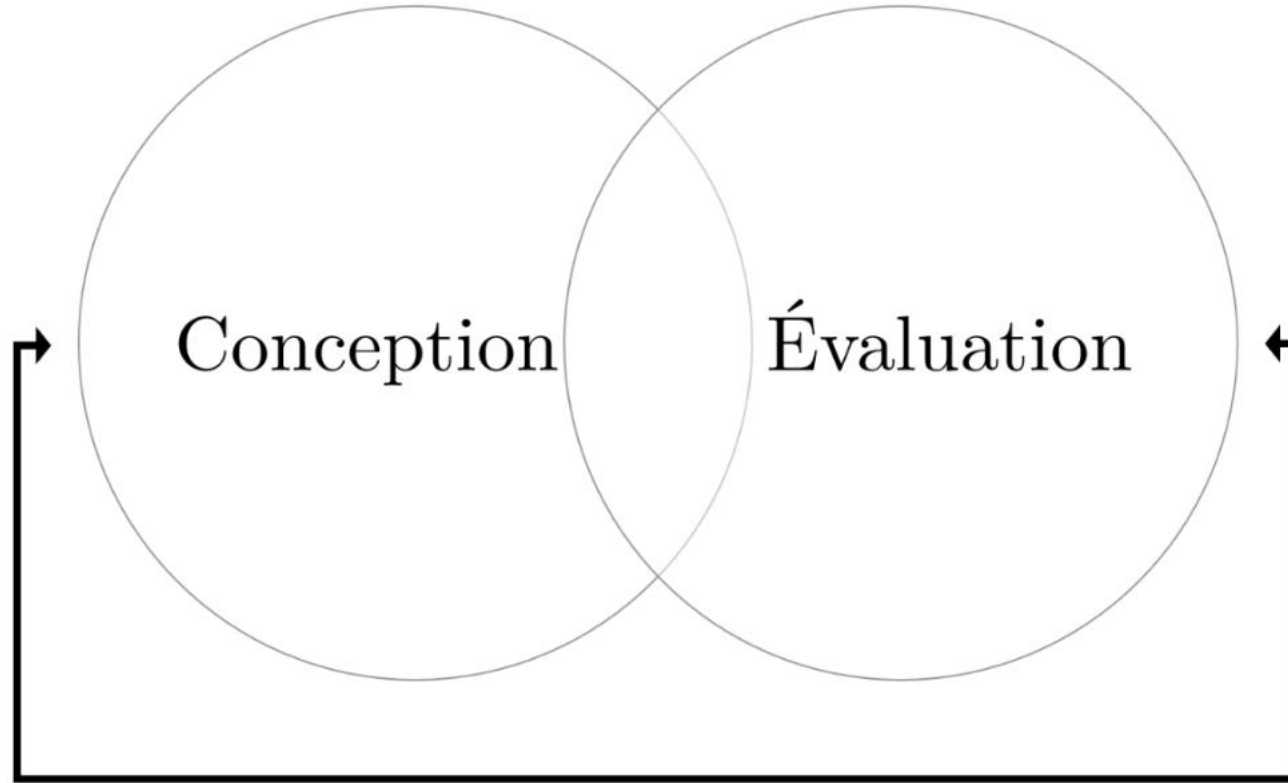


# Techniques d'évaluation en UX

Mattia A. Fritz

TECFA, Université de Genève

# Deux grandes fonctions



Un team/professionnel/projet de UX peut s'occuper exclusivement/principalement de l'une ou de l'autre, ou proposer les deux en même temps.

# Évaluation

L'évaluation mesure l'adéquation entre les attentes des concepteurs/stakeholders d'une part, et l'utilisation et perception du public cible de l'autre.



Performance

Perception

# Quelques méthodes

## ➤ Oculométrie (*Eye-tracking*)

Mesure physiologique de type performance.

## ➤ Test A/B

Mesure statistique (souvent) de type performance.

## ➤ Test des 5 seconds

Mesure perceptive de la première impression.

## ➤ Échelles utilisabilité/UX

Mesures perceptives standardisées et validées empiriquement.

## ➤ Test utilisateur

Technique d'évaluation holistique combinant performance et perception.

