

# IFT 2905: INTERFACES PERSONNE-MACHINE

Évaluation: interviews et sondages

<http://tiny.cc/ift2905>

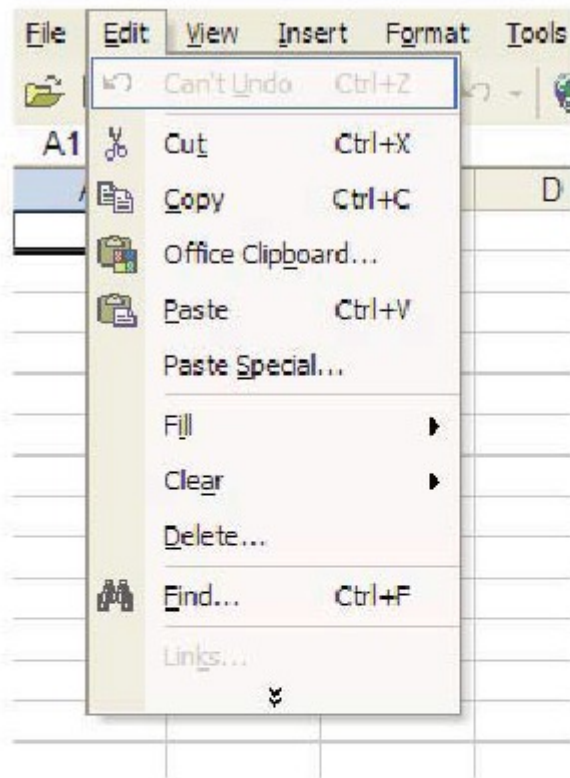


Image © Katerina Kamprani, [www.theuncomfortable.com](http://www.theuncomfortable.com)

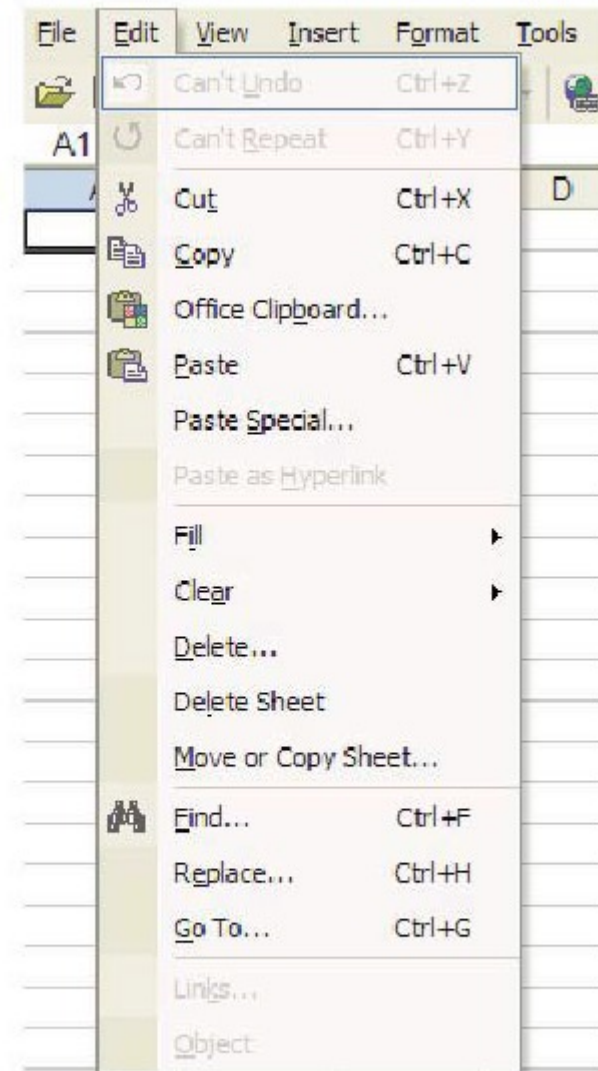
## Mikhail Bessmeltsev

Certaines diapos de Yoon, McLean et Thomaszewski

# HORREUR OU SPLENDEUR?

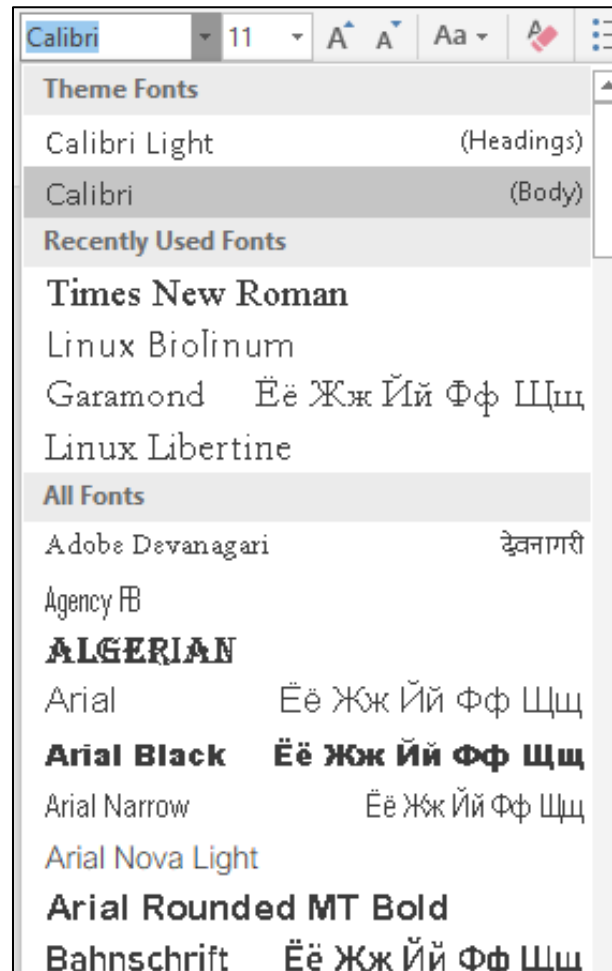


**Menu  
Adaptatif**



. All rights reserved.

