

# IFT 2905: INTERFACES PERSONNE-MACHINE

Évaluation: observation

<http://tiny.cc/ift2905>



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

## Mikhail Bessmeltsev

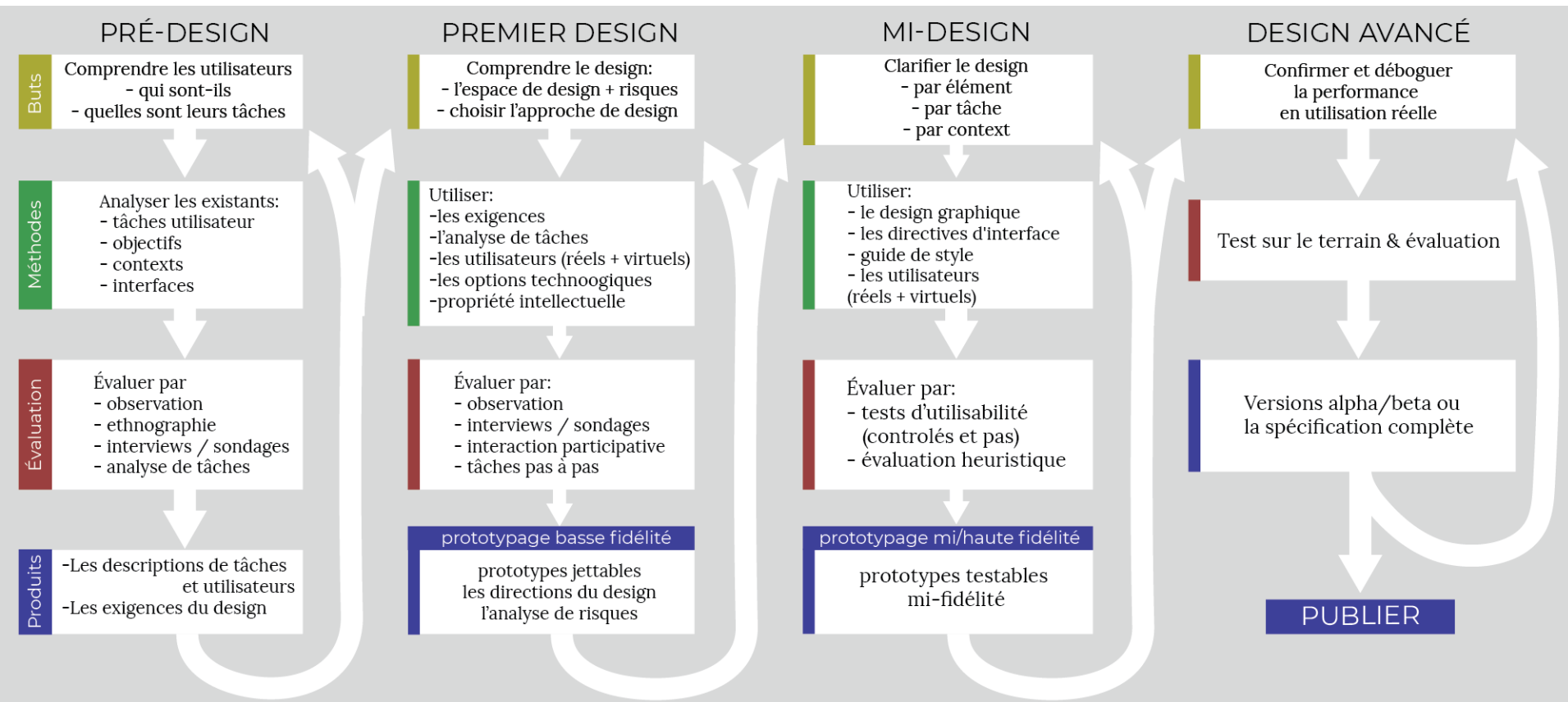
Certaines diapos de Yoon, McLean et Thomaszewski

# CETTE SEMAINE

- Micro Devoir 6
  - Jeudi (18 mars), pendant la demo
  - Jusqu'au **vendredi (19 mars)** minuit
- Devoir 3: jeudi, 18 mars
- Devoir 4: disponible bientôt
  - après 18 mars