# Oculométrie

Mesure la position et le mouvement des yeux :

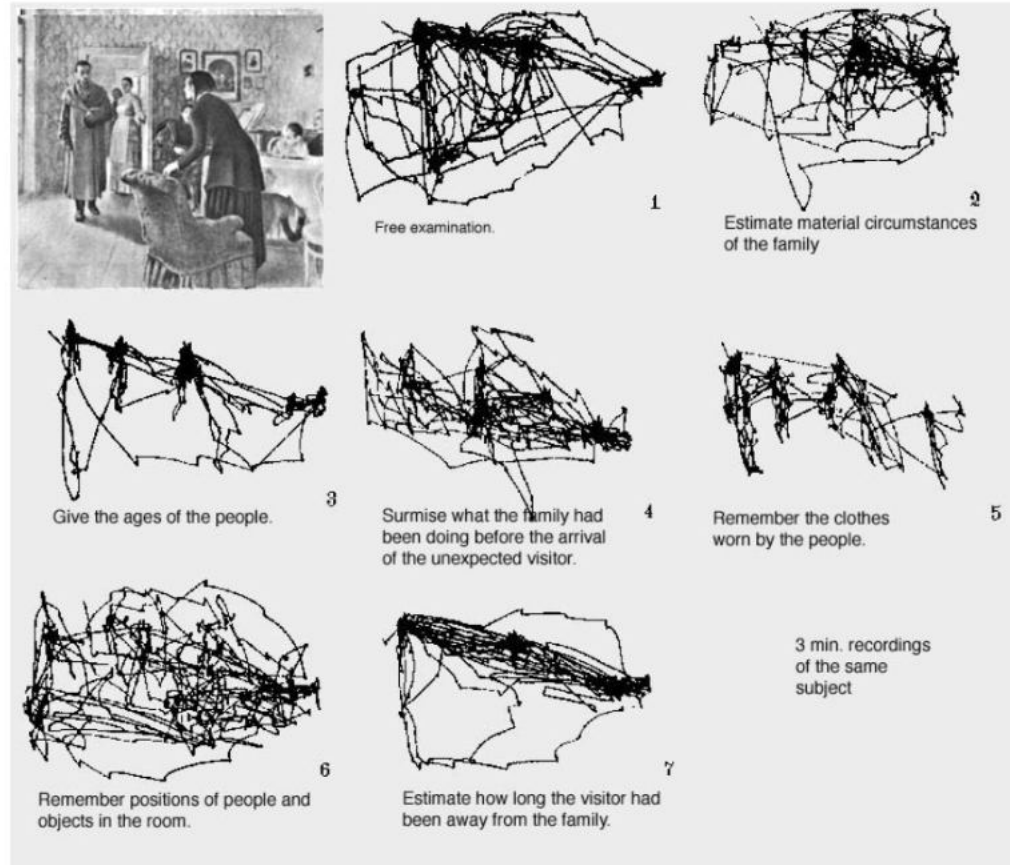
## ➤ Fixations

*Information processing.* Temps pendant lequel la pupille *ne bouge pas* (en réalité elle bouge toujours). En général, indication que la personne traite l'information.

## ➤ Saccades

*Information seeking.* Temps pendant lequel la pupille se déplace grâce aux mouvements oculaires. En général, indication que la personne cherche de l'information à traiter.

# Yarbus (1967)



Lucs-kho at [English Wikipedia](#), Public domain, via Wikimedia Commons

# Manifestation d'intérêt



## Tâche 1 sur 4

Lire | Résoudre | Solution

Enigme #1 : Échiquier est l'intérieur de l'échiquier. Le joueur, il réutilise les les... possible de... pondre. Co... les avec les neuf bouts de pierres qu... l'abbaye?