## HORREUR OU SPLENDEUR?

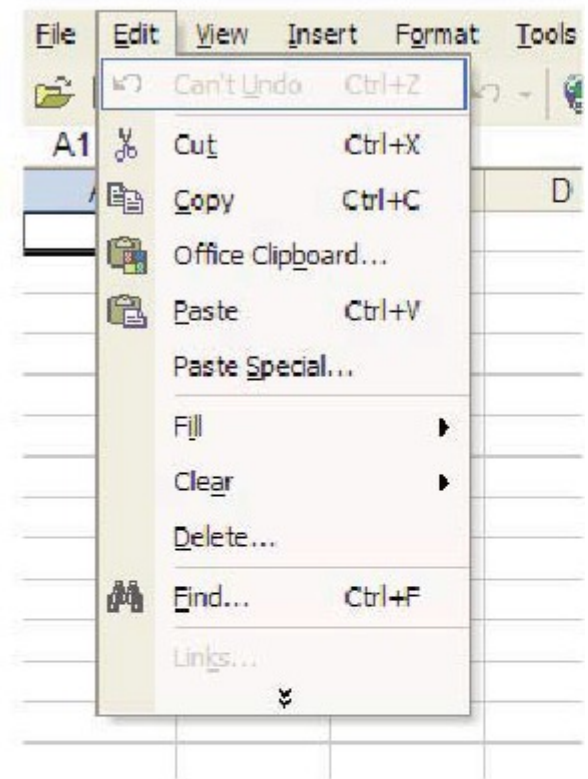


© Microsoft, 2021. All rights reserved.

# MENUS ADAPTATIFS

la simplicité vs l'efficacité vs visibilité

- complexe et imprévisible
- mémoire spatiale?
- manque de contrôle de l'utilisateur



Countries
Canada
<b>Uganda</b>
Jamaica
Mexico
Netherlands
France
Australia
Germany
China
Indonesia
Chile
Egypt
Russia
Zambia
Guatemala
Bolivia
Italy

(a)

Countries	
Uganda	↓ ↑
France	↓ ↑
Germany	↓ ↑
Chile	↓ ↑
Canada	↑
Netherlands	↑
Australia	↑
China	↑
Indonesia	↑
Egypt	↑
Jamaica	↑
Russia	↑
Zambia	↑
Guatemala	↑
Mexico	↑
Bolivia	↑
Italy	↑

(b)

