

Votre **méthode** de conduite de la recherche


N.Mandran (LIG Grenoble)
Janvier-Février 2020



Pour citer ce document : N.Mandran, Janvier-Février 2020, « Votre méthode de conduite de la recherche en EIAH ». Formation Ecole Doctorale de Lille



Quelques mots de présentation ...

- Nadine Mandran, Méthodologue 
 - Statistiques, SHS, Démarche qualité
 - Ingénieur encadrant des doctorants sur les méthodes de production et analyse de données
 - Spécialisée dans les méthodes de conduite de la recherche en informatique centrée humain - THEDRE
- Mathieu Vermeulen
 - Responsable cellule TICE
 - Spécialisé dans la conception de jeux sérieux centrée utilisateurs – metadesign –
 - Créateur de la méthode ALPES, pédagogie par projets
- Estelle Prior
 - Chargée d'étude projet APACHES
 - Amélioration et évaluation de THEDRE et de sa méthode d'enseignement mais aussi de ALPES



Une thèse ?? Thesis ??

Une thèse //Thesis

- à identifier un sujet, // identify a problematic
- à rassembler des documents (données) sur ce sujet //gathered data and documents
- à les organiser, // analyze and manage them
- à construire une proposition, à l'évaluer // build a contribution and test it
- se faire comprendre, défendre ses idées //make yourself understood,defend your ideas

Aussi // Also

- on oubliera ainsi de parler comme un poète // not a poem
- pas non plus obligé de tout lire sur le sujet mais le cœur de cible // not exhaustive literature but 'target' literature

Une thèse est une expérience de travail, mais c'est aussi un pari, un jeu, une chasse au trésor //A thesis is a work experience, but it is also a bet, a game, a treasure hunt.

Le vrai défi dans une thèse, c'est de la vivre comme un défi, de se mettre soi-même au défi.

The real challenge in a thesis is to live it as a challenge, to challenge oneself.





Seulement 3 ans pour tout ça

Une thèse //Thesis

- à identifier un sujet, // identify a problematic
- à rassembler des documents (données) sur ce sujet //gathered data and documents
- à les organiser, // analyze and manage them
- à construire une proposition, à l'évaluer // build a contribution and test it
- se faire comprendre, défendre ses idées //make yourself understood,defend your ideas

LIRE, CONCEVOIR, IMAGINER, DEVELOPPER, TESTER, REDIGER, REDIGER, REDIGER, RELIRE, RECONCEVOIR, AMELIORER, RETESTER, REDIGER, REDIGER, REDIGER, ...

READ, DESIGN, IMAGINE, DEVELOP, TEST, WRITE,WRITE,WRITE, REREAD, REDESIGN, ENHANCE, RETEST, WRITE,WRITE,WRITE





Seulement 3 ans // Only 3 years

Une thèse // Thesis

- à identifier un sujet, // identify a problematic
- à rassembler des documents (données) sur ce sujet // gathered data and documents
- à les organiser, // analyze and manage them
- à construire une proposition, à l'évaluer // build a contribution and test it
- se faire comprendre, défendre ses idées // make yourself understood, defend your ideas

READ, DESIGN, IMAGINE, DEVELOP, TEST, WRITE, WRITE, WRITE, REREAD, REDESIGN, ENHANCE, RETEST, WRITE, WRITE, WRITE

The real challenge in a thesis is to live it as a challenge, to challenge oneself.



PILOTER VOTRE THESE // DRIVE YOUR THESIS

AVOIR DES GUIDES ET DES OUTILS // GUIDES AND TOOLS





Objectifs de ces deux jours

- Concevoir votre méthode de conduite de la recherche dans un cadre de qualité
- Former à l'expérimentation avec des humains en contexte de laboratoire ou in situ
- Former aux méthodes de production de données des SHS





Avant propos ...

Méthodologie ou méthode ?

- Méthodologie : étude de la méthode, conception de méthodes

« **Methodology** is the systematic, theoretical analysis of the methods applied to a field of study. It comprises the theoretical analysis of the body of methods and principles associated with a branch of knowledge. Typically, it encompasses concepts such as paradigm, theoretical model, phases and quantitative or qualitative techniques » (Berg, 2009).

- Méthode : ensemble de tâches, d'outils, de procédures assemblés et organisés dans le temps pour atteindre un but.

Method is “an integrated collection of procedures, techniques, product descriptions, and tools, for effective, efficient, and consistent support of the engineering process” (Harmsen et al., 1997).

- **Nous utilisons uniquement le terme méthode que nous qualifions par d'autres termes : méthode de conduite de la recherche, méthode de production de données, méthode d'analyse de données ...**



Brainstorming

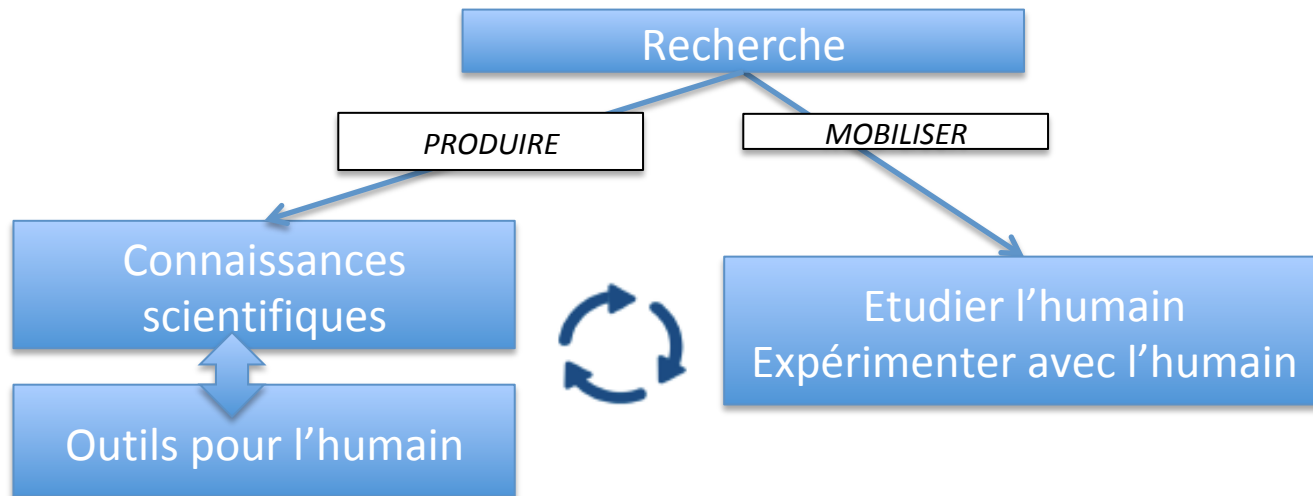
Si votre thèse était un animal ?

Si votre thèse était une couleur ?



L'ancrage de notre méthode de conduite de la Recherche en Informatique Centrée Humain

- **Un cadrage scientifique global, les sciences de l'artificiel (H.Simon 2004)**
 - **Un outil** que l'humain va pouvoir mobiliser : un EIAH, un dispositif d'enseignement, un Dashboard, ...
 - Des **connaissances scientifiques** basées sur l'activité humaine : modèle de l'apprenant, modèle de l'enseignant, modèle pédagogique, modèle de traces, ...



- Herbert A. Simon. 2004. *Les Sciences de l'artificiel* (traduction J.L. Lemoigne Éd. rev. et complétée ed.). Folio, Paris •



Exemple : simulateur TELEOS

ANR TELEOS, V.LUENGO ET AL.

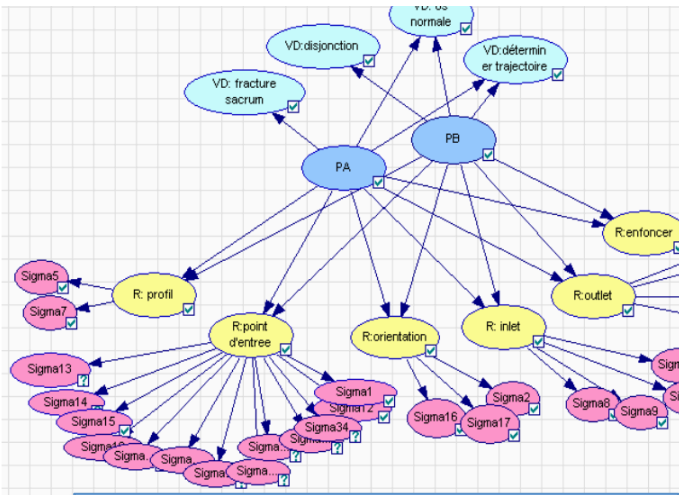


Simulateur pour l'apprentissage de la chirurgie percutanée :

L'outil activable produit par la recherche



Exemple : simulateur Teleos



Modèle de l'apprenant



Le simulateur



Exemple : simulateur TELEOS

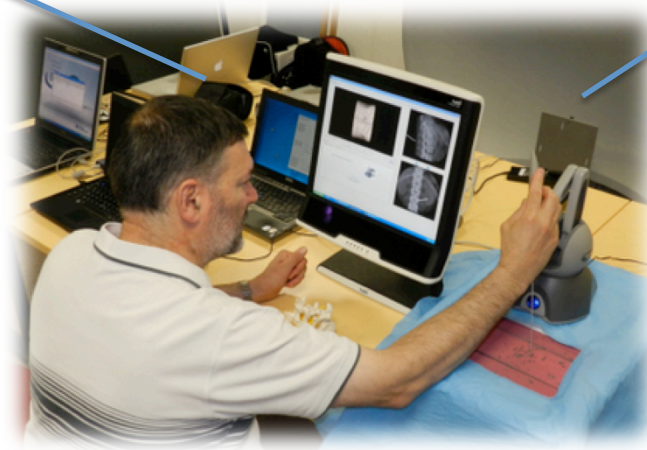
L'outil activable est décomposable en sous-parties

L'interface :

- Terminologie
- Fonctionnalités
- ...

Le bras haptique :

- Manipulation
- Résistance
- ...