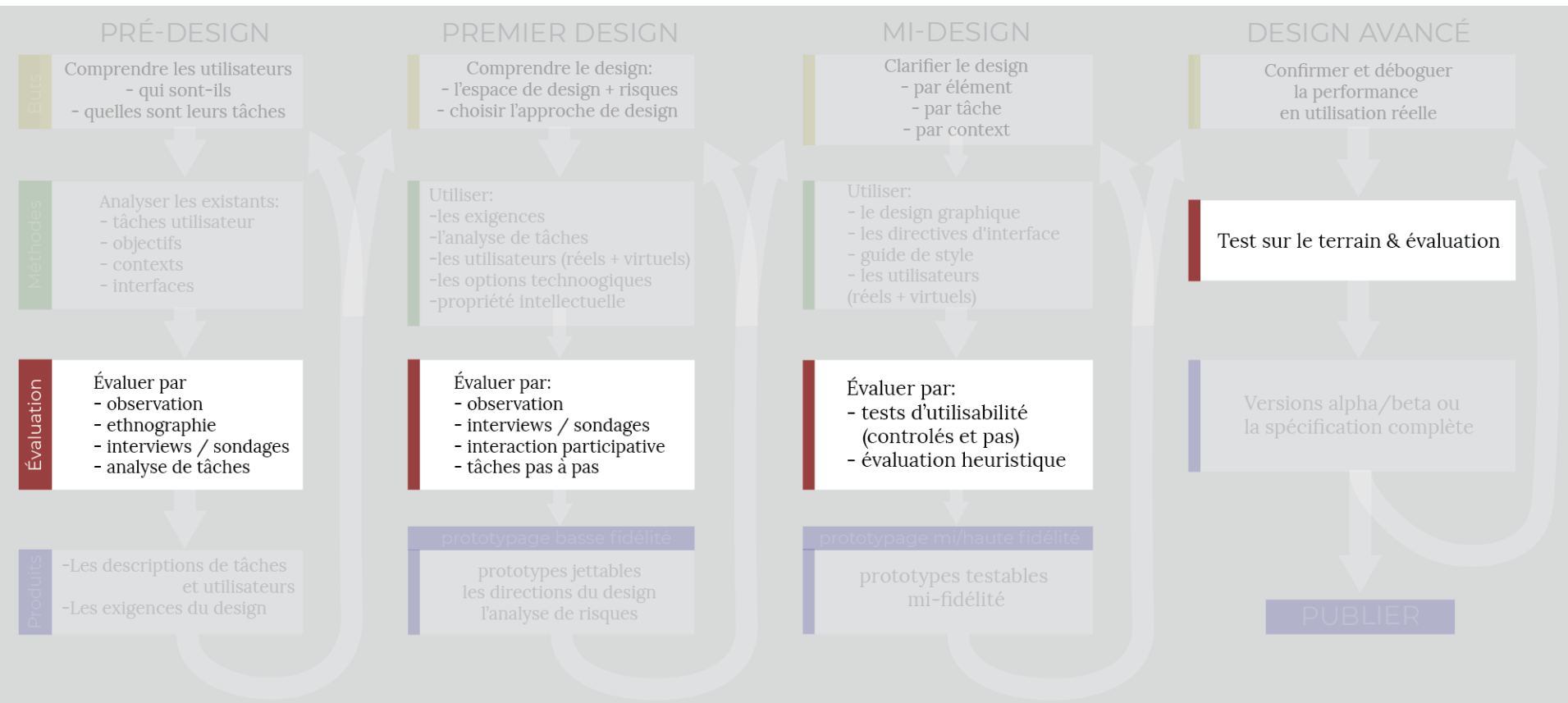
Rappel

# DESIGN CONCEPTUEL



Rappel

# DESIGN CONCEPTUEL





**EXPÉRIENCES SUR LES HUMAINS**

# EXPÉRIENCES SUR LES HUMAINS



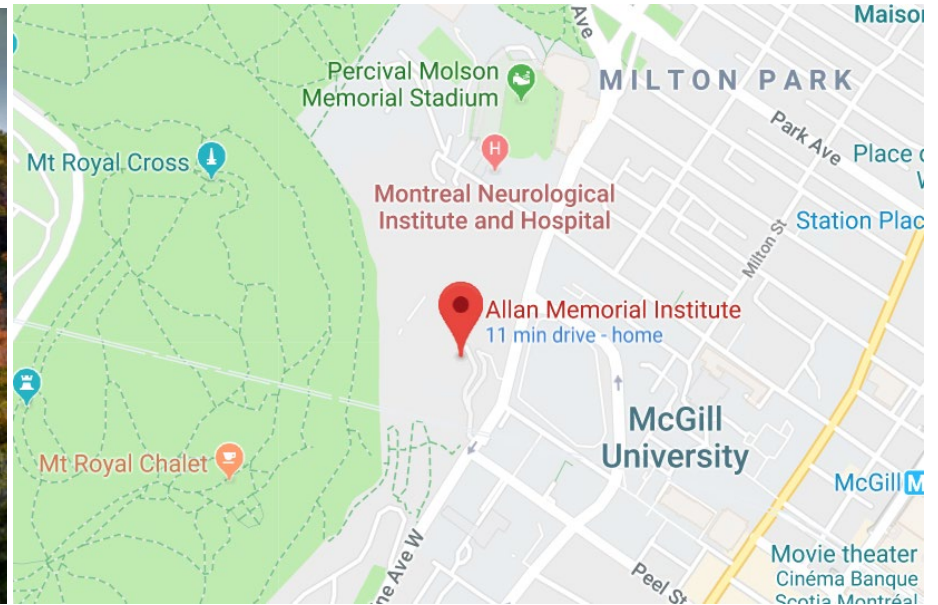
# ETHIQUE

Les usagers sont des êtres humains

Les cobayes humains ont été sérieusement abusés dans le passé

- Camps Nazi (1940-1945)
- Tuskegee syphilis study (1932-1972)
- MIT recherche sur les isotopes radioactif (1940-1960)
- Yale (1950) sur les chocs électriques (Milgram)
- McGill (1950-1962) expériences sur le LSD et la privation sensorielle

# PROJET MK ULTRA





# Declassified:

## Mind Control at McGill

by Julie Vanderperre

**T**he Allan Memorial Institute is located in an ominous mansion, formerly known as Ravenscrag, that looms over Rue McTavish at the foot of Mont Royal. The sinister stone building, [said to be haunted](#), is befitting of the grisly experiment that occurred within its walls from 1957 to 1964: Project MK ULTRA. The Central Intelligence Agency (CIA) mind control project used unconsenting patients to test the effects of sensory deprivation, LSD, electroshock therapy, and other methods of controlling the human psyche. Although it may sound like something out of a dystopian sci-fi novel, these experiments were conducted at McGill, with devastating effects on the patients involved.