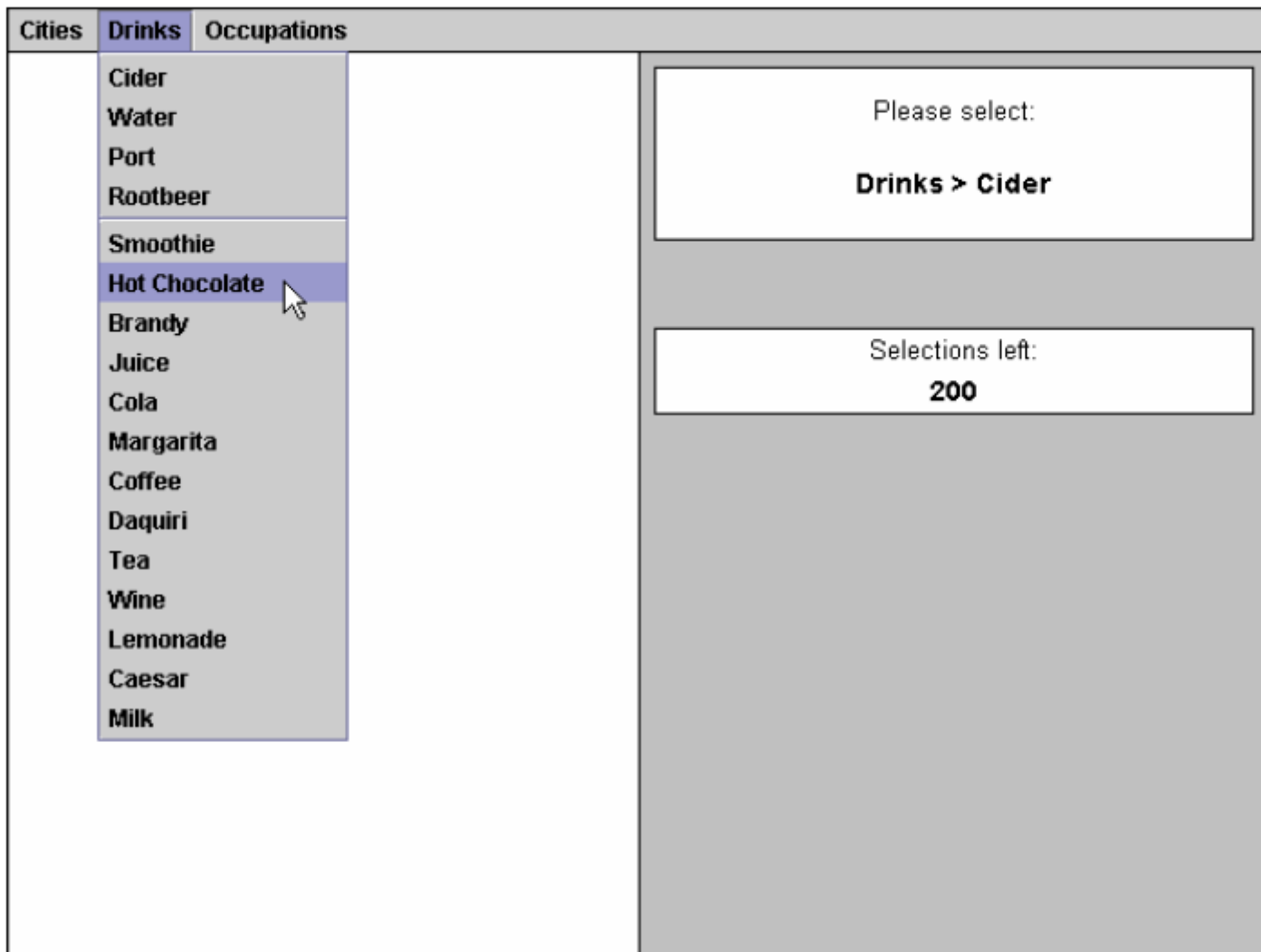
Move items up and down in top partition

Move items from bottom to top partition

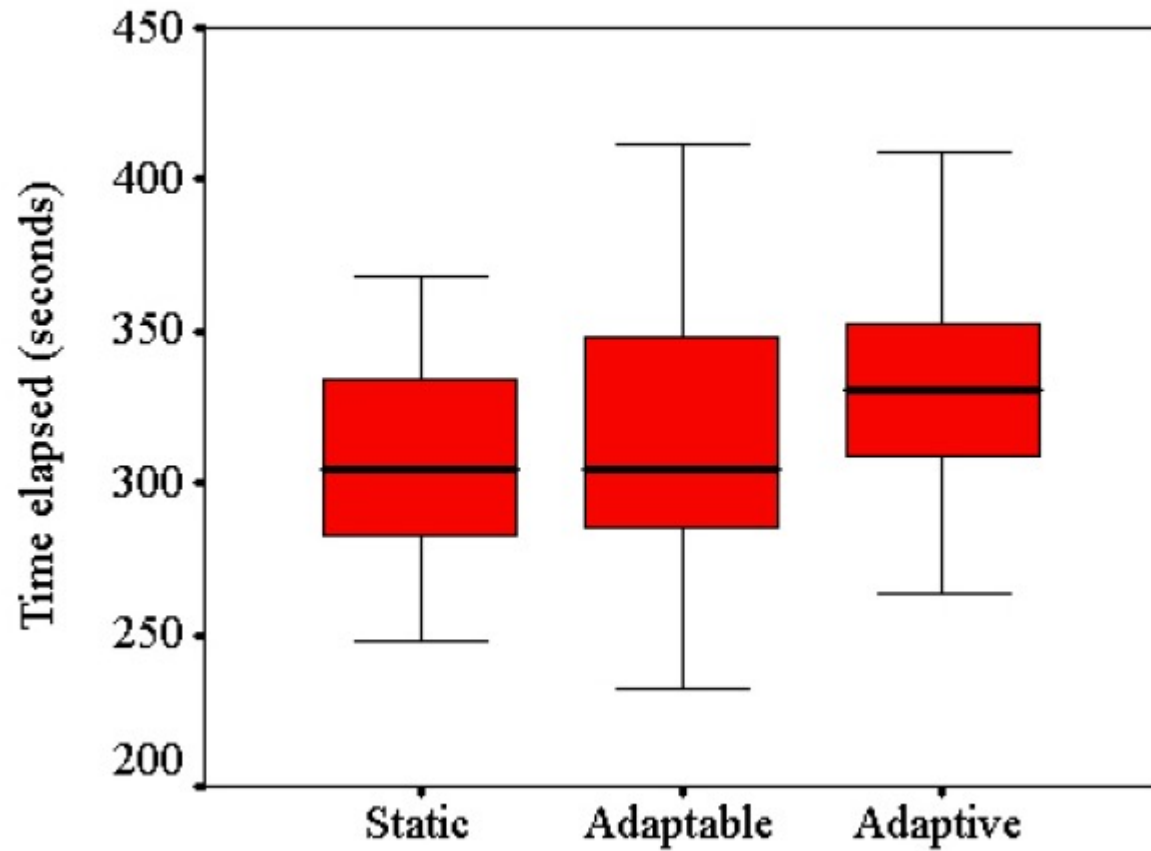
# Statique | Adaptable

## A Comparison of Static, Adaptive, and Adaptable Menus

Leah Findlater and Joanna McGrenere  
 Dept. of Computer Science, University of British Columbia  
 Vancouver, B.C., Canada  
 {lkf, joanna}@cs.ubc.ca



**Figure 4. The experimental system, showing the prompt area on the right-hand side and the menus at the top-left.**



**Table 3. Chi-square statistic for qualitative results (S=static; AV=adaptive; AB=adaptable) (N=27).**

Dependent Variable	Ranked 1 <sup>st</sup> (fre-			df	Chi-square	Sig. Level
	S	AV	AB			
Preferred overall	4	8	15	2	6.89*	.032
Most efficient	6	5	16	2	8.22*	.016
Fewest errors	9	6	10	2	1.04	.595
Most frustrating	10	15	1	2	11.62*	.003
Initially easiest to use	5	4	17	2	12.08*	.002

\* Significant at the .05 level.



# INTERVIEWS

Sont bonnes pour

- Explorer les enjeux
- Mieux comprendre les tâches
- Comprendre le processus de réflexion de l'utilisateur

Peuvent être utilisées à n'importe quelle étape du processus de design

# INTERVIEWS

## + Sont bonnes pour explorer les enjeux

- Plus interactives que les observations
- Plus flexible que les questionnaires/sondages
  - nous pouvons suivre les enjeux au fur et à mesure qu'ils surviennent

## - Ce que l'utilisateur dit est subjectif

- Ce qu'il fait  $\neq$  ce qu'il dit
- Peut oublier et rationaliser

## - Ce que l'observateur écrit est subjectif

## - Peuvent être intimidantes pour l'utilisateur

## - Prennent du temps (faire + analyser)

# **INTERVIEWS SONT AJUSTABLES → UTILES**

On peut changer:

- **Nombre de personnes**
  - 1-1, paires, groupes
- **Types**
  - Structurée, semi-structurée, déstructurée
- **Duration et profondeur/largeur**
- **Lieu**
  - Réalité ou laboratoire

+ on peut les combiner avec d'autres techniques

# ENREGISTREMENT DES OBSERVATIONS

## Crayon et papier

- Difficile à enregistrer tous les détails

## Enregistrement audio

- Tous les détails

## Enregistrement vidéo

- Si le langage corporel est important

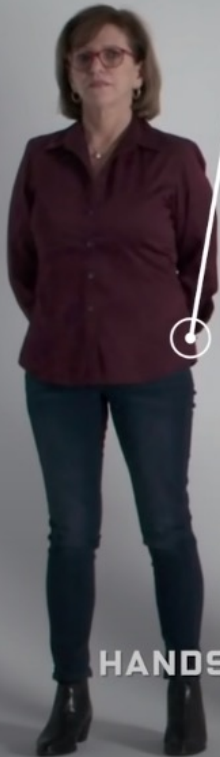


# ENREGISTREMENT DES OBSERVATIONS

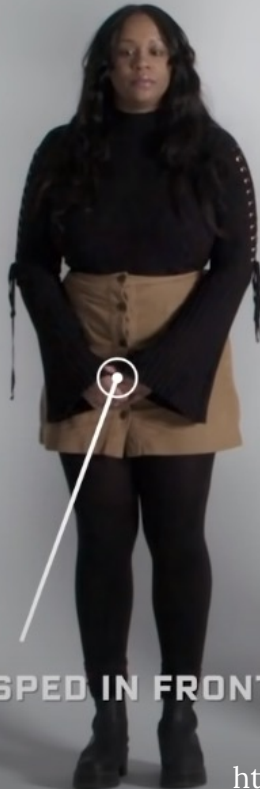
## Enregistrement vidéo

- Si le langage corporel est important

HANDS BEHIND BACK



HANDS CLASPED IN FRONT



HANDS IN POCKETS



ARMS TO THE SIDE



LEANING



# ENREGISTREMENT DES OBSERVATIONS

## Enregistrement vidéo

- Si le langage corporel est important

Presque toutes les interprétations du langage corporel relèvent de la **pseudoscience**

HANDS BEHIND BACK

HANDS IN POCKETS

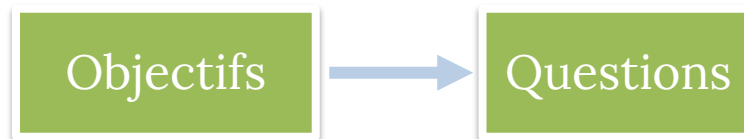
LEANING

ARMS TO THE SIDE

HANDS CLASPED IN FRONT

# AVANT L'INTERVIEW

- **Préparer les questions**
  - Elles doivent soutenir vos objectifs d'évaluations
  - Elles doivent être faciles à comprendre et répondre
- **Soyez prêtes**
  - Vérifier l'équipement
  - Ayez tous les documents nécessaires
- **Faites les projets pilotes**
  - Améliorer les questions
  - Il faut pratiquer!