Exemple : simulateur Teleos

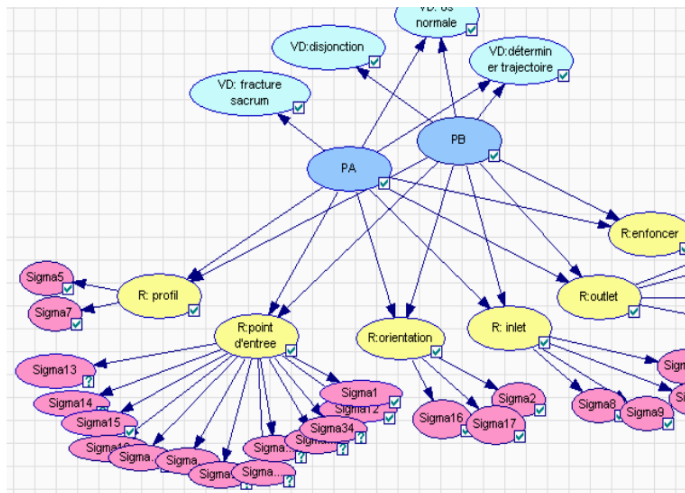
Composants Activables (CA)

L'interface :

- Terminologie
- Fonctionnalités
- ...

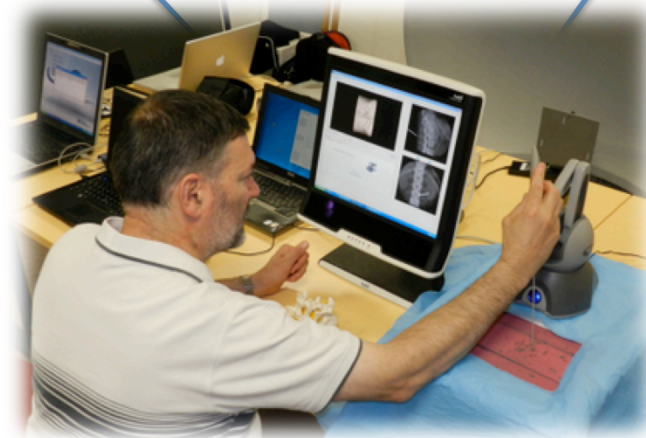
Le bras haptique :

- Manipulation
- Résistance
- ...



Modèle de l'apprenant

Connaissances scientifiques (KS)



Le simulateur

Outils Activables (OA)



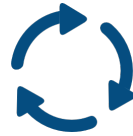


Exemple : simulateur Teleos



T1 : In situ

OA n'existe pas



T2 : Maquettage

OA existe statique



T3 : Simulateur

OA existe dynamique

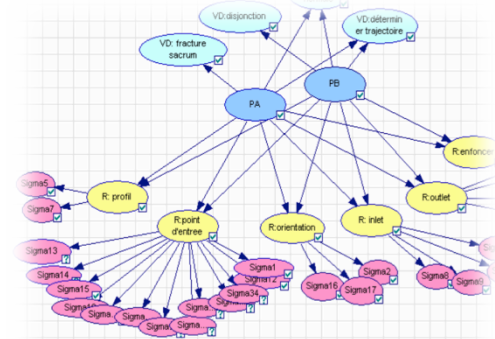


Figure 2 - le réseau des conceptions



Exemple : simulateur Teleos



T1 : In situ

OA n'existe pas



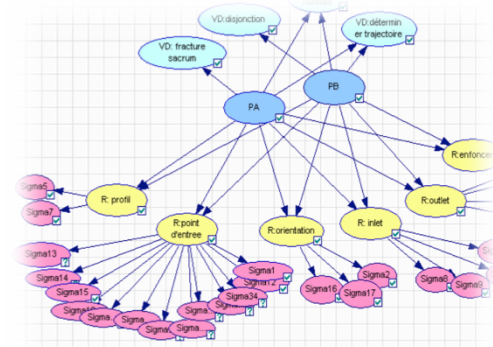
T2 : Maquettage

OA existe statique



T3 : Simulateur

OA existe dynamique





L'ancrage de notre proposition de conduite de la recherche

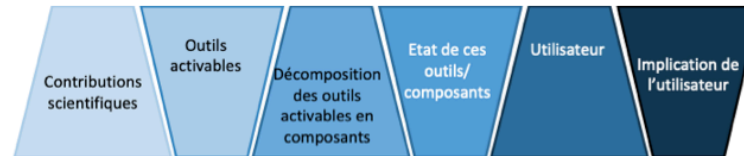
- **Trois états de l'outil activable : utilisable par un humain**
 - **Rien n'existe**
 - L'observation du terrain va permettre de commencer à élaborer un outil
 - **Statique**
 - Un outil existe mais il est sous un format non numérique. Exemples : un dictionnaire, des maquettes papier, ...
 - **Dynamique**
 - Quand des outils sont disponibles sous forme numériques
 - Exemples : une application, un site web, un robot, ...
- **Des outils décomposables : découper l'outil en composants de manière à avoir des objets d'études précis**
 - Exemple un site web : une terminologie, des fonctionnalités, une palette graphique, ...



Pour vous : Connaissances et outils



« Spécifier la contribution scientifique et les outils associés »



Pourquoi : Travail préparatoire pour identifier ce qui sera produit pendant le travail de thèse et identifier les outils à tester.
Quand : Pendant la définition de la problématique et revenir dessus après chaque expérimentation.

Contribution(s) scientifique(s)	Outil(s) activable(s) qui embarque(nt) la/les contribution(s) scientifique(s)	Décomposition des outils activables en composants	Etat de ces outils Existent-ils ? Si oui, sont ils en format numérique ? Ou en format statique ?	Quels utilisateurs ? Pourquoi ?
Exemple : Modèle de l'apprenant – V1	Plateforme de type LMS – V1 -	Terminologie employée Fonctionnalités Les ressources fournies Les indicateurs de learning analytics	Existe sous format papier Existe sous format papier Existe en format papier Existe en format papier	Enseignants, évaluer et améliorer les composants
Exemple : Modèle de l'apprenant – V2	Plateforme de type LMS – V2 -	Fonctionnalités Les ressources fournies Les indicateurs de learning analytics	Existe en numérique	Enseignants, évaluer avec eux





Debriefing

- Types de KS
- Types de OAs et de Cas
- Etat des OAs
- Et l'humain
- Incrémentale et Itérative ?



Notation

De 0 à 5 sur le, les outil(s) présenté(s)

Discussion ...



- Qu'est ce qui est répétable ou reproductible ?
 - Protocole de production de données ?
 - Les données ?
 - Les résultats techniques ?
 - Le résultats scientifiques ?



Répétabilité vs Reproductibilité

- ISO 3534-1

- L'idée de répétabilité est simple



- Il n'y a aucun problème, tout est ok, cependant, j'ai beau exécuter "tout pareil" sur un même échantillon, j'ai un aléa sur les résultats des mesures répétées. Pourtant c'est la même balance avec le même étalonnage, la même pipette à quelques secondes d'intervalle, la même température ...

- L'idée de reproductibilité est plus complexe



- On va s'intéresser aux aléas quand les conditions de mesures varient. Le plus souvent, ce qui est intéressant, c'est l'étude de cet aléa quand ce qui varie est le laboratoire où sont réalisées les mesures.



Répétabilité vs Reproductibilité

- ISO 3534-1

- **Conditions de répétabilité**

- Conditions où les résultats d'essais indépendants sont obtenus par la même méthode sur des individus d'essai identiques dans le même laboratoire, par le même opérateur, utilisant le même équipement et pendant un court intervalle de temps.

- **Conditions de reproductibilité**

- Conditions où les résultats d'essai sont obtenus par la même méthode sur des individus d'essais identiques dans différents laboratoires, avec différents opérateurs et utilisant des équipements différents



Répétabilité vs Reproductibilité

- Quasiment impossible en RICM ou toutes autres sciences impliquant l'humain...
- Les protocoles de production des données sont reproductibles (?) :
 - Ils peuvent être réutilisés avec d'autres utilisateurs dans des conditions différentes ou similaires
 - Nécessite de tracer le processus de production et de traitement des données
 - Nécessité de tracer les évolutions de la production scientifique (connaissance scientifique et des outils)



Cadre épistémologique

- Méthode de conduite de la recherche demande un positionnement épistémologique [Avenier & Thomas 2015]
- «L'épistémologie : étude de la constitution des connaissances valables» [Le Moigne 1995]
- Pour [Le Moigne 1995], cette définition pose des questions :
 - Qu'est-ce que la connaissance ?
 - Comment est-elle constituée ?
 - Comment apprécier sa valeur ou sa validité ?



Cadre épistémologique

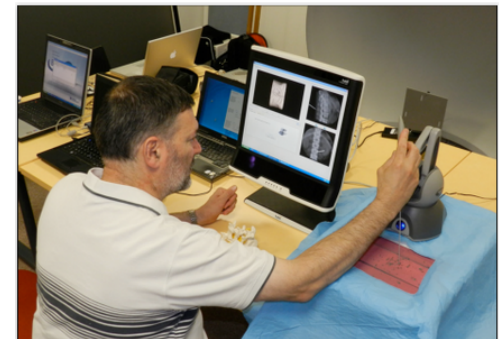
- **Constructivisme pragmatique** : [Avenier & Thomas 2015]
 - Hypothèse 1 : Des **représentations du réel** existent
 - Hypothèse 2 : **L'humain va exprimer ses connaissances du monde**
 - Hypothèse 3 : L'outil activable a **une finalité dans un contexte** donné
 - Hypothèse 4 : La construction de l'instrument est **incrémentale**
 - Hypothèse 5 : La **question de recherche peut évoluer**, en accord avec les contextes académique, technique et sociétal et les résultats du terrain





Cadre épistémologique

- Valeur
 - Contribution : Incrémenter des connaissances existantes
 - Qualité du construit : Adéquation de l'outil activable aux activités humaines
- Validité
 - Mise à l'épreuve : Phase expérimentale où l'outil est construit et évalué
 - Multiplicité des données : Une large diversité des méthodes de production (qualitatif et quantitatif)
 - Fiabilité des données : Volume et qualité des données suivis par des indicateurs





Cadre épistémologique

- **Un cadrage épistémologique**
- **Constructivisme pragmatique** [Avenier & Thomas 2015]
 - Contribution : Incrémenter des connaissances existantes
 - Qualité du construit : Adéquation de l'outil activable aux activités humaines
 - Mise à l'épreuve : Phase expérimentale où l'outil est construit et évalué
 - Multiplicité des données : Une large diversité des méthodes de production (qualitatif et quantitatif) et de raisonnement
 - Fiabilité des données : Volume et qualité des données suivis par des indicateurs
 - Documenter et tracer le travail de recherche : protocole de production et de traitement des données, l'évolution de la connaissance et es outils activables



Traceable Human Experiment Design Research

THEDRE (Mandran 2018)

