1

Conception d'expériences

- Expériences contrôlées
- Quasi-Expériences
- Etudes d'utilisabilité

Michel Beaudouin-Lafon
Université Paris-Sud

Cours de Wendy E. Mackay (I.N.R.I.A.)

2

Expériences contrôlées

But :
Est-ce que le traitement X cause l'effet Y?

Facteur clé :
Affectation aléatoire à des groupes

Test d'hypothèse
Comparer des hypothèses alternatives
Contrôler les conditions pour isoler les variables

Analyse de corrélations
Mesurer le degré de corrélation entre deux facteurs
Connaître l'un aide à prédire l'autre

3

A quoi peuvent servir les expériences

Explorer une nouvelle technique d'interaction
Interaction bi-manuelle (Buxton)

Choisir entre des alternatives de conception
Menus "pop-up" ou déroulants

Evaluer un système
Les utilisateurs peuvent-ils effectuer une certaine tâche ?

Comparer deux systèmes
Mac vs. PC

Comparer des techniques d'évaluation
"Design walkthroughs" vs. GOMS

4

Conception classique d'expériences

Formuler des modèles (ou des théories)
Décider de leur application restreinte ou générale

Faire des hypothèses explicites
Opérationnaliser le comportement
Prédire les résultats

Déterminer le sens statistique
Degré de confiance dans les résultats

1

Concevoir une expérience simple

1. **Spécifier l'hypothèse**
   Que comparer-t-on ?
2. **Spécifier les variables indépendantes**
   Qu'est-ce qui varie ?
3. **Opérationnaliser le comportement**
   Qu'est-ce qu'on étudie ?
4. **Spécifier les variables dépendantes**
   Qu'est-ce qu'on mesure ?
5. **Spécifier les procédures**
   Quels sont les groupes expérimentaux et de contrôle ?
6. **Identifier les tests statistiques appropriés**
   Y a-t-il une différence ?

2

Exemple d'hypothèse

Comparer les menus pop-up et déroulants

Il n'y a pas de différence de performance

entre les utilisateurs

en temps et taux d'erreur

pour la sélection d'un item dans un menu pop-up ou un menu déroulant,

quelle que soit l'expérience antérieure d'utilisation de la souris ou d'autres types de menus par l'utilisateur.

3

Spécifier les variables indépendantes

Varier les items systématiquement pour créer différentes conditions

Variables indépendantes :

Type de menu :	pop-up, déroulant
Nombre d'items :	3, 6, 9, 12, 15
Sujet :	expert, novice, intermittent

4

Spécifier les variables dépendantes

Identifier des éléments mesurables qui dépendent du comportement du sujet

Variables dépendantes :

- Temps pour sélectionner un item
- Erreurs

1

Affecter les sujets

- But : Contrôler la variabilité des sujets
- Clé : Affectation aléatoire à des groupes
- Reconnaître les classes et en faire une variable indépendante
  - Exemple : degré d'expertise des sujets
- Minimiser les différences entre groupes
  - Exemple : distribution homogènes des "bons" sujets
- Utiliser suffisamment de sujets
  - Déterminé par le type de test statistique, le degré de confiance nécessaire et la variabilité des données

2

Conduire l'expérience

- PRÉDIRE:
  - Ecrire ce que l'on attend comme résultat !
- Contrôler les facteurs de biais
  - Utiliser des instructions identiques
  - L'expérimentateur ne doit pas savoir de quel groupe il s'agit
    - Expérimentation en "double aveugle"
  - Minimiser les contacts informels
- Garder les mêmes conditions d'un sujet à l'autre

3

Utiliser des méthodes statistiques

- S'assurer que le test statistique est valide
  - Distribution de population (normale)
  - Type de données (ordinale)
  - Procédure d'échantillonage (aléatoire)
  - Taille de l'échantillon (n=30)
- Déterminer le degré de confiance de vos résultats
  - "L'hypothèse selon laquelle l'expérience de l'utilisation de la souris ne fait pas de différence est rejetée avec un niveau de 0.05"
- Interpréter vos résultats

4

Problèmes de fiabilité

- Les résultats seraient-ils les mêmes si l'expérience était répétée ailleurs ?
- Qu'en est-il des differences individuelles ?
  - Si un utilisateur est 10 fois plus rapide que les autres ?
- Utiliser un nombre "raisonnable" de sujets
- Utiliser des statistiques appropriées et déterminer le degré de confiance
  - p = .05 : on est sûr à 95% que ces résultats sont dus à une différence réelle entre le groupe expérimental et le groupe de contrôle

1

Problèmes de validité

- L'expérience mesure-t-elle quelque chose de pertinent en dehors du laboratoire ?
- Les utilisateurs sont-ils typiques ?
- Les tâches sont-elles typiques ?
- L'environnement physique est-il différent?
- L'influence sociale est-elle différente ?