# PENDANT L'INTERVIEW

- **Appliquer une structure**
  - début, milieu et fin clairs
- **Donner du contexte aux participants**
  - Expliquer pourquoi vous êtes là, qu'est-ce que vous voulez apprendre
- **Utiliser des accessoires**
  - Prototypes, photos, ...
  - Éviter contextes abstraits



# ÉCOUTER!

- Soyez intéressé, attentif et flexible
- Établir un contact visuel
- Donner le temps de réfléchir
- Renvoyer aux choses déjà dites
- Contrôler les conversations s'ils sont hors sujet

Plus de détails (en anglais):

[https://youtu.be/9t-\\_hYjAKww](https://youtu.be/9t-_hYjAKww)  
<https://youtu.be/WPrAWW4YGPw>

[https://www.captionfy.com/video/youtu  
be/9t-\\_hYjAKww?caption=fr](https://www.captionfy.com/video/youtu<br/>be/9t-_hYjAKww?caption=fr)

# TYPES D'INTERVIEWS

1. Non-structurée
2. Semi-structurée
3. Structurée
  
4. En groupe
5. Rétrospective (l'utilisateur se souvient)
6. Contextuelle (l'utilisateur est en train de travailler)

# SONDAGES/ QUESTIONNAIRES

<p><b>Quelle est votre garniture de pizza préférée ?</b></p> <p><input type="radio"/> Fromage</p> <p><input type="radio"/> Champignons</p> <p><input checked="" type="radio"/> Anchois</p> <p><input type="radio"/> Saucisse</p> <p><input type="radio"/> Cœurs d'artichaut</p>	<p><b>Quelles garnitures de pizza aimez-vous ?</b></p> <p><input type="checkbox"/> Fromage</p> <p><input checked="" type="checkbox"/> Champignons</p> <p><input type="checkbox"/> Anchois</p> <p><input checked="" type="checkbox"/> Saucisse</p> <p><input type="checkbox"/> Cœurs d'artichaut</p>
---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

# SONDAGES

(=questionnaires, pas de différence pour nous)

## Similaires aux interviews

- questions fermées et ouvertes

## Plus de participants / à distance

Peuvent collecter les données subjectives, objectives, qualitatives et quantitatives

- Évaluer pour mieux comprendre (tôt)
- Évaluer prototypes (souvent utilisé avec d'autres formes d'évaluation)

# SONDAGES

- + réalisation n'est pas chère
- + ne nécessite pas la présence d'un évaluateur
- + la plupart des résultats peuvent être quantifiés

- Préparation coûte cher
  - Questions sont difficiles à bien faire
- peuvent avoir un faible taux / faible qualité des réponses
- Difficile à faire un suivi approfondi

# COMMENT CRÉER

- **Établir un objectif**
  - Comment allez-vous utiliser les résultats?
  - Quelle information est nécessaire?
  - Comment allez-vous les analyser?
- **Déterminer l'audience**
  - E.g. 50-1000 utilisateurs
  - Échantillonnage
- **Tester tout**
  - le libelle
  - le temps
  - la validité
  - l'analyse

# QUESTIONS

On ne peut pas les clarifier

- contrairement aux interviews

Conseilles:

- Courtes!
- Clair comment répondre
- Une seule question à chaque fois  
*PAS: « À quelle fréquence avez-vous utilisé le système et qu'est-ce que vous en appréciez? »*
- Éviter les biais
- Éviter toute ambiguïté
- Piloter!



«Dans quelle mesure notre application est-elle utile?»

# BIAIS

«Dans quelle mesure notre application est-elle utile?»

Question insidieuse

# BIAIS

~~« Dans quelle mesure notre application est-elle utile? »~~

« Avez-vous trouvé l'application utile? »

# BIAIS

Quel impact notre produit a-t-il sur votre vie?

- extrêmement positif
- très positif
- plutôt positif
- pas positif

échelles déséquilibrées

# BIAIS

Quel impact notre produit a-t-il sur votre vie?

- extrêmement positif
- très positif
- plutôt positif
- pas positif

échelles déséquilibrées

# BIAIS

Quel impact notre produit a-t-il sur votre vie?

~~extrêmement positif~~

~~très positif~~

~~plutôt positif~~

~~pas positif~~

très positif

plutôt positif

plutôt négatif

très négatif



# BIAIS

1. Combien de secondes vous faut-il pour résoudre l'équation suivante:

$$y(2x^4 + y) \frac{dy}{dx} = (1 - 4xy^2)x^2$$

Votre réponse: \_\_\_\_\_

2. Comment évaluez-vous vos compétences en résolution de mathématiques?

- Très fortes
- Fortes
- OK/Meh
- Faibles
- Très faibles



Effet d'amorçage ("priming")

# BIAIS

1. Combien de secondes vous faut-il pour résoudre l'équation suivante:

$$y(2x^4 + y) \frac{dy}{dx} = (1 - 4xy^2)x^2$$

Votre réponse: \_\_\_\_\_

2. Comment évaluez-vous vos compétences en résolution de mathématiques?

- Très fortes
- Fortes
- OK/Meh
- Faibles
- Très faibles

# VALIDITÉ

Vos questions testent-elles ce que vous voulez?

On peut l'augmenter par:

- Pilotage
- Triangulation
  - Tester l'hypothèse avec plusieurs questions ou types d'évaluation
- Utiliser des questionnaires préalablement validés

# TEMPS

Les questionnaires sont limités par leur longueur et leur complexité

Pas possible de demander tout ce que vous voulez

Concentrer sur les choses les plus importants à évaluer, mais ne vous concentrez pas trop sur ce que vous vous attendez à voir

- exclusion potentielle d'autres explications

# ADMINISTRER DES QUESTIONNAIRES

- **En personne**
  - + taux d'achèvement élevé
  - Prends beaucoup de temps
- **Ramener à la maison**
  - faible taux d'achèvement
- **Email**
  - + laisse plus de temps aux utilisateurs
  - les utilisateurs ont tendance à donner plus de réponses de forme libre
  - le taux de réponse varie
- **Web**
  - + vérifier automatiquement
  - + facile à analyser (peut être)
  - le taux de réponse varie

# EXEMPLE: CONTEXTE

Comment administrer?