<http://www.mcgilltribune.com/mind-control-mcgill-mk-ultra/>

# ETHIQUE

Les usagers sont des êtres humains  
Les cobayes humains ont été sérieusement  
abusés dans le passé

Camps Nazi (1940-1945)

Tuskegee syphilis study (1932-1972)

MIT recherche sur les isotopes radioactif (1940-1960)

Yale (1950) sur les chocs électriques (Milgram)

McGill (1950-1962) expériences sur le LSD et la privation sensorielle

Aujourd'hui, la recherche impliquant des  
humains doit être approuvée par un comite  
d'éthique

Code de Nuremberg

UdeM : Comite universitaire d'éthique de la recherche (CUER)

# CODE DE NUREMBERG



WIKIPÉDIA  
L'encyclopédie libre

## Les expériences médicales acceptables [ modifier | modifier le code ]

La traduction moderne de référence<sup>18</sup> du code de Nuremberg, faite depuis le texte du jugement, est la suivante pour les 10 articles :

1. **Le consentement volontaire du sujet humain est absolument essentiel.** Cela veut dire que la personne concernée doit avoir la capacité légale de consentir ; qu'elle doit être placée en situation d'exercer un libre pouvoir de choix, sans intervention de quelque élément de force, de fraude, de contrainte, de supercherie, de duperie ou d'autres formes sournoises de contrainte ou de coercition ; et qu'elle doit avoir une connaissance et une compréhension suffisantes de ce que cela implique, de façon à lui permettre de prendre une décision éclairée. Ce dernier point demande que, avant d'accepter une décision positive par le sujet d'expérience, il lui soit fait connaître : la nature, la durée, et le but de l'expérience ; les méthodes et moyens par lesquels elle sera conduite ; tous les désagréments et risques qui peuvent être raisonnablement envisagés ; et les conséquences pour sa santé ou sa personne, qui pourraient possiblement advenir du fait de sa participation à l'expérience. L'obligation et la responsabilité d'apprécier la qualité du consentement incombent à chaque personne qui prend l'initiative de, dirige ou travaille à l'expérience. Il s'agit d'une obligation et d'une responsabilité personnelles qui ne peuvent pas être déléguées impunément ;
2. L'expérience doit être telle qu'elle produise des résultats **avantageux pour le bien de la société, impossibles à obtenir par d'autres méthodes** ou moyens d'étude, et pas aléatoires ou superflus par nature ;
3. L'expérience doit être construite et fondée de façon telle sur les résultats de l'expérimentation animale et de la connaissance de l'histoire naturelle de la maladie ou autre problème à l'étude, que les résultats attendus justifient la réalisation de l'expérience ;
4. L'expérience doit être conduite de façon telle que soient **évités toute souffrance** et toute atteinte, physiques et mentales, non nécessaires ;
5. Aucune expérience ne doit être conduite lorsqu'il y a une raison a priori de croire que la **mort ou des blessures invalidantes** surviendront ; sauf, peut-être, dans ces expériences où les **médecins expérimentateurs servent aussi** de sujets ;
6. Le niveau des risques devant être pris ne doit jamais excéder celui de l'importance humanitaire du problème que doit résoudre l'expérience ;
7. Les dispositions doivent être prises et les moyens fournis pour protéger le sujet d'expérience contre les éventualités, même ténues, de blessure, infirmité ou décès ;
8. Les expériences ne doivent être pratiquées que par des personnes scientifiquement qualifiées. Le plus haut degré de compétence professionnelle doit être exigé tout au long de l'expérience, de tous ceux qui la dirigent ou y participent ;
9. Dans le déroulement de l'expérience, le sujet humain **doit être libre de mettre un terme à l'expérience** s'il a atteint l'état physique ou mental où la continuation de l'expérience lui semble impossible ;
10. Dans le déroulement de l'expérience, le scientifique qui en a la charge doit être **prêt à l'interrompre à tout moment**, s'il a été conduit à croire — dans l'exercice de la bonne foi, de la compétence du plus haut niveau et du jugement prudent qui sont requis de lui — qu'une continuation de l'expérience pourrait entraîner des blessures, l'invalidité ou la mort pour le sujet d'expérience.

# L'UTILISATEUR SOUS PRESSION

L'utilisateur est le centre d'attention

observé par des étrangers ou une camera  
doit faire des tâches qui ne lui sont pas familières devant une interface qui  
peut être mauvaise

L'utilisateur peut ressentir différentes formes  
de pression

Angoisse de performance  
Impression de passer un test d'intelligence  
Comparaison avec les autres  
Se sentir stupide devant des observateurs

La pression est donc bien réelle

il y aura une influence sur les résultats

# CONSENTEMENT

- **Volontaire**
  - volontairement, sans contrainte ni pression pour consentir.
- **Éclairé**
  - Les participants doivent comprendre l'objectif de la recherche, ses risques prévisibles et des avantages potentiels.
- **En cours**
  - Nouvelle information → « Vous voulez continuer? »
  - Peut arrêter n'importe quand