Votre réponse

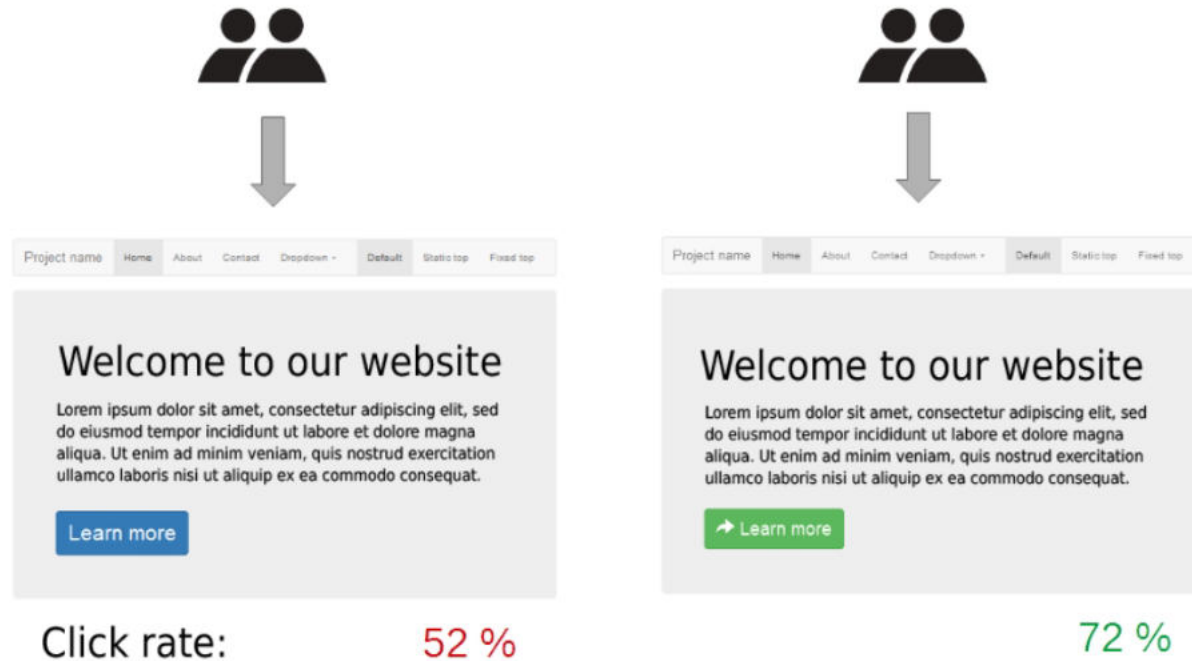
Votre réponse est automatiquement sauvegardée à la fin du temps pour résoudre l'énigme.

Test utilisateur sur un **outil d'awareness émotionnelle** dans un mémoire de Master (Fritz, 2015).



# Test A/B

Répartition aléatoire à une version de l'interface différente.