1. Vous voulez demander aux patients qui viennent d'essayer une nouvelle insuline pompe ce qu'ils en ont pensé
2. Vous souhaitez diffuser largement un questionnaire sur les problèmes de santé généraux aux personnes âgées de 20 à 55 ans

# AUTRES CONSIDÉRATIONS

- soyez attentif à vos répondants et au contexte dans lequel vous y accédez
- atteintes à la vie privée: faites attention à ce que vous demandez
- capacité: limitations comme l'alphabétisation et le handicap
- motivation
  - pourquoi le répondant devrait-il s'en soucier?
  - ont généralement besoin d'offrir quelque chose en retour
  - ... Mais attention à ne pas introduire de biais

# AUTRES TESTS UTILISATEURS

## Etude sur le terrain

Trouver des problèmes en contexte

Evaluer un prototype fonctionnel, dans un vrai contexte, sur des vraies tâches

Observations surtout qualitatives

## Expérience contrôlée

Tester une hypothèse (interface A est plus rapide que interface B)

Evaluer un prototype fonctionnel dans un environnement contrôlé (laboratoire) sur des tâches choisies

Observations surtout quantitatives (temps, taux d'erreur, satisfaction, ...)

# TYPES D'ÉVALUATION

Evaluation **formative**

pendant les itérations du développement

Evaluation **sommative**

à la **fin** du développement



# EVALUATION FORMATIVE

## « POURQUOI? »

- Souvent qualitative
- Trouver les problèmes pour la prochaine itération
- Evaluer un prototype sur une tâche précise

# EVALUATION FORMATIVE

## « POURQUOI? »

- Souvent qualitative
- Trouver les problèmes pour la prochaine itération
- Evaluer un prototype sur une tâche précise

### Caractéristiques

- Souvent fait lors des tests de prototypes en papier avec de *faux* usagers
- L'environnement peut être contrôlé (comme un bureau)
- L'évaluateur choisit les tâches à tester provenant de l'analyse des tâches faite au départ

# EVALUATION SOMMATIVE

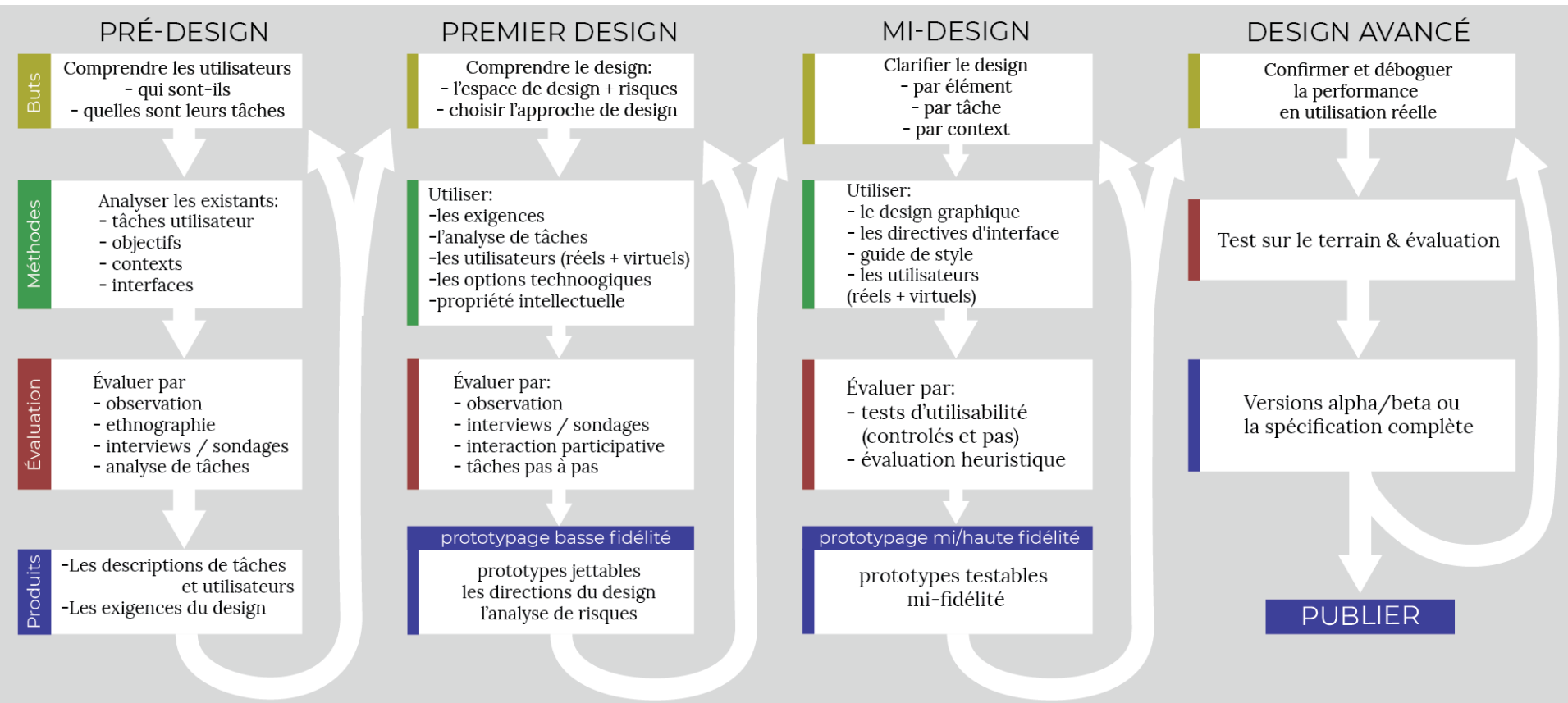
- Souvent quantitative
- Justifier que votre application est mieux (qu'avant, que les concurrents, ...)
- Evaluer un produit complet

	<b>formative</b>	<b>sommative</b>
genre d'apprentissage	exploration	évaluation
type de données	qualitative	quantitative
niveau de contrôle	moins	plus
formalité	plutôt moins	plutôt plus
phase	pré/premier/mi-design	mi-design/design avancé
coût	souvent moins cher	souvent cher
tâches utilisateur	relativement ouvertes	attribuées