2

"Between-Subjects Design"

- Choisir une population
  - Sélectionner aléatoirement 2 groupes
- Les exposer tous les deux à des conditions identiques, MAIS
  - Le groupe expérimental reçoit la variable indépendante, pas le groupe de contrôle
- Suppositions :
  - D'autres variables non contrôlées sont distribuées aléatoirement entre les deux groupes
  - La seule différence systématique entre les deux groupes est la variable indépendante
- Utiliser des tests statistiques appropriés pour distinguer les deux groupes

3

"Within-Subjects Design"

- Choisir une population
  - Exposer chaque sujet successivement à toutes les valeurs de la variable indépendante
  - Contrôle automatique de la plupart des autres variables
  - Permet d'utiliser un nombre plus faible de sujets
- Utiliser des tests appropriés pour distinguer les valeurs de la variable indépendante

4

L'apprentissage est difficile à mesurer

- Expériences contrôlées plus difficiles à concevoir
  - Expérience de type "ABA" impossible
  - Effets de "primacy" et "recency"
- Expérience "between-subjects" plus probable
- Pré-test et post-test nécessaires
  - Differentiate between background and material learned

1

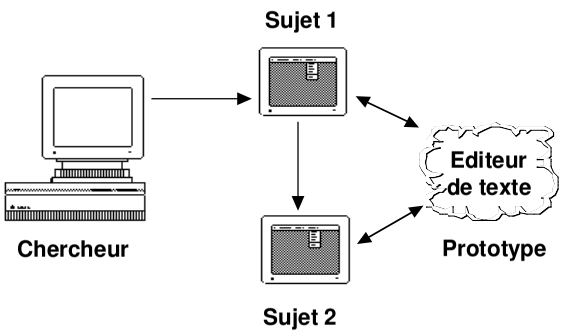
Expérience complexe

Evaluer l'efficacité d'un Tuteur Electronique

- **Expérience à la fois "within-subjects" et "between-subjects"**
- **Inclut la technique du Magicien d'Oz**
- **Contrôlée selon plusieurs dimensions :**
  - **Selon les tâches d'un sujet**
  - **Entre les tâches de paires de sujets**
  - **Entre les tâches de groupes de sujets**

2

Contrôle couplé



3

Résumé

- **Expériences contrôlées :**
  - **Manipuler des variables indépendantes**
  - **Observer les effets sur les variables dépendantes**
  - **Tirer des conclusions à partir de tests statistiques significatifs**
  - **Permettent un plus grand contrôle mais peuvent ne pas être généralisables à des situations réelles**

4

Exemples de sources d'erreurs

- **Sujets**
  - **Le sujet s'attend à un résultat particulier**
  - **Le sujet déforme sa réponse en fonction de connaissances préalables**
- **Expérimentateurs**
  - **Le sujet est influencé par l'expérimentateur**
    - **Utiliser le Double Aveugle**
- **Situation**
  - **Mesurer un phénomène différent**
  - **Désaccord entre observations répétées**

1

Problèmes classiques des expériences contrôlées

- **Mesure-t-on ce que l'on croit mesurer ?**
  - Evaluation du test de lecture par Sidman
- **Y a-t-il un effet dû aux conditions expérimentales ?**
  - Effet Hawthorne, Hopitâl, Placebo

2

Quasi-Expérience

- **Situation:**
  - Avoir des traitements, des mesures de résultats et des unités expérimentales
  - Affectation aléatoire impossible
- **Problème:**
  - Comment faire de l'inférence causale dans un contexte réel ?
- **Solution:**
  - Rendre explicites les "menaces d'invalidité"

3

Quasi-Expérience

- **Fondée sur les mêmes suppositions que les statistiques traditionnelles**
- **Identifier les sources potentielles d'interprétation erronée**
- **Trianguler, trianguler, trianguler !**

4

Etudes d'utilisabilité

- **Adaptations de différentes techniques à l'Interaction Homme-Machine**
- **Peut être quantitative ou qualitative**
- **Peut être formelle ou informelle**
- **Exemples:**
  - "Design Walkthroughs"
  - "Claims analysis"
  - Evaluation heuristique
  - GOMS
  - QOC: Questions, Options, Critères

Faire une hypothèse

But :

- Choisir entre des alternatives

Procédure

- Décrire l'objectif de conception
- Décrire trois méthodes d'implementation
- Quels sont les compromis entre les méthodes ?
- Décrire une expérimentation :
  - Variables indépendantes ?
  - Variables dépendantes ?
  - Hypothèse nulle ?
  - Prédiction ?

1

2

Faire une expérimentation

But :

- Déterminer le meilleur choix de conception

Procédure

- Spécifier l'hypothèse nulle
- Opérationaliser les variables dépendantes
- Identifier les variables indépendantes
- Préparer l'environnement pour tester chaque condition
- Utiliser au moins 3 sujets
- Analyser les résultats
  - Les différences sont-elles significatives ?

3

4