Pour

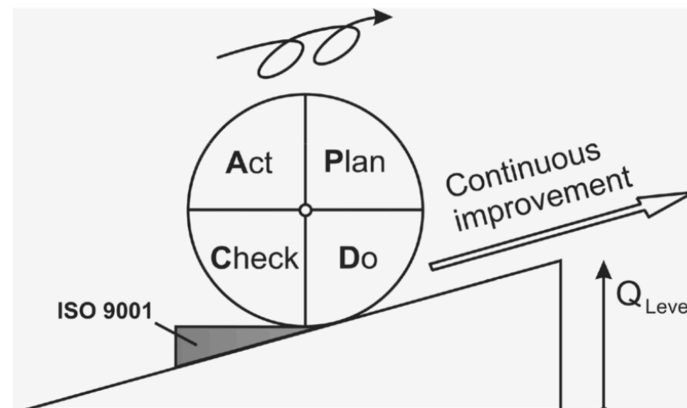
les recherches qui proposent des modèles de l'activité humaine en intégrant l'humain

des recherches qui développent des outils activables supports à l'activité humaine

+ de THEDRE Processus de guidage et outils

Avec un modèle de processus de conduite de la recherche ancré dans la démarche qualité et des indicateurs de traçabilité (Plan, Do, Check, Act – cycle de Deming)

Des guides pour accompagner l'utilisateur de THEDRE



4 actions et un processus d'amélioration continue [Sokovic et al. 2010]





Traceable Human Experiment Design Research

- **Démarche centrée utilisateurs (Norme Iso, Mandran et al 2013)**
 - 3 tâches :
 - **Observer, analyser** l'activité
 - Quand « aucune » connaissance ou outil n'est disponible
 - Se référer à des études existantes pour créer un outil d'observation
 - Aller sur le terrain observer, recueillir toutes informations ou documents utiles
 - **Co-construire** avec l'humain,
 - Quand des connaissances ou des outils sont disponibles et que l'on veut les co-construire avec les utilisateurs
 - Nécessite de continuer leur développement
 - Procéder par réunion de travail collaboratif, « interactivité »
 - **Evaluer** le dispositif avec l'humain
 - Quand des outils ou connaissances sont suffisamment aboutis pour être testés
 - Procéder par tests utilisateurs ou analyse de traces



Théorie et Méthode

Théorie

Cadre scientifique général
Sciences de l'artificiel , Naturelles, SHS, ...

Posture épistémologique : Quoi ? Sur
quelle réalité ? Quelle validité

Cadre scientifique spécifique

Méthode

Processus et Guides
THEDRE



Démarche centrée
Utilisateur

Démarche qualité/
traçabilité/indicateurs



Importance de cet ancrage

- De **distinguer recherche et terrain** tout en tenant compte de leur interdépendance
- **Intégrer tous les acteurs** de la mise en œuvre d'un dispositif ou d'une recherche
- **Contributions et outils activables : Identifier ce qui est mis en place**
 - **modèles théoriques** : pédagogiques, d'apprenants, d'enseignants, ..
 - **intentions** des outils pédagogiques
 - **intentions** des situations pédagogiques
 - **aspects techniques** des outils pédagogiques
 - **aspects techniques** des outils numériques
- De **découper les outils à construire et à évaluer** en composants pour préciser les objectifs
- De **tracer le travail** réaliser et ainsi répondre aux différents acteurs impliqués dans la thèse : enseignants, apprenants, concepteur technique, chercheurs, ... par des indicateurs/mesures ...



Entrer dans ce « monde » ?



Théorie

Cadre scientifique général
Sciences de l'artificiel, Naturelles, SHS, ...

Posture épistémologique : Quoi ? Sur
quelle réalité ? Quelle validité

Cadre scientifique spécifique

Méthode

Processus et Guides
THEDRE 

Démarche centrée
Utilisateur

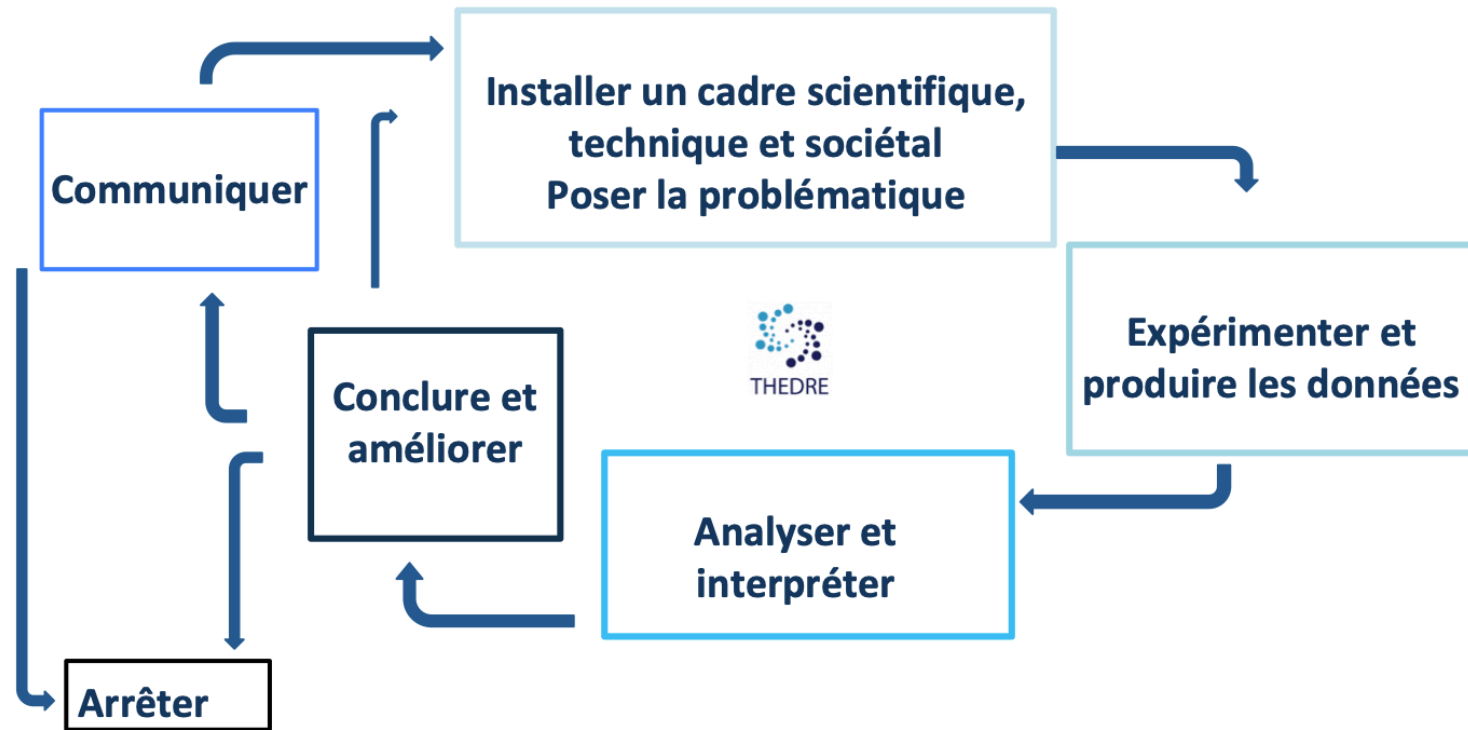
Démarche qualité/
traçabilité/indicateurs

Un modèle adaptable
Des Guides
Logigrammes





Une méthode adaptable





Guides et outils



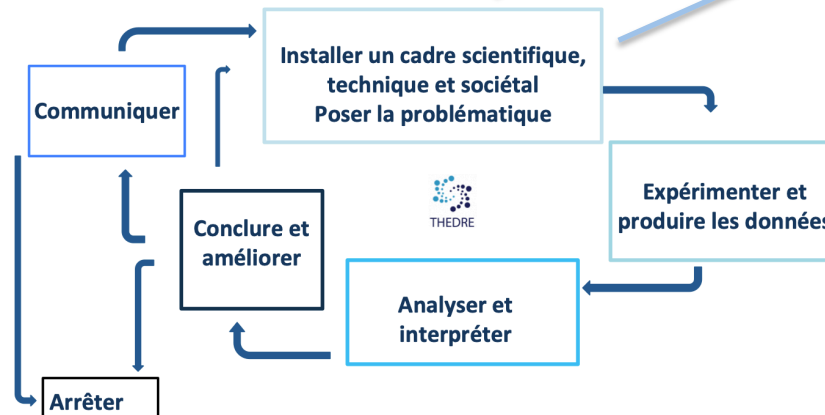
Pour bien commencer un travail de thèse et construire sa problématique



Spécifier la contribution scientifique et les outils associés



Spécifier les indicateurs et les mesures à prendre





Guides et outils



Pour bien commencer un travail de thèse et construire sa problématique



Spécifier la contribution scientifique et les outils associés



Spécifier les indicateurs et les mesures à prendre



Lister les critères pour choisir une/des méthodes de production de données



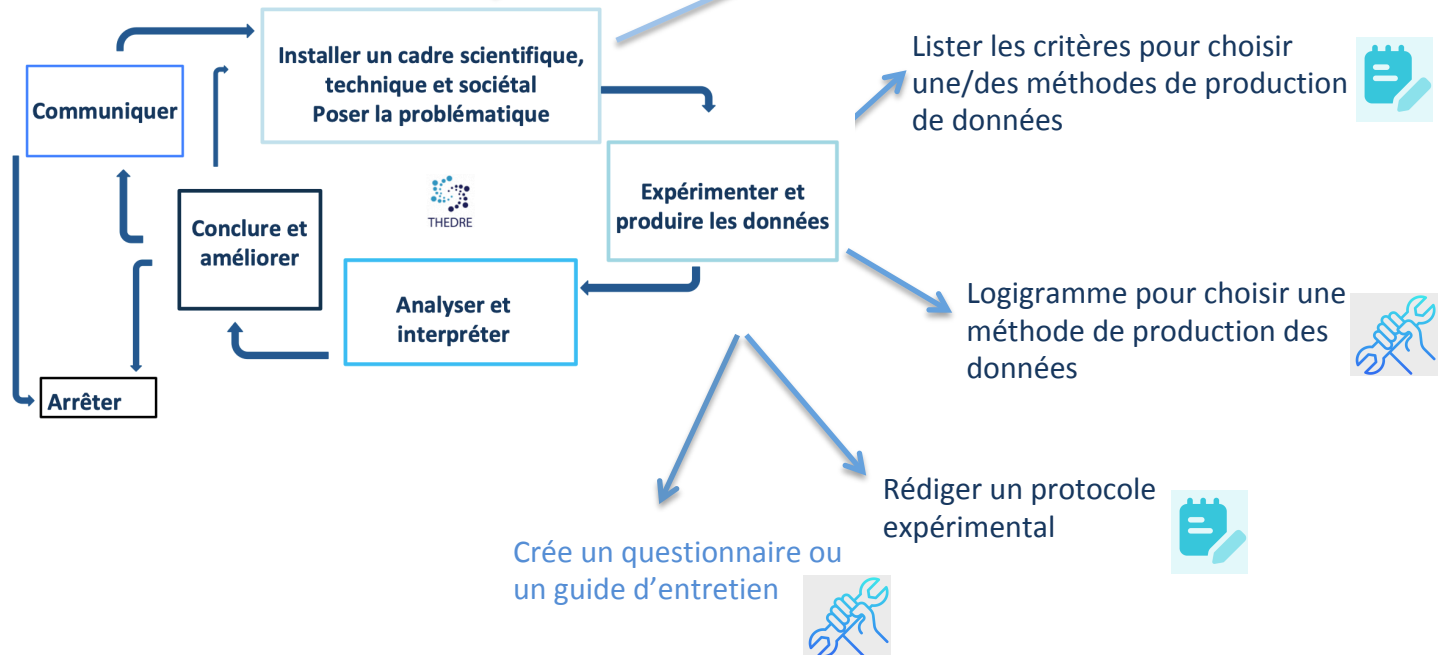
Logigramme pour choisir une méthode de production des données