Dérivé d'une diapo de "Evaluating User Interfaces"  
par University of Minnesota

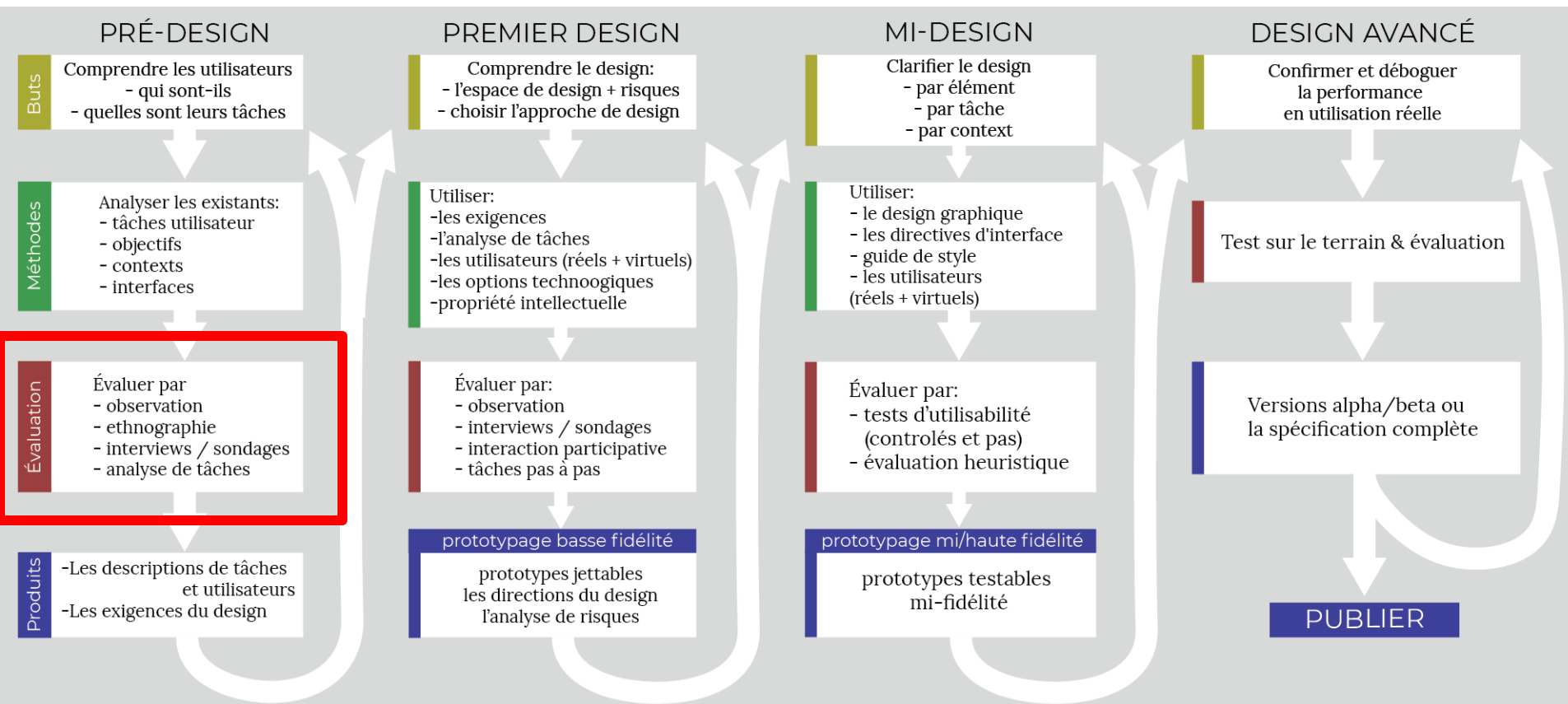
Rappel

# DESIGN CONCEPTUEL



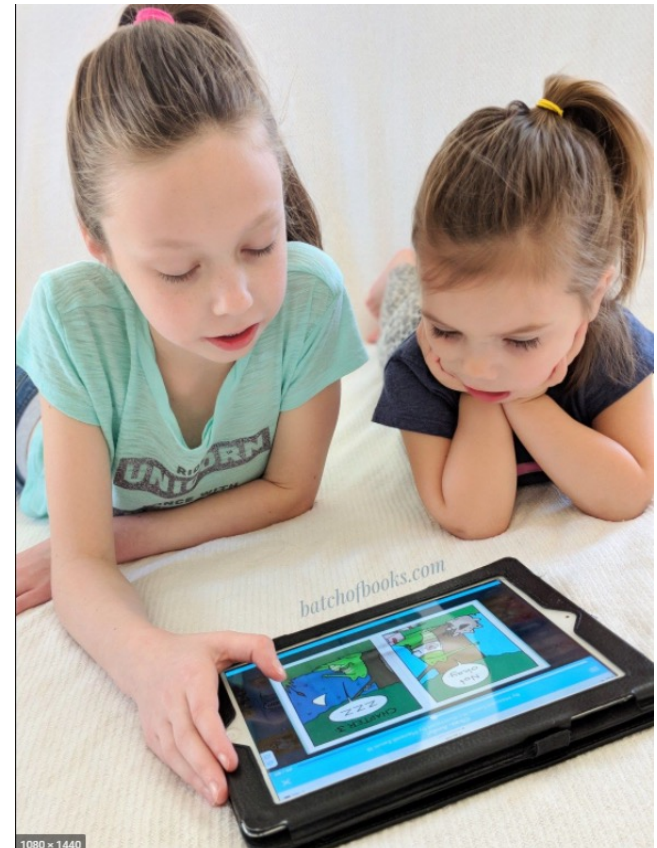
# EXEMPLE

## Ebook pour les enfants

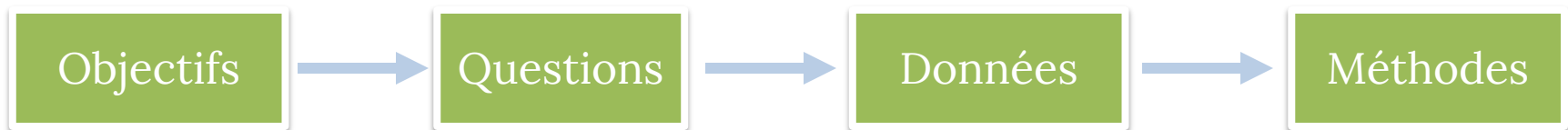


« Amazon souhaite développer une tablette et «apprendre à lire» applications destinées aux enfants. Amazon a déjà une gamme de tablettes et de liseuses électroniques, mais n'a pas encore étudié dans quelle mesure ces produits sont utilisables ou réussis avec les enfants. »

- Les besoins humains?
- Les parties intéressées?
- Les tâches concrètes?
- Qu'est-ce que on veut apprendre?
  - Les objectifs d'évaluation
- Qui seront nos participants?