# CONFIDENTIALITÉ ET SÉCURITÉ

Les participants peuvent s'attendre à ce que toute information personnelle qui pourrait être utilisée pour les identifier:

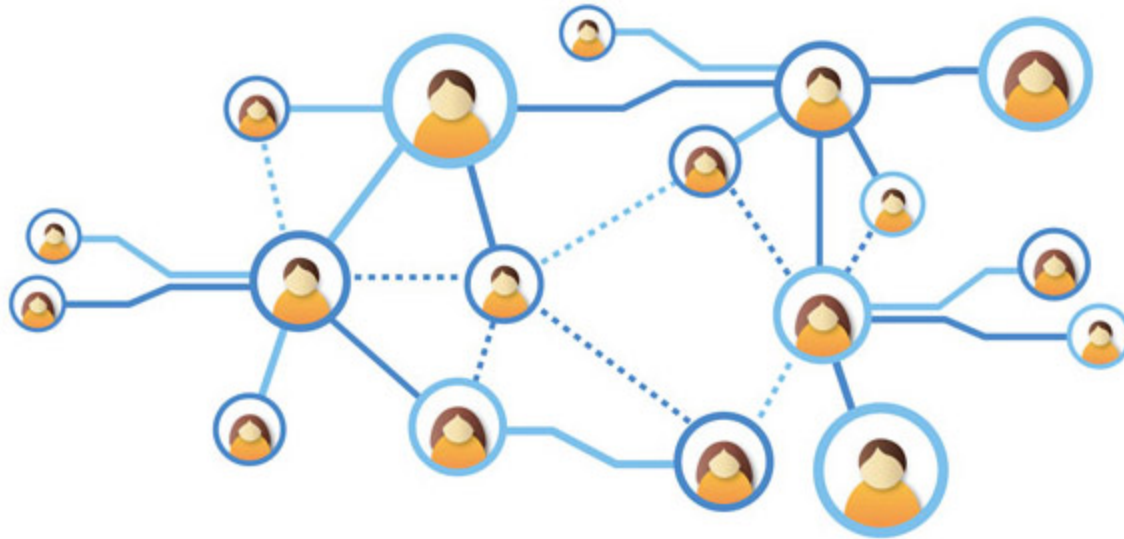
- ne sera pas lié à d'autres informations les concernant
  - par exemple, les dossiers de santé, d'emploi ou d'éducation
- ne sera pas être mis à la disposition de quiconque externe
- sera stocké en toute sécurité

# CONFIDENTIALITÉ ET SÉCURITÉ

## Catégories d'information:

- Informations d'identification directe
  - Nom, numéro d'assurance sociale, etc.
- Informations d'identification indirecte
  - « Si j'en collectionne une grande partie, je peux l' identifier »
  - date de naissance, lieu de résidence ou caractéristique personnelle unique
- Informations codées
  - Identifiants directs remplacés par des codes
  - Codes stockés séparément
- Informations anonymisées
  - Codes ne sont pas stockés
- Informations anonymes
  - Identifiants n'ont jamais été collectés

# EXAMPLE





# De-anonymizing Social Networks

Arvind Narayanan and Vitaly Shmatikov  
The University of Texas at Austin

## Abstract

*Operators of online social networks are increasingly sharing potentially sensitive information about users and their relationships with advertisers, application developers, and data-mining researchers. Privacy is typically protected by anonymization, i.e., removing names, addresses, etc.*

*We present a framework for analyzing privacy and anonymity in social networks and develop a new re-identification algorithm targeting anonymized social-network graphs. To demonstrate its effectiveness on real-world networks, we show that a third of the users who can be verified to have accounts on both Twitter, a popular microblogging service, and Flickr, an online photo-sharing site, can be re-identified in the anonymous Twitter graph with only a 12% error rate.*

*Our de-anonymization algorithm is based purely on the network topology, does not require creation of a large number of dummy “sybil” nodes, is robust to noise and all existing defenses, and works even when the overlap between the target network and the adversary’s auxiliary information is small.*

associated with individual nodes are suppressed. Such suppression is often misinterpreted as removal of “personally identifiable information” (PII), even though PII may include much more than names and identifiers (see the discussion in Appendix B). For example, the EU privacy directive defines “personal data” as “any information relating to an identified or identifiable natural person [...]; an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity” [Eur95].

Anonymity has been unquestioningly interpreted as equivalent to privacy in several high-profile cases of data sharing. After a New York court ruling ordering Google to hand over viewing data of over 100 million YouTube users to Viacom and the subsequent protests from privacy advocates, a revised agreement was struck under which Google would anonymize the data before handing it over [Swe08]. The CEO of NebuAd, a U.S. company that offers targeted advertising based on browsing histories gathered from ISPs, dismissed privacy concerns by saying that “We don’t have any raw data on the identifiable individual. Everything is

# CONFLIT D'INTÉRÊT

« Si vous étiez dans une étude comparant la convivialité de deux logiciels ... et que vous entendiez que le chercheur avait investi de l'argent dans l'un d'entre eux, feriez-vous confiance aux résultats de l'étude? »

# PLUS D'INFORMATION

<https://tcps2core.ca/welcome>

# RESPECT DE L'UTILISATEUR

L'utilisateur doit être traité avec respect

- Ne pas gaspiller son temps
- Préparer tout d'avance

Rendre l'usager confortable

- Cela réduit beaucoup la pression ressentie

Informé le mieux possible l'usager

- Mais attention de ne pas biaiser les tests
- Ne pas cacher inutilement des choses