Rédiger un protocole expérimental



Crée un questionnaire ou un guide d'entretien





Guides et outils



Pour bien commencer un travail de thèse et construire sa problématique



Spécifier la contribution scientifique et les outils associés



Spécifier les indicateurs et les mesures à prendre



Lister les critères pour choisir une/des méthodes de production de données



Logigramme pour choisir une méthode de production des données



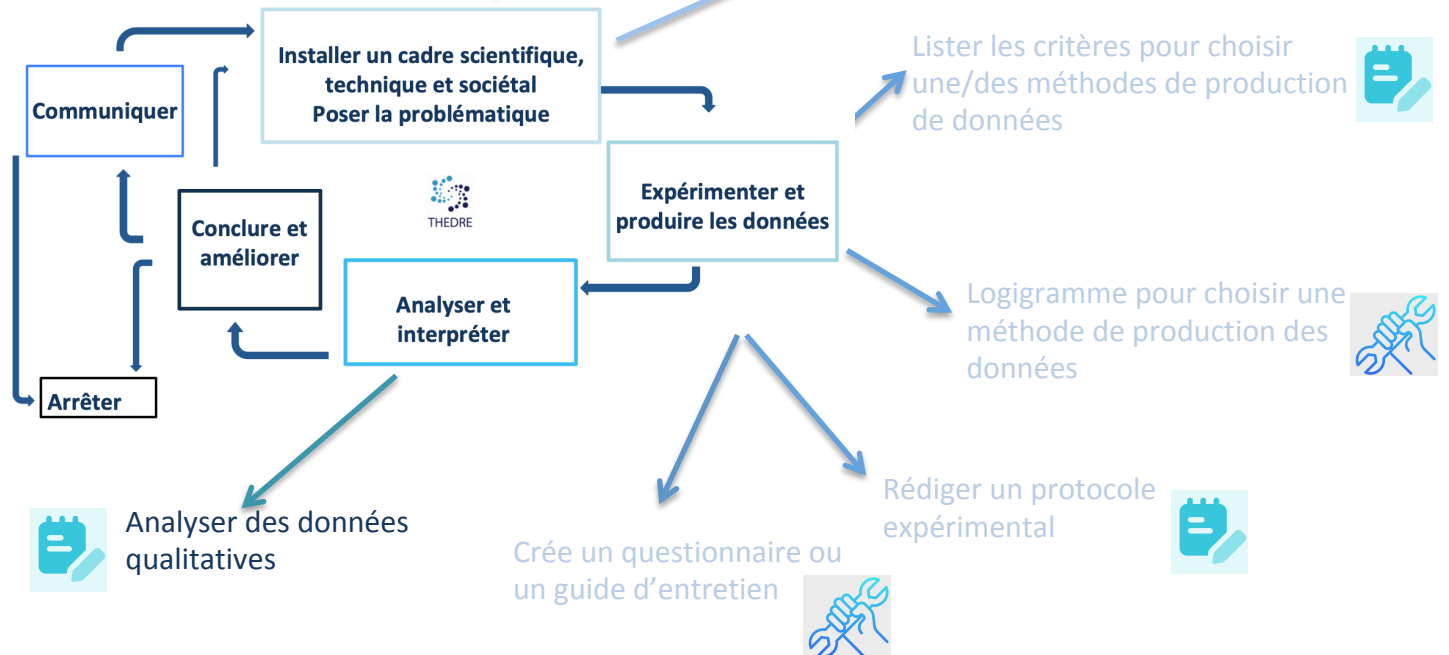
Rédiger un protocole expérimental



Crée un questionnaire ou un guide d'entretien



Analyser des données qualitatives





Guides et outils



Pour bien commencer un travail de thèse et construire sa problématique



Spécifier la contribution scientifique et les outils associés



Spécifier les indicateurs et les mesures à prendre



Lister les critères pour choisir une/des méthodes de production de données



Logigramme pour choisir une méthode de production des données



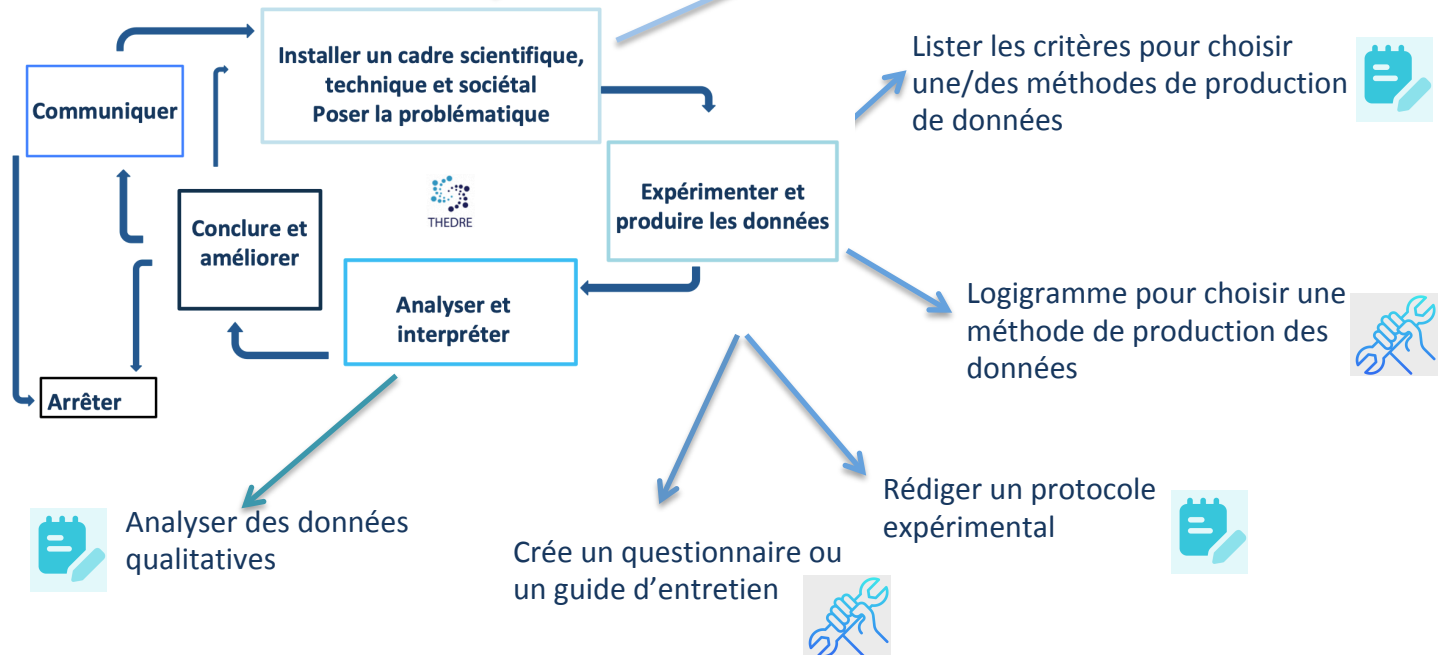
Rédiger un protocole expérimental



Crée un questionnaire ou un guide d'entretien

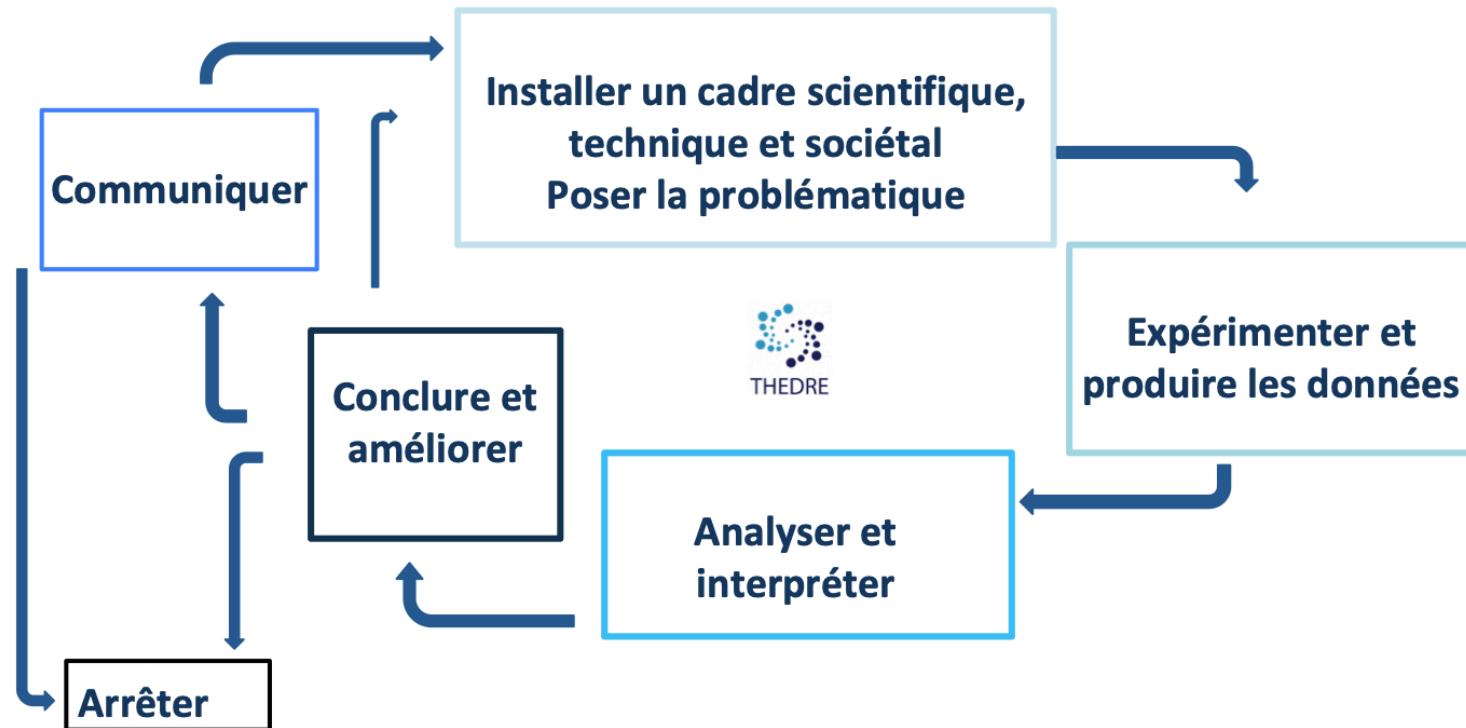


Analyser des données qualitatives





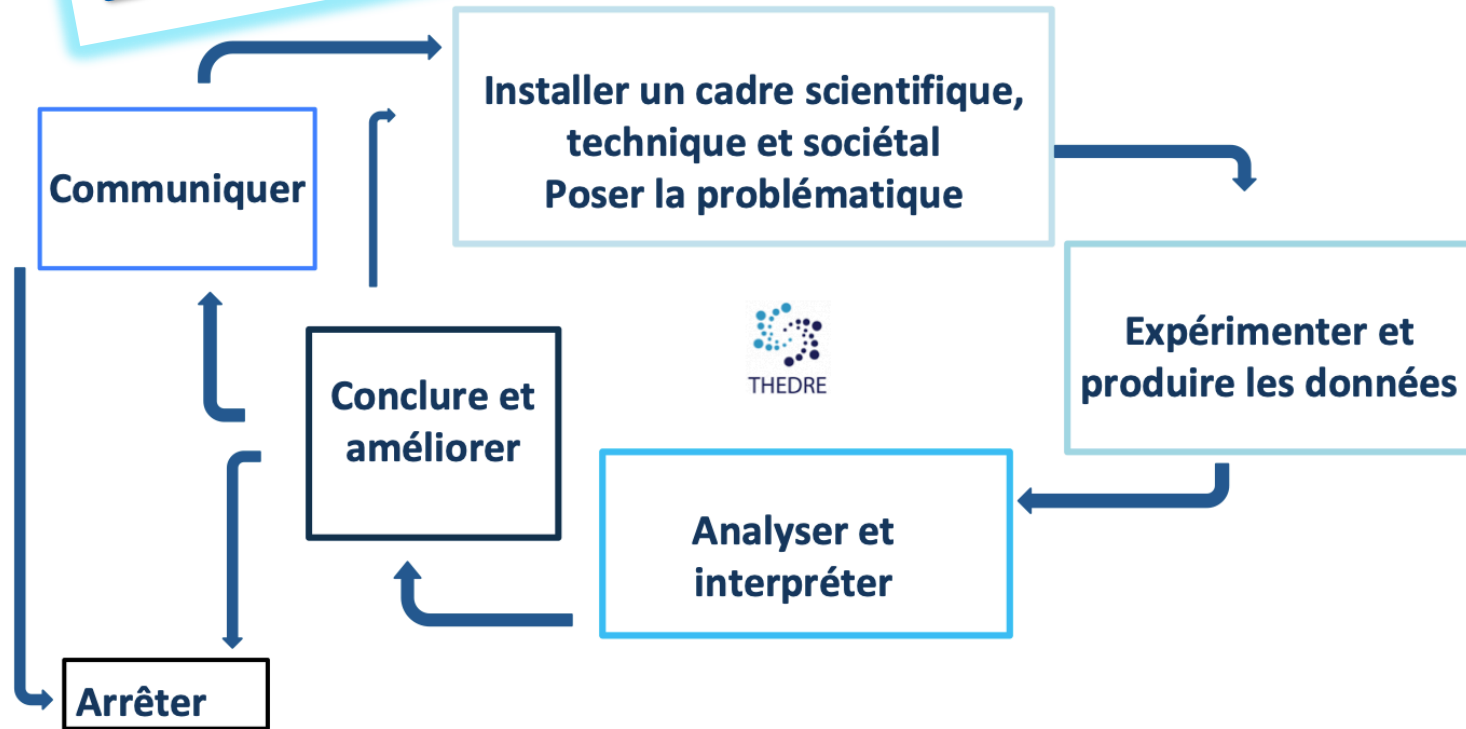
Une méthode adaptable





Une méthode adaptable

Debriefing





Notation

De 0 à 5 sur le, les outil(s) présenté(s)



Importance des tâches

- Avec des gommettes rouge et orange venir indiquer les tâches qui vous semblent «bloquantes » et « à risque »





Importance des tâches

Debriefing

- Avec les gommettes rouge et orange venir indiquer les tâches qui vous semblent « bloquantes » et « à risque »





- Pause de midi ...



Indicateurs et Mesures

- Un indicateur répond à une question centrale, ou ne prend un véritable sens que lorsqu'il mesure un objectif concret.
- Un indicateur aide à prendre une décision, à suivre un projet.
- Un indicateur est construit à partir d'une ou plusieurs mesures
- Un indicateur peut prendre plusieurs formes : visualisation, tableaux, valeurs, intervalles, seuils, ...
- Les mesures sont produites sur le terrain avec des méthodes de production de données qualitatives ou quantitatives.
- La mesure est associée à un outil de mesures, une unité

- Exemples :
 - L'indicateur est la moyenne d'une classe pour un enseignant, les mesures sont les notes des élèves. Domaine de validité : intervalle de 1 à 20
 - La note est un indicateur pour l'élève, c'est aussi la mesure. Unité : intervalle de 1 à 20
 - Indicateur de satisfaction d'un outil, la mesure : note de de satisfaction des utilisateurs, unité échelle de Lickert
 - Indicateur de performance : mesure de rapidité d'exécution, unité le temps



Indicateurs et Mesures

- Indicateurs d'activités : IA
 - Ce qui est fait, ce sont des actions
 - Fait/pas Fait
 - Exemple : Créer un outil, rédiger le protocole expérimental, Lire la bibliographie,...
- Indicateurs de production : IP
 - Ce qui est produit
 - Produit/non produit
 - Exemple : les données expérimentales, une synthèse de la bibliographie, ...
- Indicateurs d'objectifs : IO