# COMMENT FAIRE L'ÉVALUATION?



+itérations



## 3 types d'objectifs à pré-design

- **Comprendre quel besoins humains sont/ne sont pas soutenus par les systèmes existants**
  - examiner dans quelle mesure les applications de lecture existantes sont intéressantes pour les enfants
- **Découvrir les problèmes avec l'interface**
  - P.e., comprendre quelles fonctionnalités des lecteurs électroniques existants que les enfants ont du mal à utiliser, et pourquoi
- **Préparer des descriptions de tâches pour ces besoins humains**
  - prioriser et justifier

## Objectifs

- examiner dans quelle mesure les liseuses existantes sont-elles intéressantes pour les enfants
- comprendre quelles fonctionnalités des liseuses existantes les enfants ont du mal à utiliser, et pourquoi

## Questions

*De quelles questions avons-nous besoin pour atteindre ces objectifs?*

- à quelle fréquence et à quels étapes les enfants demandent-ils de l'aide à leurs parents
- comment les enfants aiment les liseuses
- combien de temps les enfants lisent-ils une liseuse par rapport aux livres ordinaires sans s'ennuyer
- ...



*Quels types de données répondraient à ces questions?*

- à quelle fréquence et à quels étapes les enfants demandent-ils de l'aide à leurs parents
- comment les enfants aiment les liseuses
- combien de temps les enfants lisent-ils une liseuse par rapport aux livres ordinaires sans s'ennuyer
- ...



*Quels types de données répondraient à ces questions?*

**quantité**

- **à quelle fréquence** et à quels étapes les enfants demandent-ils de l'aide à leurs parents
- comment les enfants aiment les liseuses
- combien de temps les enfants lisent-ils une liseuse par rapport aux livres ordinaires sans s'ennuyer
- ...

Objectifs



Questions



Données

*Quels types de données répondraient à ces questions?*

*catégorie, histoire*

- à quelle fréquence et à **quels étapes** les enfants demandent-ils de l'aide à leurs parents
- comment les enfants aiment les liseuses
- combien de temps les enfants lisent-ils une liseuse par rapport aux livres ordinaires sans s'ennuyer
- ...



*Quels types de données répondraient à ces questions?*

- à quelle fréquence et à quels étapes les enfants de l'aide à leurs parents
- **Opinion, préférences**
- **comment les enfants aiment les liseuses**
- **combien de temps les enfants lisent-ils une liseuse par rapport aux livres ordinaires sans s'ennuyer**
- ...



*Quels types de données répondraient à ces questions?*

- à quelle fréquence et à quels étapes les enfants demandent-ils de l'aide à leurs parents
- **Quantité/temps passé** ; aiment les liseuses
- **combien de temps** les enfants lisent-ils une liseuse par rapport aux livres ordinaires sans s'ennuyer
- ...



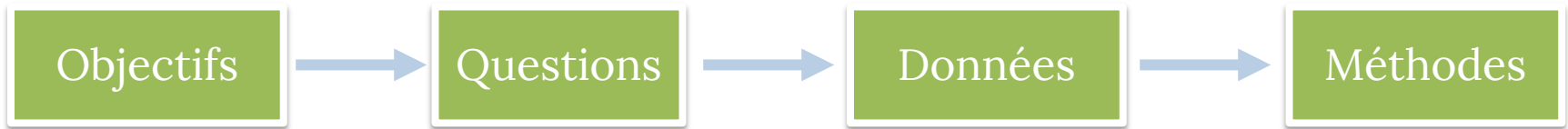
*Quels types de données répondraient à ces questions?*

- à quelle fréquence et à quels étapes les enfants demandent-ils de l'aide à leurs parents
- comment les enfants aiment les liseuses
- combien de temps les enfants lisent-ils une liseuse par rapport aux livres ordinaires sans s'ennuyer
- ...

**Types de données:**

durée, nombre, notes, classement, préférence, catégorie, opinion, histoire, ...





*Quelles méthodes d'évaluation pourraient vous fournir ces types de données*

- Nombre → **observation contrôlée**
- Temps passé → **observation contrôlée, journaux de données (logs), questionnaires/sondages**
- Histoires → **interviews non structurés**
- Préférences → **interviews, questionnaires/sondages**
- Des avis → **interviews, questionnaires/sondages**
- Processus de pensée → **interviews, observation (penser à voix haute)**

# COMMENT CHOISIR LES MÉTHODES?

- Validité interne
  - Répond-elle vraiment à cette question?
- Réalisme = la validité écologique
  - Les résultats s'appliqueront-ils dans le monde réel?
- Validité Externe = Généralisabilité
  - Les résultats s'appliqueront-ils à d'autres situations?
- Approches disruptives (ou pas)
- Temps, coût, compétence, ressources

# EST-CE QUE LES ENFANTS SONT PLUS FATIGUÉS D'ÉCRAN ÉLECTRONIQUE?



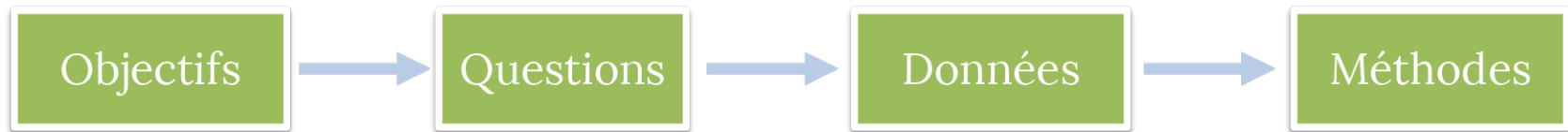
Image © <https://www.scholastic.com/parents/books-and-reading/raise-a-reader-blog/7-ways-to-find-books-your-kids-will-love.html>



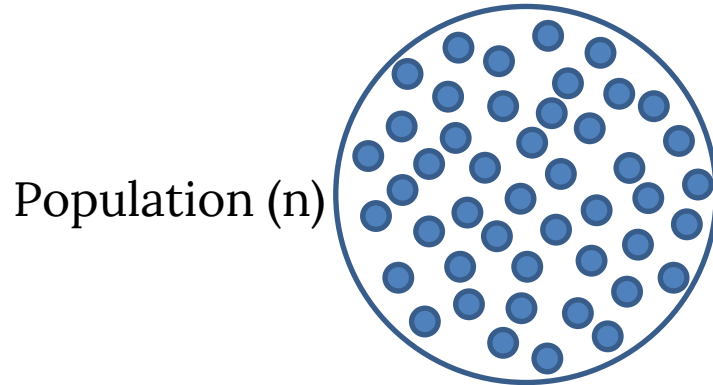
Image © batchofbooks.com

# PROJET PILOTE

- **une étude préliminaire à petite échelle**
  - Vérifier que l'étude fonctionne bien
- **« échouer rapidement, échouer souvent »**
  - « *Fail fast, fail often* »
  - Corriger le plan d'évaluation après le projet pilote
  - Les nouvelles données souvent sont incompatibles avec les anciennes données
- **Peut avoir échantillonnage plus facile**
  - Les participants de projet pilote ne peuvent pas utilisés dans l'étude principale