Préserver la vie privée de l'usager

- Ne pas publier les résultats en identifiant les usagers

L'usager doit pouvoir arrêter à tout moment

# AVANT LE TEST

## Temps

- Faire un test pilote du matériel et des tâches

## Confort

- *"Ce n'est pas vous qu'on teste, c'est l'interface."*
- *"Si vous avez un problème, c'est la faute du système. C'est justement ce qu'on cherche à trouver."*

## Vie privée

- *"Les résultats seront confidentiels."*

## Information

- Résumer le but du test
- Informer sur l'enregistrement et sur les observateurs
- Répondre aux questions avant de commencer

## Contrôle

- *"On arrête quand vous voulez."*

# PENDANT LE TEST

## Temps

- Éliminer les tâches inutiles

## Confort

- Atmosphère calme
- Prendre des pauses
- Ne jamais paraître déçu (de l'interface)
- Donner les tâches une à la fois
- Commencer par les tâches faciles

## Vie privée

- Le patron/directeur/professeur de l'utilisateur ne devrait pas assister

## Information

- Répondre aux questions (si pertinentes)

## Contrôle

- L'utilisateur choisit de continuer ou non

# APRÈS LE TEST

## Confort

Mentionner à l'utilisateur en quoi sa participation a été utile

## Information

Répondre aux questions qui auraient pu biaiser l'expérience ("*C'est quoi ce bouton?*")

## Vie privée

Protéger l'identité des utilisateurs dans les résultats

Ne pas envoyer les vidéos sur YouTube

# OBSERVATION

Pourquoi évaluer?

- Chaque étapes de processus du design a son propre but
- Les méthodes d'évaluation doivent aider ce but
- Évaluer tout au long du processus

Parties intéressées différents → besoins différents

Beaucoup de méthodes → différentes **forces** et **faiblesses**



# OBSERVATION DIRECTE

Trois approches communes:

- observation simple
- penser à voix haute
- deux utilisateurs (apprentissage par co-découverte)

+ objectif: on voit ce que l'utilisateur fait

+ souple: peut être fait dans un contexte réel ou contrôlé

+ beaucoup de données

-le fait que l'utilisateur est observé change son comportement

-beaucoup de données à analyser

-couteux et chronophage

# OBSERVATION SIMPLE

marshmallow kid experiment



- L'utilisateur reçoit une tâche
- Le facilitateur regarde juste

IGNITERMEDIA.COM

# OBSERVATION SIMPLE

marshmallow kid experiment



- L'utilisateur reçoit une tâche
- Le facilitateur regarde juste

- Aucune compréhension de la façon dont l'utilisateur pense ou pourquoi

IGNITERMEDIA.COM

# PENSER À VOIX HAUTE

L'utilisateur doit penser à voix haute

- Dire ce qu'il pense
- Dire ce qu'il veut faire
- Dire pourquoi il a fait chaque action

Très populaire!

- L'usager se sent inconfortable
- Penser a haute voix peut modifier le comportement
- Rend la concentration plus difficile



# DEUX UTILISATEURS

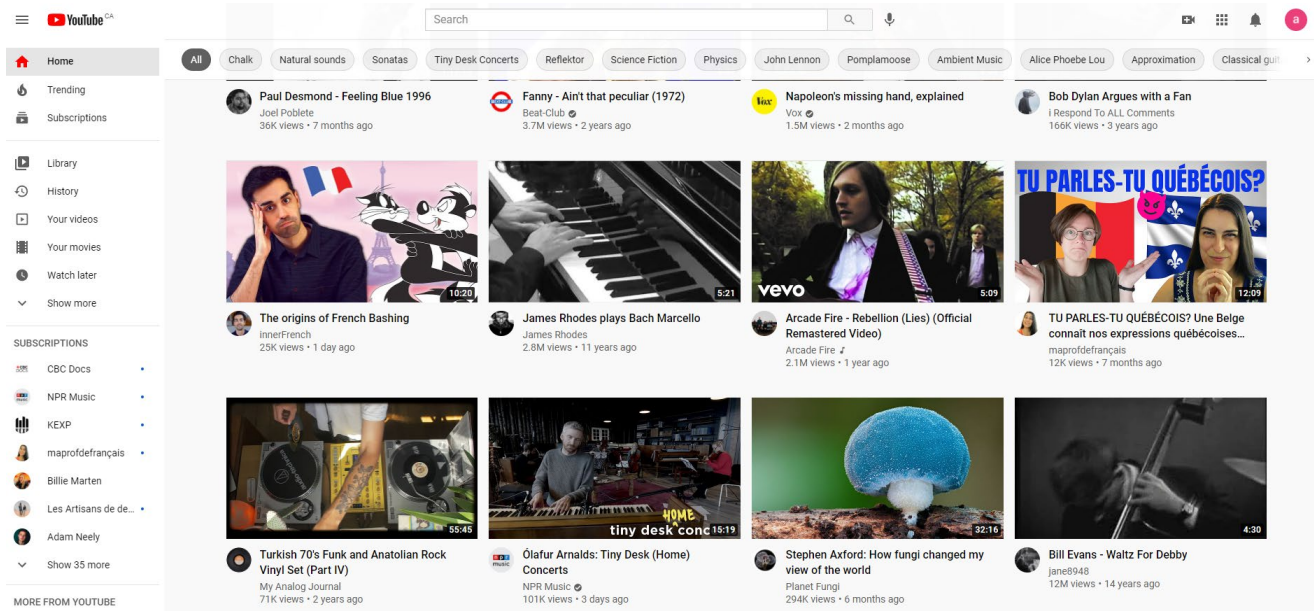
*apprentissage par co-découverte*

- Deux utilisateurs vont probablement se parler naturellement
- Aussi appelé co-découverte, ou interaction constructive