 - Fixés a priori avant le démarrage d'un projet pour savoir si le projet a atteint son but
 - Exemple : Avoir suivi un dispositif sur 2 ans, nombre d'utilisateurs >1000 étudiants



Indicateurs et Mesures

- Indicateurs
 - ⇒ Suivre l'activité et le processus de recherche
- Mesures
 - ⇒ Prise de mesures lors de l'expérimentation

Propriétés des indicateurs et des mesures : **SMART**
Spécifique, Mesurable, Acceptable, Réaliste, Temporel



Indicateurs et Mesures

- Indicateurs
 - ⇒ Suivre l'activité et le processus de recherche
- Mesures
 - ⇒ Liées à l'expérimentation

Propriétés des indicateurs et des mesures : **SMART**
Spécifique, Mesurable, Acceptable, Réaliste, Temporel



Indicateurs et Mesures

- Utilisabilité d'un dispositif mobile
 - Répondre à une question expérimentale liée à la problématique de recherche => mesure
 - Répondre à un objectif de la recherche => indicateur d'objectif
 - Note d'utilisabilité > 7





Indicateurs et Mesures



Indicateurs d'Activités



Indicateurs de Production



Indicateurs d'Objectifs



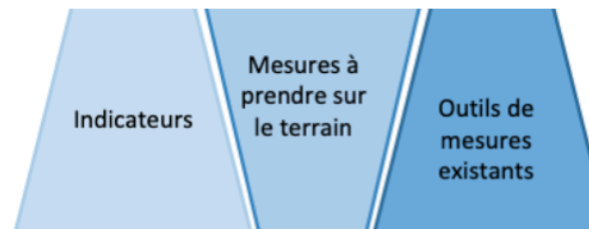
Mesures expérimentales

Plus document pour décrire les indicateurs





Guide « Identifier les indicateurs et les mesures »



Pourquoi : Travail sur la création des indicateurs d'objectifs, permet d'approfondir les mesures qui seront prises pendant la production des données, permet d'affiner le protocole expérimental, de suivre l'activité et le processus de recherche

Quand : En début de thèse, puis revenir dessus tous les 6 mois ou après chaque expérimentation

Indicateurs d'activités, i.e. pour rendre compte des actions faites au cours du travail de thèse.

- Quelles sont les actions que vous devez réaliser pendant votre thèse ? (Exemple : lecture de la biblio, concevoir les outils, ...)
- Quels sont les indicateurs d'activités (nombre , ratio, ..) que vous allez suivre ? (exemple : nombre de références pertinentes lues, nombre d'utilisateurs dans les expériences, ...)

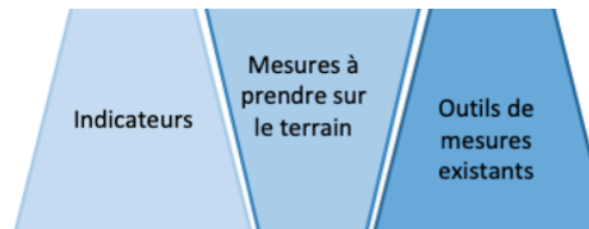
Indicateurs de production, i.e. pour rendre compte des productions faites au cours du travail de thèse.

- Qu'est ce que vous allez produire pendant votre thèse (exemple : livrables, protocoles



Debriefing

Guide « Identifier les indicateurs et les mesures »



Pourquoi : Travail sur la création des indicateurs d'objectifs, permet d'approfondir les mesures qui seront prises pendant la production des données, permet d'affiner le protocole expérimental, de suivre l'activité et le processus de recherche

Quand : En début de thèse, puis revenir dessus tous les 6 mois ou après chaque expérimentation

Indicateurs d'activités, i.e. pour rendre compte des actions faites au cours du travail de thèse.

- Quelles sont les actions que vous devez réaliser pendant votre thèse ? (Exemple : lecture de la biblio, concevoir les outils, ...)
- Quels sont les indicateurs d'activités (nombre , ratio, ..) que vous allez suivre ? (exemple : nombre de références pertinentes lues, nombre d'utilisateurs dans les expériences, ...)