Maxime Lorant, CC BY-SA 4.0, via Wikimedia Commons



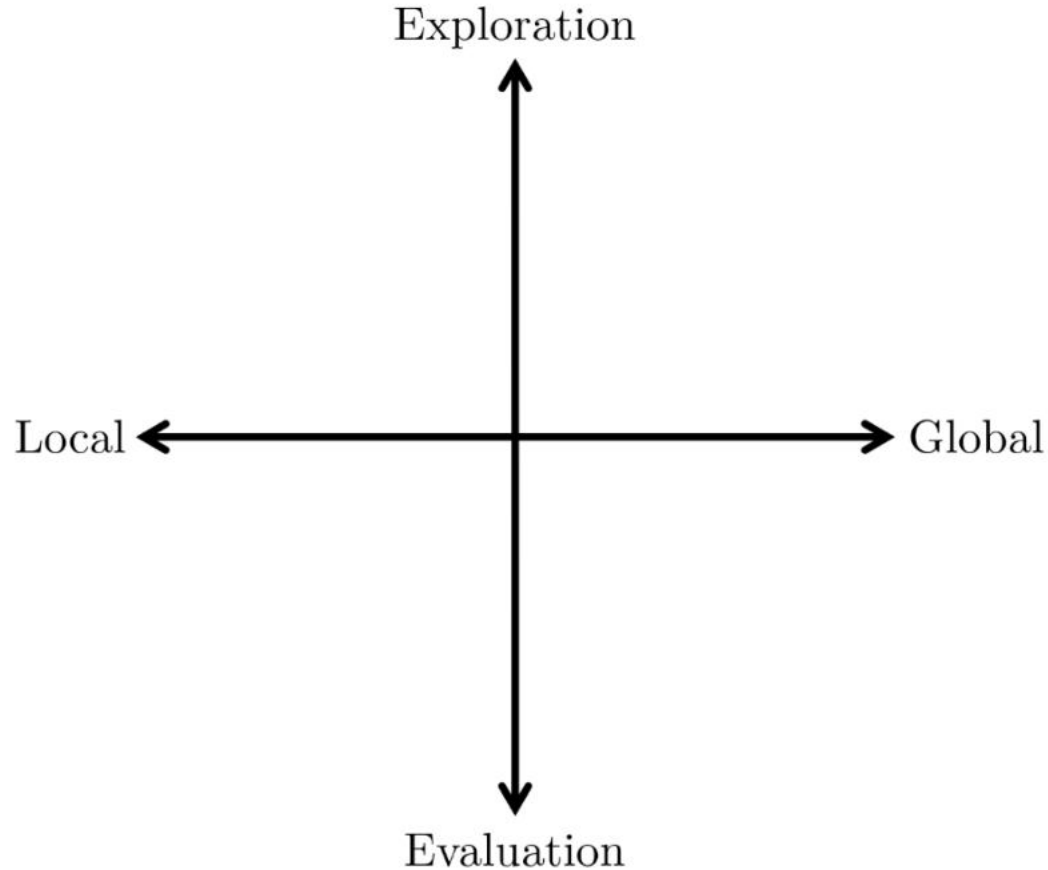
# Parfois avec test statistique

Version	Utilisateurs	Oui	Non	%
A	500	91	409	18.20%
B	500	119	381	23.80%

$$\chi^2(1, N = 1000) = 4.39, p = .036$$

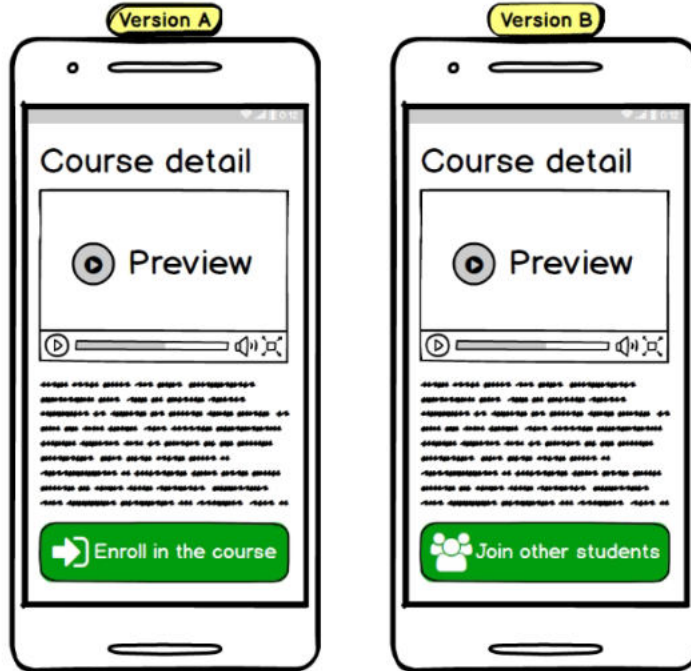
Dans la version B, les utilisateurs ont eu un comportement d'intérêt (e.g. cliquer sur un bouton, visiter une page, s'inscrire à une *mailing list*, ...) avec une fréquence de 5 points de pourcentage plus élevée que pour la version A.

