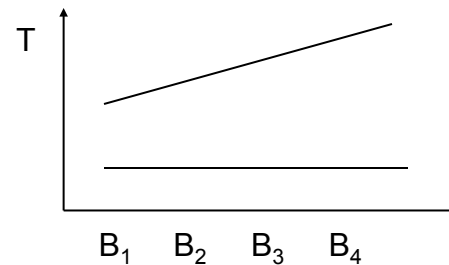
Effet significatif du périphérique. Pas d'autre effet



Effet significatif des périphériques et des blocs. Pas d'interaction



Interaction significative. Pas d'autre effet



Effet significatif des périphériques. Interaction

Test de sphéricité

98

```
kable(anova$`Mauchly's Test for Sphericity`)
```

	Effect	W	p p<.05
4	block	0.9395387	0.6462531
5	distance	0.2812080	0.0001391 *
7	task:block	0.8117481	0.2322464
8	technique:block	0.6412757	0.0445978 *
9	task:distance	0.7087711	0.0898549
10	technique:distance	0.7522581	0.1363227
11	block:distance	0.4064775	0.2134726
12	task:technique:block	0.9161025	0.5415126
13	task:technique:distance	0.2067270	0.0000161 *
14	task:block:distance	0.4338710	0.2668584
15	technique:block:distance	0.4676308	0.3396913
16	task:technique:block:distance	0.3352582	0.1039435

Test de sphéricité

99

A list containing one or more of the following components:

ANOVA A data frame containing the ANOVA results.

Mauchly's Test for Sphericity

If any within-Ss variables with >2 levels are present, a data frame containing the results of Mauchly's test for Sphericity. Only reported for effects >2 levels because sphericity necessarily holds for effects with only 2 levels.

Sphericity Corrections

If any within-Ss variables are present, a data frame containing the Greenhouse-Geisser and Huynh-Feldt epsilon values, and corresponding corrected p-values.

<https://cran.r-project.org/web/packages/ez/ez.pdf>

ANOVA avec deux facteurs intra-sujets et un facteur

- Même conception que précédemment
- Facteur inter-sujets: contre-balancement des périphériques

Aussi...

- 4 facteurs...
- ... n facteurs

Tests non paramétriques

Quand les utiliser?

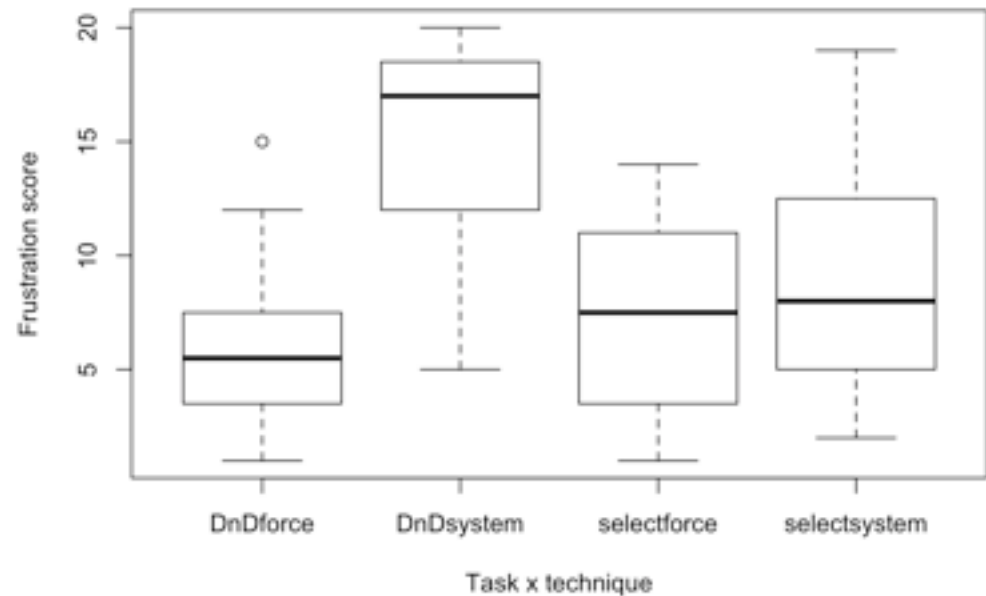
103

- Données de type ordinal : échelle de Likert
- Les hypothèses de l'ANOVA ne sont pas respectées
 - ▣ d'autres solutions existent (transformation log, ART...)
- Limite: pas d'analyse des interactions possible
- Souvent utilisé en IHM: Friedman (within subject)

Exemple: analyse de NASA-TLX

104

participant	DnDforce	DnDsystem	selectforce	selectsystem
0	4	20	5	13
1	6	9	4	11
2	1	17	2	2
3	3	13	3	2
4	8	20	11	6
5	12	20	9	12
6	3	17	8	19
7	15	17	12	15
8	4	19	14	17
9	7	18	13	7
10	6	15	2	2
11	5	11	8	8
12	7	7	11	5
13	5	5	4	5
14	1	17	1	10
15	9	14	7	8



Exemple: analyse de NASA-TLX

105

```
Friedman rank sum test
```

```
data: data.matrix(data_tr)
```

```
Friedman chi-squared = 21.939, df = 3, p-value = 6.717e-05
```

```
Pairwise comparisons using Wilcoxon rank sum test
```

```
data: mdata$score and mdata$tasktech
```

	DnDforce	DnDsystem	selectforce
DnDsystem	0.00034	-	-
selectforce	1.00000	0.00128	-
selectsystem	0.70051	0.02153	1.00000

```
P value adjustment method: bonferroni
```

Coder l'expérience

- **Toujours enregistrer les données brutes!**
 - Evite les erreurs
 - Evite les oublis

Conclusion

107