Indicateurs de production, i.e. pour rendre compte des productions faites au cours du travail de thèse.

- Qu'est ce que vous allez produire pendant votre thèse (exemple : livrables, protocoles



Notation

De 0 à 5 sur le, les outil(s) présenté(s)



Deux Guides



- Bien commencer son travail de thèse, le suivre et rédiger son introduction de thèse
- *Décrire le contexte d'enseignement, le terrain (EIAH, Sciences de l'éducation, ...)*



Deux Guides



- Bien commencer son travail de thèse, le suivre et rédiger son mémoire de thèse
- *Décrire le contexte de l'enseignement, le terrain (EIAH, Sciences de l'éducation, ...)*

Debriefing



Notation

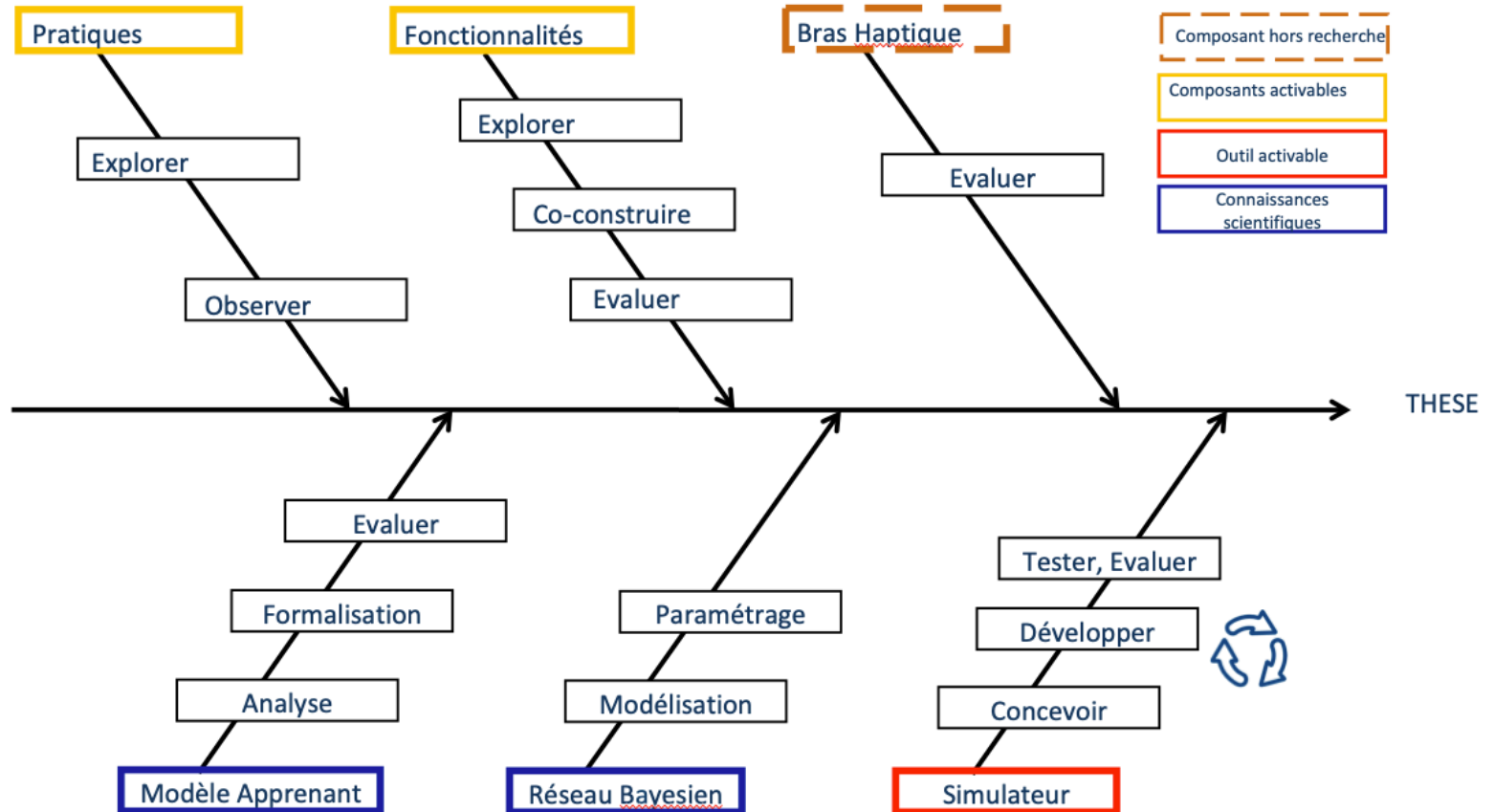
De 0 à 5 sur le, les outil(s) présenté(s)



- Fin de la 1^{ère} journée

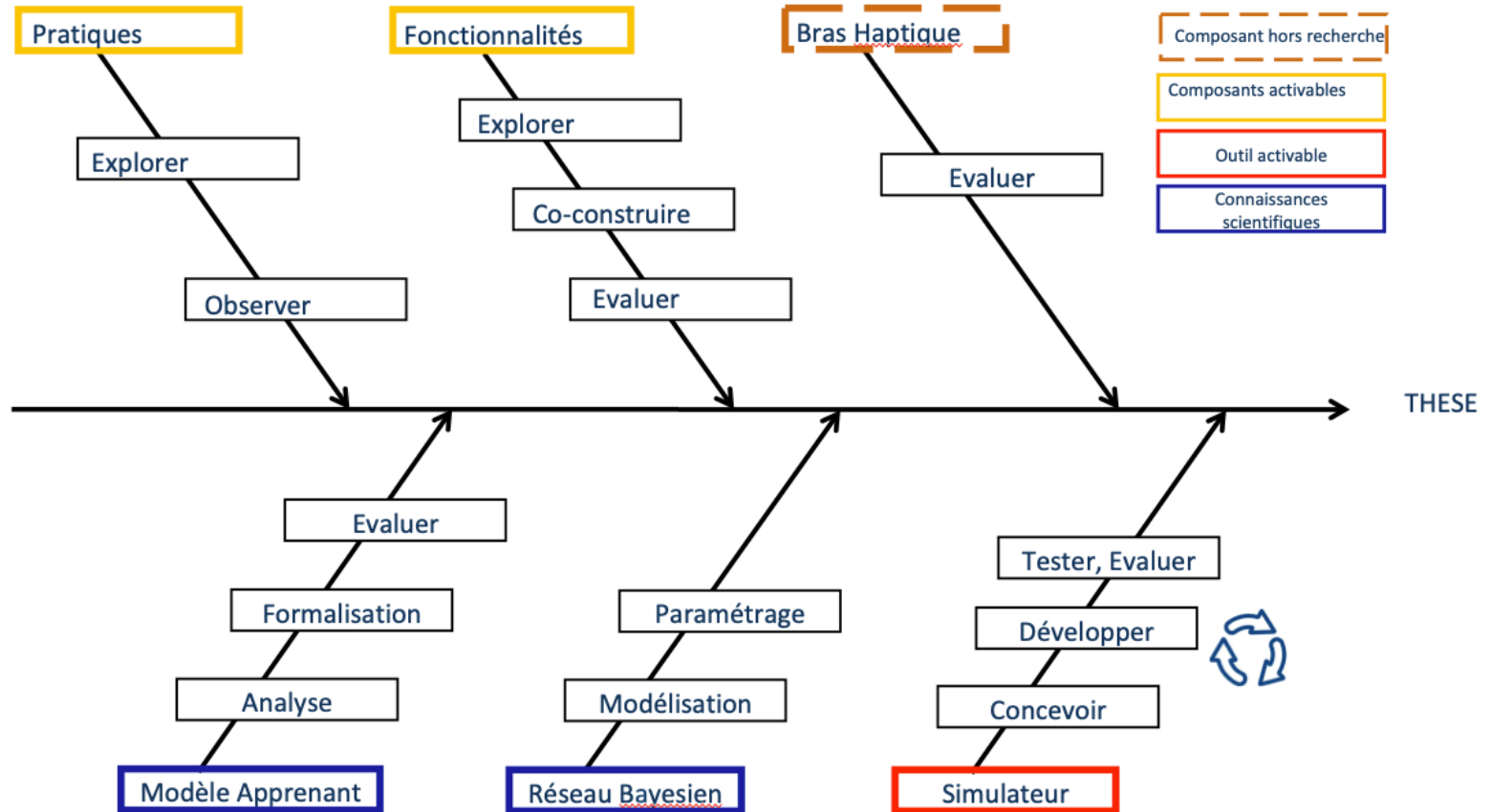


Une vue synthétique





Une vue synthétique





Notation

De 0 à 5 sur le, les outil(s) présenté(s)



- Tests sur 
Le cahier numérique collaboratif pour écrire et apprendre les sciences



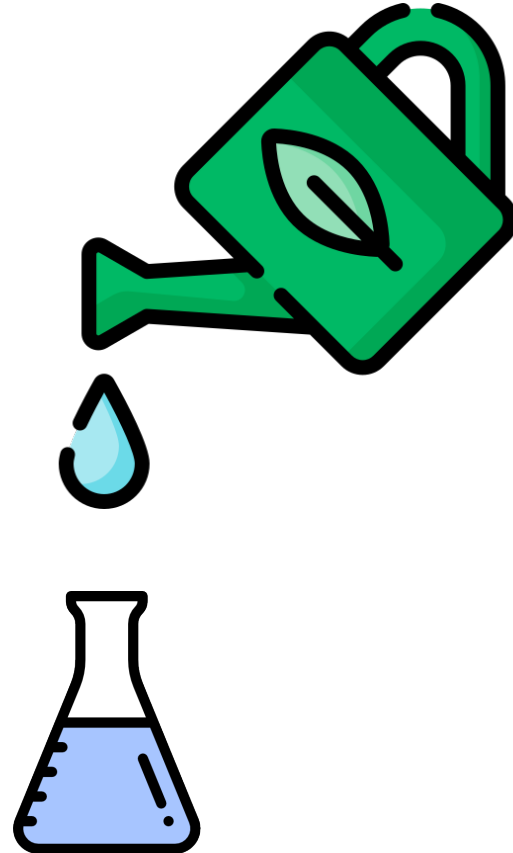
Deuxième journée

- Questions ?
 - Connaissances et Outils activables
 - Méthode Itérative et Tâches
 - Acteurs et Rôles
 - Indicateurs et mesures
 - Synthèse et Vision Globale



Mesurer avec l'humain ??

Avoir le bon outil comme
en biologie, chimie,
mécanique, etc ...