# Différents types de test A/B



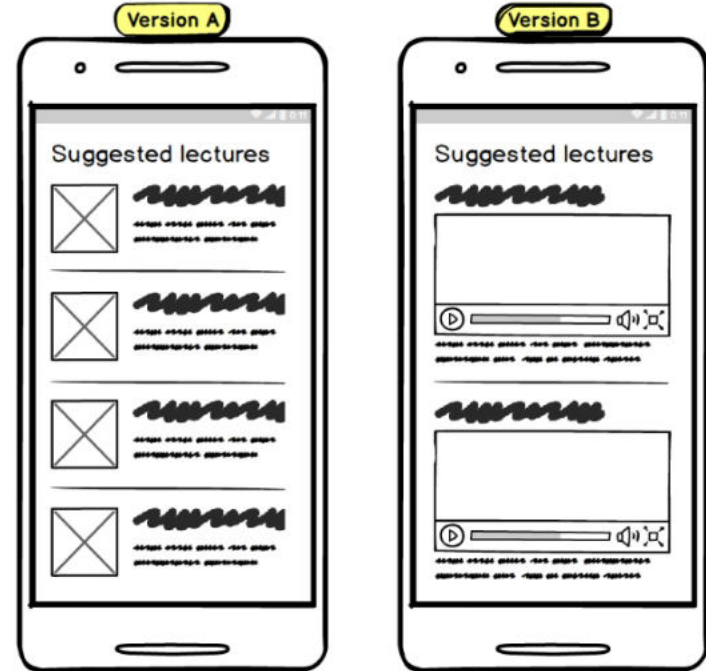
Tiré de *Designing with Data* (King et al., 2017)

# Local vs. Global



Local

Différence dans la labellisation d'un bouton



Global

Différence dans l'affichage des éléments

# Exploration vs. Evaluation

## ➤ Exploration : sans hypothèse précise

Le team UX utilise les données du test A/B pour recueillir plus d'informations. La **cause** d'une éventuelle préférence d'une version plutôt que l'autre n'est pas établie en amont. Les résultats du test sont difficilement généralisables à d'autres choix de design/conception.

## ➤ Evaluation : avec hypothèse précise

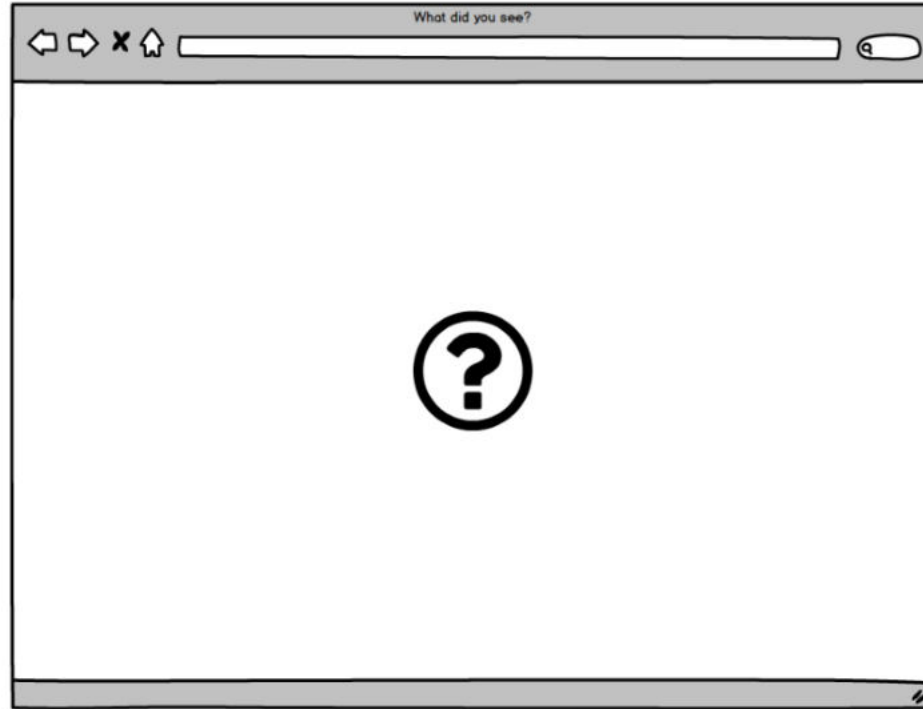
Le team UX a une idée à propos d'un **mécanisme causale** qui pourrait expliquer la préférence d'une version plutôt que l'autre et veut tester la validité de cette hypothèse. Les résultats du test sont plus facilement généralisables à d'autres choix qui impliquent le même mécanisme causale.

# Tests séquentiels ou parallèles



Les tests peuvent être menés en séquence ou en parallèle. À l'image, le deuxième test exploite les résultats du test précédent.

# Test des 5 seconds



Demander de dessiner l'interface et de deviner les objectifs du site après une brève exposition.