presque aussi commun que l'approche  
avec un seul usager

# OBSERVATION À DISTANCE

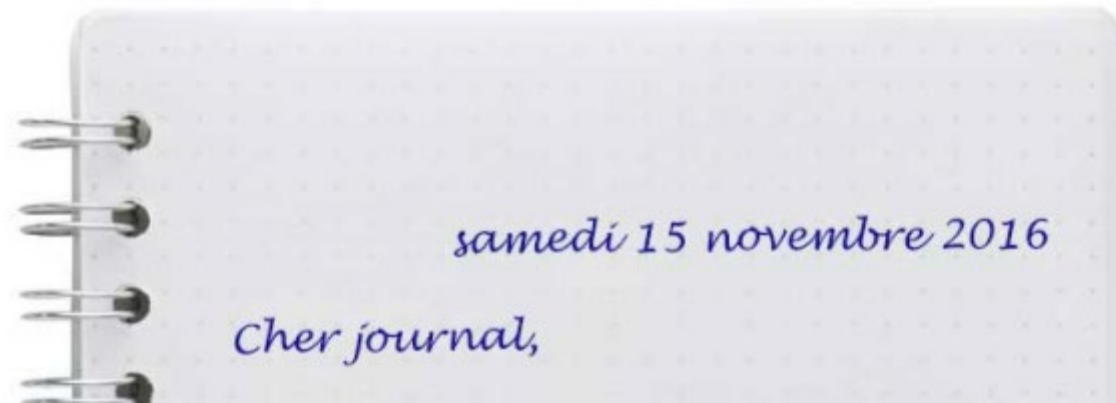
- Nécessaire pour tous maintenant (IFT2905!)
- L'écran d'enregistrement nécessite un consentement
- Vérifier que l'audio fonctionne
- Respectez la vie privée de l'utilisateur



# OBSERVATION INDIRECTE

- journaux et autodéclarations
  - + à distance
  - + moins de problèmes de confidentialité
  - biaisé
  - Certains choses peut être oubliées
- journaux d'interaction (logs)
  - + discret
  - beaucoup de données à analyser

Habituellement utilisé avec d'autres techniques



# ENREGISTREMENT DES OBSERVATIONS

## Crayon et papier

- un formulaire préparé d'avance peut être utile

## Enregistrement audio

- pratique pour *penser à haute voix*

## Enregistrement vidéo

- Souvent 2 cameras (visage de l'utilisateur, et écran)
- L'utilisateur peut ne pas aimer être filmé
- Parfait pour des observateurs dans une pièce séparée
- Crée trop de donnée
- Possibilité de retour en arrière pour identifier des moments critiques

## Capture d'écran et journal des événements

- simple, pas cher, et ne dérange pas l'utilisateur



# LES RÔLES

En expérience contrôlée, il y a trois rôles distincts

- L'utilisateur
- Le facilitateur
- Observateurs

# LE FACILITATEUR

Le facilitateur est le responsable du test usager

- Explique le test
- Propose les tâches
- Incite l'utilisateur a penser a haute voix
  - "*A quoi pensez-vous ?*"
  - "*Pourquoi avez-vous fait ca ?*"
- Contrôle la session de test et empêche les observateurs d'interférer

Il aide les utilisateurs a se sortir des impasses

- mais seulement après avoir noté le problème.

# L'OBSERVATEUR

## 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 a l'ordinateur)
- Parfois positif ("Ah ! Maintenant je vois !", ...)

# TYPES DE DONNÉES

- Qualitatives
  - Répondent aux questions: Pourquoi? Comment?
  - Les mots, les images, les histoires
- Quantitatives
  - Répondent aux questions: Est-ce vrai? Combien?
  - Les nombres, tables, graphiques

Planifier quel données à enregistre et stocker

# FEUILLES DE CODAGE

Coding Sheet

Date: \_\_\_\_\_ Time: \_\_\_\_\_ Observer: \_\_\_\_\_

# - Subject 1: \_\_\_\_\_ # - Subject 2: \_\_\_\_\_

Verbal Communication	
What do participants discuss? How often are the exchanges?	

Strategies	✓	Notes
Shared Goal	<input type="checkbox"/>	
Divide and Conquer	<input type="checkbox"/>	

Roles	✓	Notes
Same	<input type="checkbox"/>	
Different Text	<input type="checkbox"/>	

Workspace	✓	Notes
One workspace	<input type="checkbox"/>	
Multiple Workspaces	<input type="checkbox"/>	