Les méthodes de production de données



- Quelles sont celles que vous connaissez ?
 - Les noter sur un post-it



Les méthodes de production de données pour les SHS



- Quelles sont celles que vous connaissez ?
 - Les noter sur un post-it
- Les classer ? Quelles différences ?

Méthode de production et analyse de données Qualitatives



- Pourquoi ? **Comprendre**, Identifier des phénomènes, Observer des utilisateurs, ..
- Quand ? Peu de connaissance sur l'objet de l'étude ou sur l'utilisateur
- Qui ? Des utilisateurs au profils variés, pour un maximum de variabilité
- Comment ? Des méthodes qui vont privilégier la narration
 - **Entretiens semi-directifs** : Faire exprimer un utilisateur sur l'objet de l'étude
 - **Observation in situ** : aller sur le terrain et observer ce qui se passe
 - **Focus-group** : réunion d'utilisateur, confrontation des idées
 - **Carnet de bords, social probes** : l'utilisateur note des informations relatives au sujet de l'étude
 - **Think aloud** : Faire verbaliser pendant une activité
 - **Auto-confrontation** : vidéo de l'activité, faire visionner à l'utilisateur et le faire verbaliser sur ces actions
- Données ? Vidéo, audio, schéma, des post-it, carnet de bord, des annotations, des documents du terrain ...
- Comment les analyser ?
 - Analyse thématique
 - Analyse textuelle (lexicale, sémantique, ...)



J.C.Creswell, Qualitative, Quantitative and mixed Methods Approaches. Research Design. Sage Publication. 2003

Paille et Muchielli. L'analyse qualitative, xxx

Statistique textuelle, L.Lebart et A.Salem, Dunod. 1994

Jean Copans, L'enquête ethnologique de terrain, Arland Colin, 2011



Méthode de production de données Quantitatives



- Pourquoi ? **Quantifier, Evaluer, Tester des hypothèses, quantifier des résultats du qualitatif**
- Quand ? On a des connaissances sur l'objet d'étude et l'utilisateur
- Qui ? Des utilisateurs « représentatifs » de la population à étudier, ou un recensement « tous les utilisateurs »
- Comment ? Des méthodes qui vont privilégier le dénombrement
 - **Questionnaire**
 - **Questionnaire baromètre**
 - **Traces d'activités**
 - **Tests utilisateurs**
- Données ? Fichiers de type ligne colonne (Individus statistiques x Variables), fichiers de traces, ...
- Comment les analyser ?
 - **Statistiques descriptives**
 - Statistiques inférentielles
 - Modélisation, Clustering, ..
 - **Visualisations**

D.Ray, Mesurer et développer la satisfaction client. Edition d'organisation
D.C.Howell. Méthodes statistiques en sciences humaines et sociale. (version française). Edition deBoeck.
G.Milliot. Comprendre et analyser les tests statistiques avec R. Edition deBoeck.
D.Grangé, L.Lebart. Traitements statistiques des enquêtes Edition Dunod.





Qualitatif // Quantitatif

A partir d'une problématique : poser des questions à explorer ou des hypothèses à valider

Le contenu des hypothèses ou des questions va guider le choix des méthodes
2 familles de méthodes de production et d'analyse des données

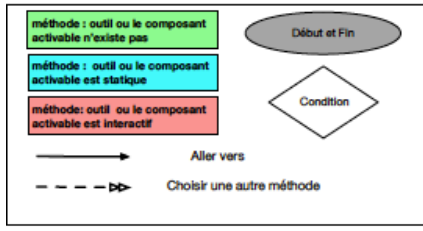
	Qualitatives	Quantitatives
	recenser le maximum d'informations, identifier des phénomènes.	Quantifier, valider ou invalider des hypothèses
Echantillon	Taille de 15 à 25 personnes. Privilégier la diversité des profils	Echantillon représentatif de la population étudiée, plan expérimental
Outils de prise de mesures	Observation in situ Entretien Focus group	Questionnaires Tests utilisateurs Traces d'activités
Données Produites	Enregistrement audio-video Schéma Annotations	Variables <u>qualitatives</u> : modalités ou catégories qui ne réfèrent pas à une quantité (e.g. catégorie professionnelle) Variables <u>quantitatives</u> : qui se réfèrent à une quantité ou à un étalon (e.g. durée, poids)
Analyse des données	Analyse thématique Analyse du discours Analyse lexicale CAQDAS Comparaison des schémas	Statistiques descriptives, visualisation des données, Inférentielles, tests d'hypothèses statistiques, Modélisation, tests d'adéquation du modèle, Clustering, Analyse temporelle, ...



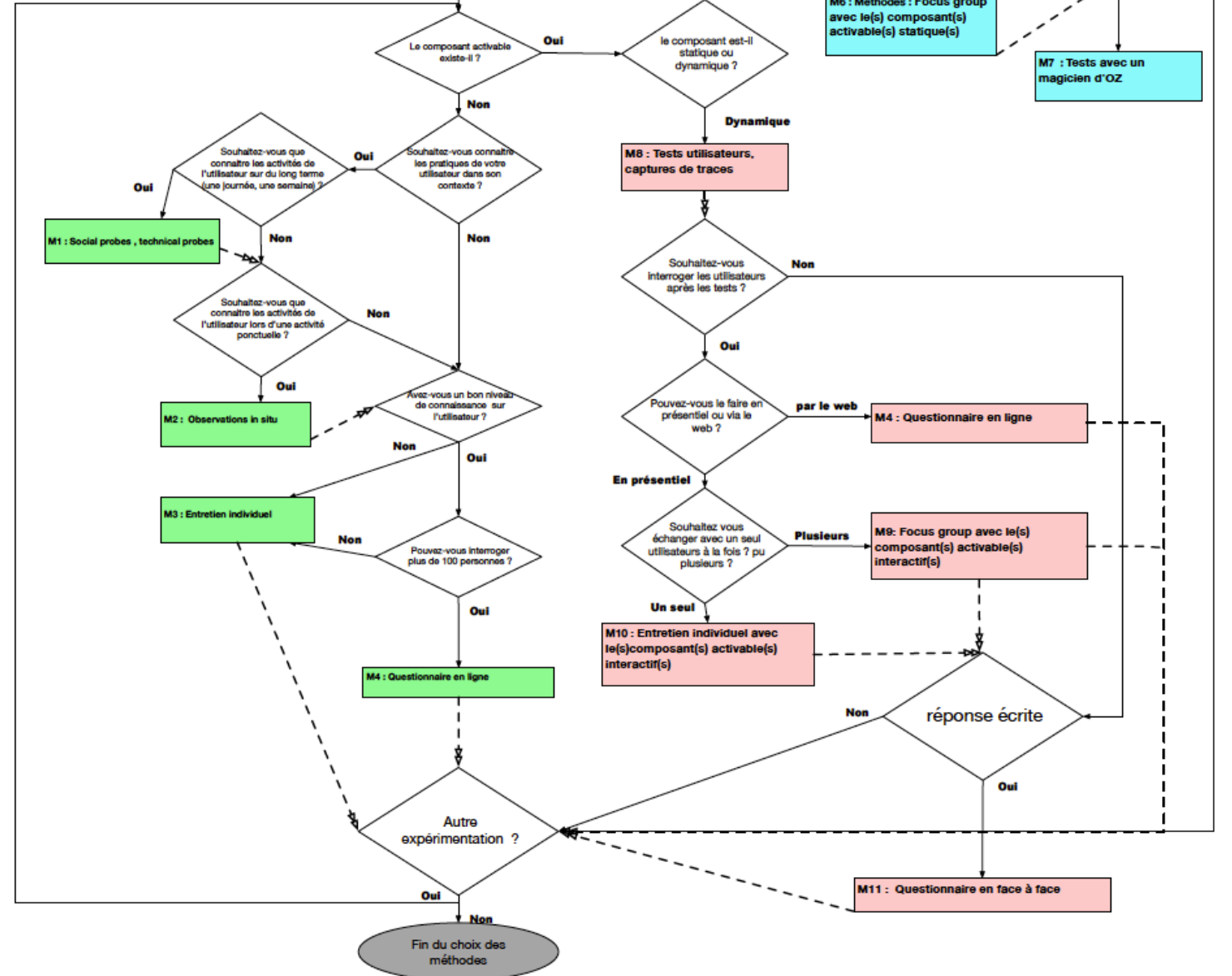


Comment expérimenter ?

- **Utilisateurs**
- Votre **niveau de connaissance** sur votre utilisateur (inconnu, connu)
- **Pourquoi** l'utilisateur : observer, co-construire, évaluer, ..
- Manière de joindre cet utilisateur, **lieu**
- **Outils activables** mis à disposition
 - Pédagogiques
 - Numériques
- Eventuellement des **composants activables**
- **Etat de ces outils** : dynamique, statique, n'existe pas
- **Temporalité** : les moments de la construction ou de l'évaluation
- Les **mesures** à prendre auprès des utilisateurs

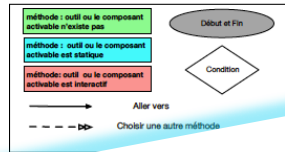


Pré-requis : Les objectifs de la recherche et des expérimentations sont définis. L'outil activable a été divisé en composants.