# Échelles utilisabilité/UX

Questionnaires standardisés et/ou validés empiriquement.

## ➤ Échelles utilisabilité

En général plus anciennes, limitées plutôt à évaluer l'efficacité, l'efficience et la satisfaction perçues.

## ➤ Échelles UX

En général plus récentes, elles utilisent une approche holistique, en intégrant les dimensions cognitive, sociale et affective.

Elles permettent la comparaison entre versions ou systèmes différents.



# System Usability Scale

J'aimerais utiliser ce système fréquemment	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Je trouve ce système inutilement complexe	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Je pense que ce système est facile à utiliser	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
J'aurais besoin d'un support technique pour pouvoir utiliser ce système	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Les différentes fonctionnalités de ce système sont bien intégrées	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ce système est truffé d'incohérences	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Le grand public peut apprendre à utiliser ce système très rapidement	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ce système est lourd à utiliser	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
J'ai confiance en ce système	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
J'ai dû apprendre beaucoup de choses avant de pouvoir utiliser le système	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

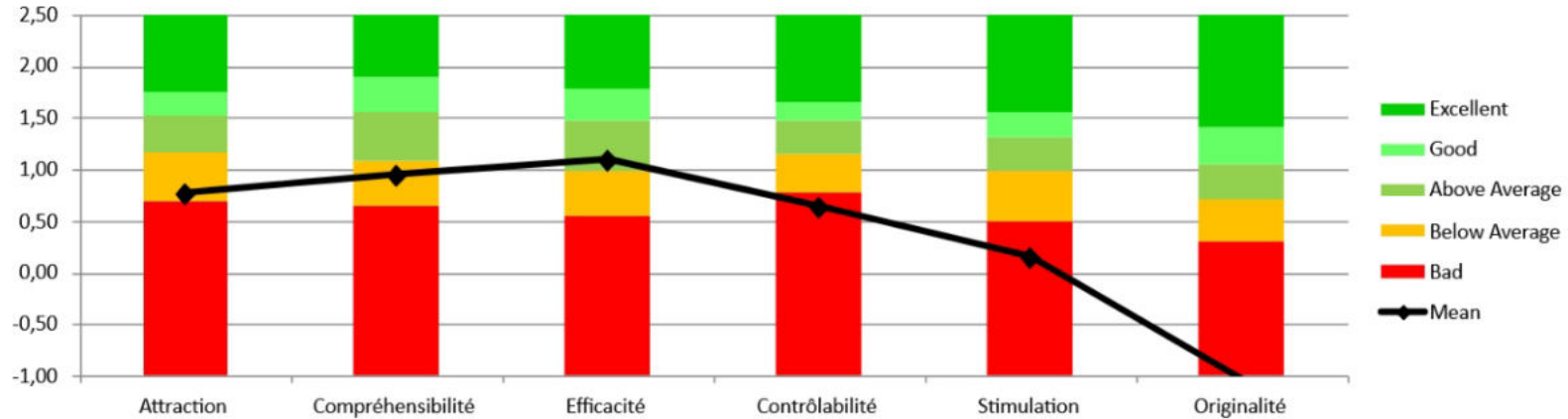
Traduction libre de l'échelle de Brooke (1996).

# User Experience Questionnaire

	1	2	3	4	5	6	7		
Agacant	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Agréable	1
Incompréhensible	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Compréhensible	2
Moderne	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Sans fantaisie	3
Appropriation simple	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Appropriation compliquée	4
Apporte de la valeur	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Peu de valeur ajoutée	5
Ennuyeux	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Captivant	6
Inintéressant	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Intéressant	7
Imprévisible	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Prévisible	8
Rapide	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Lent	9
Original	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Conventionnel	10
Rigide	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Facilitant	11
Bien	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Médiocre	12

Traduction non validée pour le moment de Laugwitz, Held & Schrepp (2008)

# Benchmark UEQ



Résultats de la passation de l'échelle UEQ avec N = 20 participant-es sur une maquette interactive de site web.

Les résultats sont comparés à un benchmark proposé avec l'échelle qui prend en compte l'évaluation de 452 produits pour un total de 20'190 passations (Schrepp, Hinderks, & Thomaschewski, 2017).