# POPULATION ET PARAMÈTRES

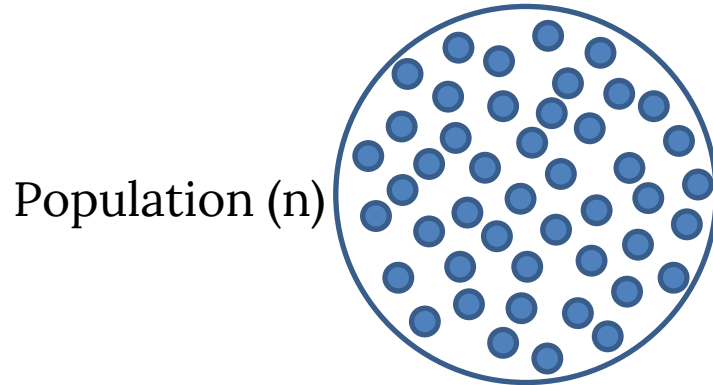


**Population:** un grand ensemble de personnes  
tous les Canadiens âgés entre 20 et 65 ans

**Paramètres:** valeur sommaire (comme une moyenne ou un pourcentage) qui décrit l'ensemble de la population

La moyenne du temps d'accès à la barre de menu  
La proportion entre utilisateur Windows et Mac

# POPULATION ET PARAMÈTRES

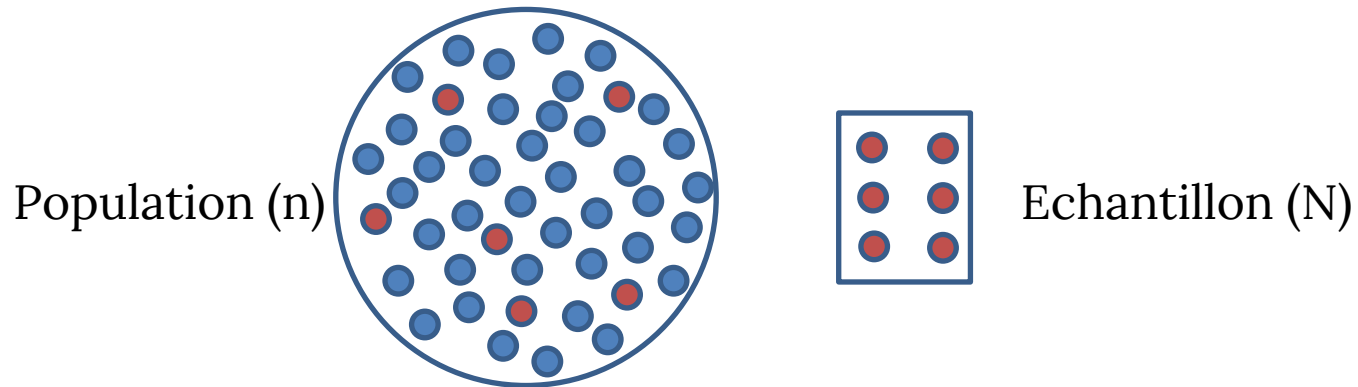


Nous ne pouvons pas connaître la valeur réelle d'un paramètre de la population

impossible de tester toute la population

**=> estimer le paramètre**

# ECHANTILLON ET STATISTIQUE



Echantillon: un groupe représentatif *tiré au sort* parmi la population.

Statistiques: valeur *récapitulative* qui décrit l'échantillon

- *e.g., moyenne ou un pourcentage*

**Rappel:**

# FEUILLES DE CODAGE

## Feuille de codage

Date: \_\_\_\_\_ Temps: \_\_\_\_\_ Observateur: \_\_\_\_\_

# - Sujet 1 : \_\_\_\_\_ # - Sujet 2 : \_\_\_\_\_

Communication verbale	
De quoi discutent les participants ? Quelle est la fréquence des échanges ?	

Stratégies	✓	Remarques
Objectif partagé	<input type="checkbox"/>	
Diviser et conquérir	<input type="checkbox"/>	

Les rôles	✓	Remarques
Même	<input type="checkbox"/>	
Différent	<input type="checkbox"/>	
Text		

Espace de travail	✓	Remarques
Un espace de travail	<input type="checkbox"/>	
Espaces de travail multiples	<input type="checkbox"/>	



**Rappel:**

# FUILLES DE CODAGE

Date: \_\_\_\_\_

Temps: \_\_\_\_\_

Observateur: \_\_\_\_\_

# - Sujet 1 : \_\_\_\_\_

# - Sujet 2 : \_\_\_\_\_

## Communication verbale

De quoi discutent les participants ?

Quelle est la fréquence des échanges ?

**Rappel:**

# FUILLES DE CODAGE

Stratégies	✓	Remarques
Objectif partagé		
Diviser et conquérir		

Les rôles	✓	Remarques
Même		
Différent		
Text		

**Rappel:**

# FEUILLES DE CODAGE

une liste de ce qu'il faut enregistrer

- réduire la subjectivité d'observateur
  - instruire l'observateur
  - faciliter l'analyse quantitative
- 
- n'a pas nécessairement de notes de forme libre (mais celles-ci peuvent être utiles)

Sur [canva.com](https://canva.com), créez une présentation à une seule diapositive avec un fond bleu clair et du texte noir, le texte doit être vertical (de haut en bas), centré, sur une seule ligne.

## L'observateur observe

- Silence!
- Ne pas aider, ne pas expliquer, ne pas montrer les erreurs
- S'asseoir sur ses mains si nécessaire

## Prendre des notes

- Etre attentif aux incidents marquants qui affectent la performance ou la satisfaction
- En général négatif (erreurs, essais répétés, insultes proférées à l'ordinateur)
- Parfois positif ("*Ah ! Maintenant je vois !*", ...)

# LE FACILITATEUR

## Le facilitateur est le responsable du test

- Explique le test
- Propose les tâches
- Incite l'utilisateur à penser à haute voix
  - « À quoi pensez-vous ? »
  - « Pourquoi avez-vous fait ça ? »
- Contrôle la session de test et empêche les observateurs d'interférer

## Il aide les utilisateurs à se sortir des impasses

- mais seulement après avoir noté le problème.