# 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)

# FEUILLES DE CODAGE

*Comment les créer?*

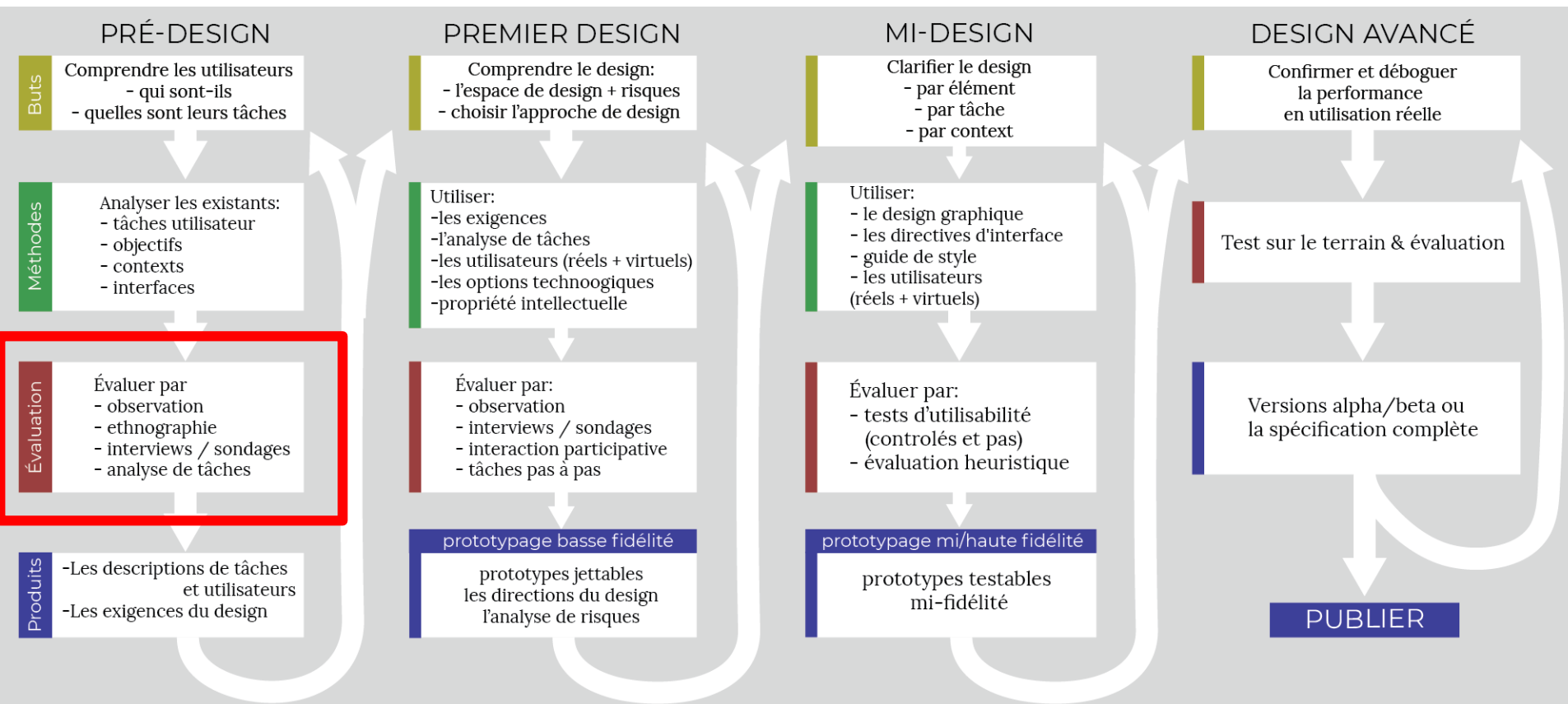
Ça dépend:

- Des objectifs de l'évaluation
- De l'étape de design
- De la méthode d'observation
- Des types de données