# Test utilisateur

Technique holistique : performance et perception.



Photo by [Annie Spratt](#) on [Unsplash](#)

# Étapes test utilisateur

## 1. Accueil

De manière similaire à une recherche scientifique, on met les participant-es dans les meilleures conditions (i.e. on teste le dispositif et non pas leurs capacités)

## 2. Scénarisation/Mise en situation

Le test s'adresse à des *persona* (profils d'utilisateurs issues de l'analyse exploratoire) qu'il faut activer/simuler.

## 3. Exécution de tâches en lien avec le scénario

Les participant-es effectuent une série de tâches authentiques en lien avec le scénario. Les passations sont en général encodées, e.g. avec enregistrement audio/vidéo et/ou oculométrie.

## 4. Évaluation post-test et débriefing

Après les tâches, les participants répondent à des questions fermées (e.g. échelles UX), semi-directifs et/ou ouvertes qui complètent les tâches.



# Tâches scénarisées



Trouvez les prix des tickets d'entrée au musée.



Vous souhaitez visiter le musée le dimanche, avec votre partenaire et vos deux enfants de 15 et 8 ans. Combien va vous coûter l'entrée au total ?

Les tâches conduisent très souvent à des réponses spécifiques, qui permettent de déterminer si la tâche est réussite.

# Synthèse des tâches

Agrégation des passations pour avoir un aperçu général des tâches scénarisées.

Table 1: Synthèse des passations

#	Temps	Réussite	Bilan
1	0m 35sec	2/3	Les utilisateurs en général trouve l'adresse, mais après plusieurs tentatives. Ils cherchent plutôt dans le footer de la page et dans la présentation du musée avant de cliquer sur Contacts.
2	1m 28sec	3/3	Les utilisateurs ont trouvé facilement le tableau du prix et, mis à part un peu de stress pour le calcul, ils ont trouvé le résultat assez vite.



# Analyse des problèmes

Identification des problèmes majeurs, souvent avec utilisation de grilles/héuristiques pour effectuer un diagnostic (e.g. Bastien & Scapin, 1993).

Table 2: Liste des problèmes majeurs rencontrés

#	Description	Diagnostic	Taux	Gravité du problème
1	Les utilisateurs ont du mal à trouver l'adresse du musée. Ils ne cliquent pas tout de suite sur Contacts et cherchent plutôt dans le footer de la page (Tâche 1)	Guidage > Incitation, et Compatibilité (habitude)	2 sur 3	Critique (1x) et Moyen (1x)

# Proposition de remédiation

“ Suite aux résultats du test utilisateur, lors de la prochaine itération, nous conseillons d'ajouter l'adresse complète, les horaires d'ouverture et les coordonnées de contacts du musée en bas de chaque page du site. ”

— Votre team UX

Avis **justifiés** et accompagnés souvent par éléments multimédia (diagrammes, séquences vidéos, *gaze path*)

# Évaluer les évaluateurs

Critères de sélections pour des utilisateurs :

## 1. Disponibilité de temps

Si le feedback/tâche à faire est trop simple et rapide, les gens ont tendance à donner des avis trop positifs ("Oui, super... allez-y!")

## 2. Disponibilité à mettre en jeu sa propre réputation

Actions concrètes qui témoignent *attachement*, par exemple retweet d'un message de l'entreprise qui lance le produit.

## 3. Disponibilité en termes d'argent

De manière paradoxale, on peut demander aux utilisateurs de payer pour participer à un test, ou les engager à acheter le produit à la fin de la conception.

# Importance de la perspective

Une évaluation n'existe pas en absolu :

- Quels sont les objectifs ?
- Quel est le contexte d'utilisation ?
- Qui est le public cible ?
- Quelles sont les conséquences envisageables ?
- Qu'est-ce qu'on ne pourra pas savoir à cette étape ?



# L'importance de la perspective



Leonardo da Vinci, CC BY-SA 4.0, via Wikimedia Commons

Rotation de l'image

# Merci pour votre attention !

Mattia A. Fritz

TECFA, Université de Genève



Présentation créée avec **Reveal.js**.