**Pilotez-les!**

# EXEMPLE

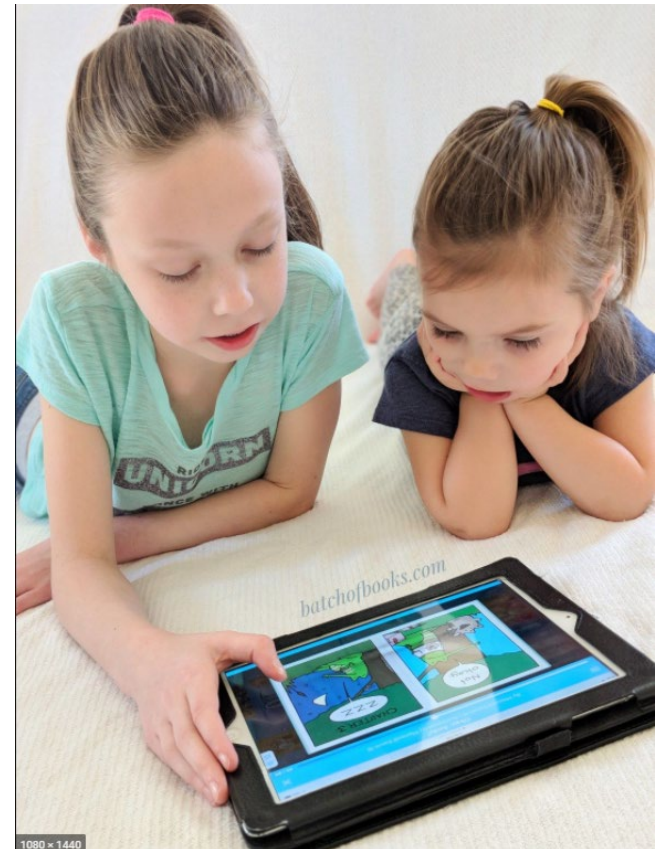
## Ebook pour les enfants



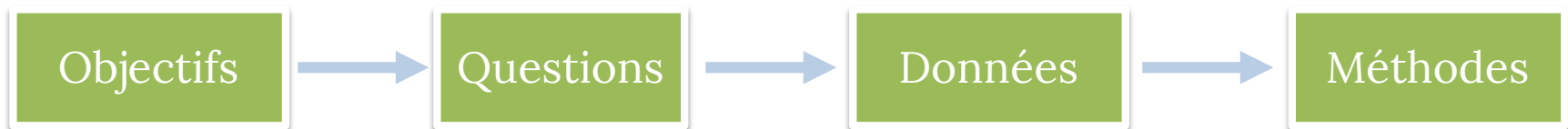


« Amazon souhaite développer une tablette et «apprendre à lire» applications destinées aux enfants. Amazon a déjà une gamme de tablettes et de liseuses électroniques, mais n'a pas encore étudié dans quelle mesure ces produits sont utilisables ou réussis avec les enfants. »

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



# COMMENT FAIRE L'ÉVALUATION?



+itérations

## 3 types d'objéctifs 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
- Découvre 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 intéressantes pour les enfants
- comprendre quelles fonctionnalités des liseuses existantes que les enfants ont du mal à utiliser, et pourquoi

## Questions

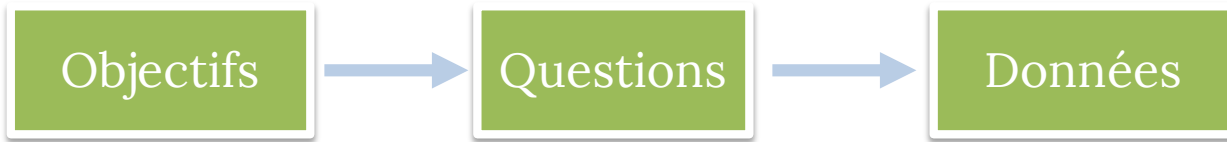
*De quelles questions avons-nous besoin pour atteindre ces objectifs?*

- 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
- ...



*Quels types de données répondraient à ces questions?*

*catégorie, histoire*

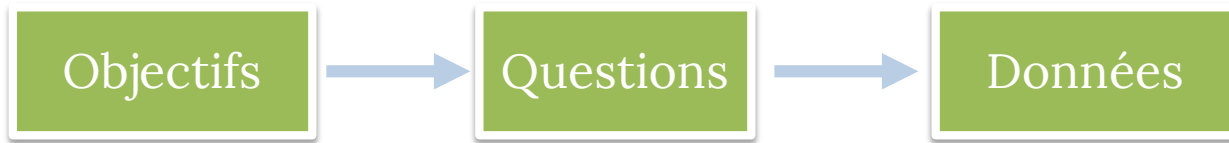
- à quelle fréquence et à **quels étapes** les enfants demandent-ils de l'aide à leurs parents
- comment les enfants aiment les liseuses
- combien de temps les enfants lisent-ils une liseuse par rapport aux livres ordinaires sans s'ennuyer
- ...



*Quels types de données répondraient à ces questions?*

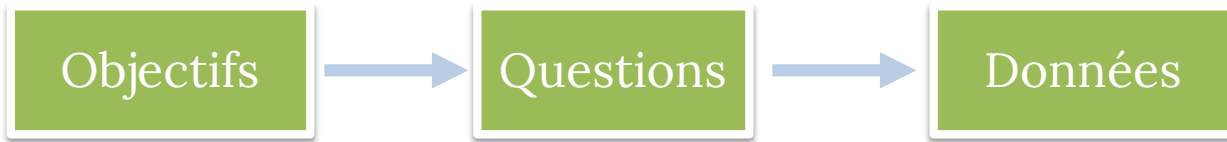
- à quelle fréquence et à quels étapes les enfants de l'aide à leurs parents
- **Opinion, préférences**
- **comment** les enfants aiment les liseuses
- combien de temps les enfants lisent-ils une liseuse par rapport aux livres ordinaires sans s'ennuyer
- ...





*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
- ...

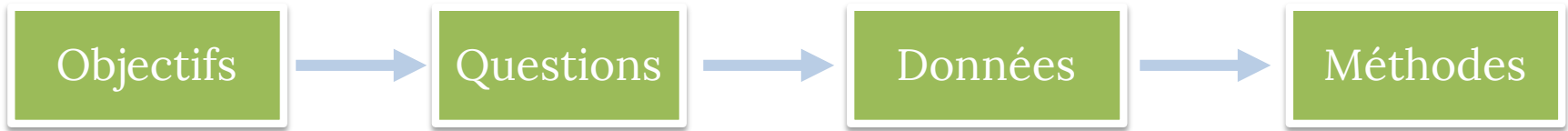


*Quels types de données répondraient à ces questions?*

- 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?

Il faut prendre en compte:

- **Validité** (Répond-elle vraiment à cette question?)
- **Réalisme** (les résultats s'appliqueront-ils dans le monde réel?)
- **Généralisabilité** (les résultats s'appliqueront-ils à d'autres situations?)
- Approches disruptives (ou pas)
- Temps, coût, compétence, ressources

# 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 « nouveaux » données souvent sont incompatibles avec les « vieux » données
- Peut avoir échantillonnage plus facile
  - Les participants de projet pilote ne peuvent pas utilisés dans l'étude principale