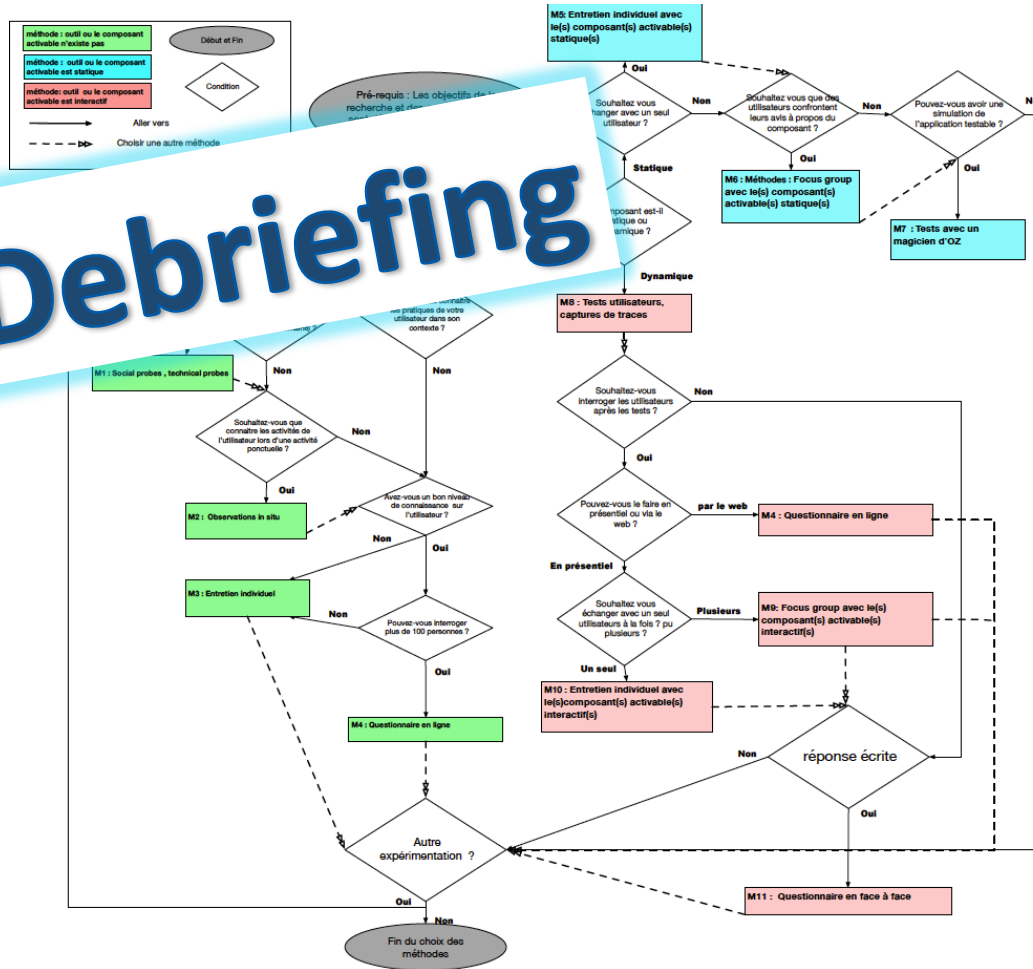




- **Par rapport à l'état actuel de votre recherche, répondre à ces questions**
 - **Quels sont vos utilisateurs ?**
 - Votre **niveau de connaissance** sur votre utilisateur (inconnu, connu)
 - **Pourquoi** l'utilisateur : observer, co-construire, évaluer, ..
 - **Lieu d'observation ou de test**
 - **Outils et/ou composants activables** mis à disposition
 - Pédagogiques
 - Numériques
 - **Etat de ces outils** : dynamique, statique, n'existe pas
 - **Temporalité** : les moments de la construction ou de l'évaluation
 - Les **mesures** à prendre auprès des utilisateurs
 - Pour Comprendre
 - Pour Quantifier
- **Choisir des méthodes de production des données avec l'aide des logigrammes MATUI**



Debriefing



Protocole expérimental

6- Protocole expérimental

Date à laquelle le document est créé // Dates des modifications successives du document

Auteur(s) du document

Objectifs et buts de l'expérimentation

- Donner un nom à l'expérimentation
- Décrire à quoi cette expérimentation va servir

Si l'évaluation se situe en contexte de recherche

- Décrire les connaissances scientifiques concernées par l'expérimentation
- Indiquer en quoi l'expérimentation va faire évoluer la connaissance scientifique
- Indiquer l'état de la connaissance scientifique (à élaborer, à améliorer, à évaluer, etc...)

Questions ou hypothèses et mesures

- Indiquer les questions et /ou les hypothèses qui devront trouver des éléments



Protocole expérimental

6- Protocole expérimental

Date à laquelle le document est créé // Dates des modifications successives du document

Auteur(s) de

Debriefing

Expérimentation
Expérimentation

à quoi cette expérimentation va servir

Si l'évaluation se situe en contexte de recherche

- Décrire les connaissances scientifiques concernées par l'expérimentation
- Indiquer en quoi l'expérimentation va faire évoluer la connaissance scientifique
- Indiquer l'état de la connaissance scientifique (à élaborer, à améliorer, à évaluer, etc...)

Questions ou hypothèses et mesures

- Indiquer les questions et /ou les hypothèses qui devront trouver des éléments



Notation

De 0 à 6 sur le, les outil(s) présenté(s)

Créer un outil de mesures en SHS

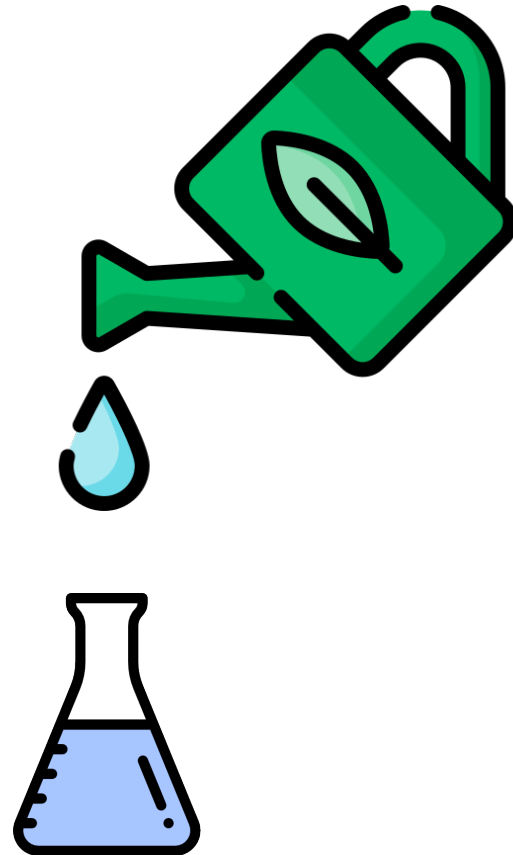
Entretien

Et/Ou

Questionnaires

Et/Ou

Traces



Entretien

Etudier les pratiques des enseignants avec des pédagogies actives et des outils numériques

Enseignement : L1, en Biologie, UE sciences expérimentales. Responsable d'UE veut introduire une nouvelle plateforme existante

Le responsable d'UE est intéressé par les méthodes Student Centered Learning (Freeman 2014)



Entretien

Lister les mesures à prendre

- Pratiques des pédagogies actives
- Avantages, inconvénients de ces pédagogies
- Usage du numérique dans la classe
- Avantages, inconvénients du numérique
- Perception *a priori* de la nouvelle plateforme
- Perception sur les pédagogies actives Student Centered Learning (autonomie, collaboration, organisation, ...)

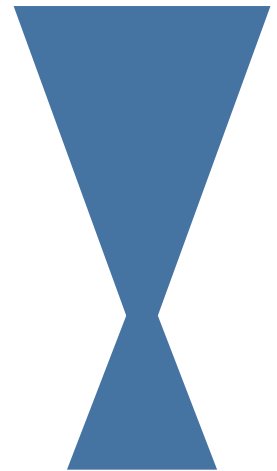
Organiser et ordonner les mesures par grands thèmes

Pédagogies actives

Usage du Numérique

Méthode Student Centered Learning

La nouvelle plateforme



Guide d'entretien 1/4

C'est un document rédigé, qui est testé

Les questions s'enchaînent, pour arriver au point essentiel de l'entretien

Faire une introduction, sur la raison de l'entretien et comment il va se dérouler

- Il faut présenter au sujet pourquoi nous lui avons demandé de participer à cet entretien
- En faisant attention à ne pas lui donner trop d'informations qu'il cherchera à valider pour vous faire plaisir pendant l'entretien
- lui indiquer les étapes de l'entretien
- Lui dire que c'est votre proposition qui est évaluée et pas lui
- Lui dire d'exprimer tout ce qu'il a envie de dire par rapport à la proposition

Guide d'entretien 2/4

- **Du spontané au suggéré**
 - Laisser libre cours au discours de la personnes
 - Si elle ne s'exprime pas sur un sujet alors lui poser la question
 - Lors de la rédaction de la grille : **envisager tous les cas de figures.**
 - *Si le sujet a parlé de la méthode Student Centred Learning (peu probable), s'il parle de laisser autonome les étudiants, ...*
 - *Si elle n'en a pas parlé*
- Le guide doit contenir les **remarques pour l'enquêteur**
 - *ENQUETEUR : Noter le déroulement du cours*
 - *ENQUETEUR : Noter les outils numériques cités*
- Le guide se termine en demandant les caractéristiques des participants
 - Année de naissance, vous êtes un homme/ une femme, votre profession, niveau d'étude, domaine d'activités, ancienneté dans la profession , votre type d'habitat, nombre d'enfants ... et toutes informations utiles pour répondre à vos questions.

Guide d'entretien 3/4

Les domaines qui sont abordés sont parfois complexes pour cela deux possibilités non exclusives :

Proposer des constats pour lancer l'entretien

Selon la directive, XYZ de 2018, Les pédagogies actives sont de plus en plus présentes dans l'enseignement supérieur. Elles sont souvent utilisées dans des contextes de spécialistes.

Proposer des définitions sur des termes qui ne doivent pas être ambigus

*Aujourd'hui, on entend beaucoup parler de **réseaux sociaux**. Dans notre étude, il s'agit des réseaux de type facebook où chacun est libre de déposer des informations qui le concerne et de les partager avec d'autres personnes. L'utilisateur de facebook peut également créer des groupes de personnes auxquels il réservera certains types d'informations .*

Guide d'entretien 4/4

- **Rédaction des questions**
 - Les questions doivent permettre des **réponses larges et ouvertes** :
 - *Avez vous déjà mis en place des pédagogies actives ?*
 - *Dans quelles circonstances avez vous mis en place des pédagogues actives ? (oui)*
 - **Ne jamais suggérer** de réponses dans les questions
 - *Pensez-vous que ce dispositif soit utile ? (non)*
 - *Que pensez-vous de ce dispositif ? (oui)*

Entretien : Passation

- **Faire un pilote**, un test avec une personne pour identifier si les questions sont comprises et si l'ordre des questions est suffisamment fluide.
- **Connaître** son guide d'entretien, à partir de 3 ou 4 passations, on le connaît
- Prévoir un **enregistrement** des entretiens, demander l'accord au moins oral aux personnes
- **Prendre des notes et laisser s'installer des silences**
- Lors de la passation de l'entretien, l'interviewé peut parfois aborder les différents éléments prévus par la grille sans forcément respecter l'ordre que vous avez prévu
- Peu importe : l'entretien est semi-directif ce sera l'enquêteur de conduire l'entretien sans pour autant être dirigiste.
- Laisser le sujet exprimer ses idées même si elles ne sont pas dans l'ordre de la grille. L'enquêteur doit veiller à ce que tous les points soient abordés.
- **Ne pas se justifier** en cours d'entretien, attendre la fin de l'entretien pour expliquer à

Questionnaire 1/2

- **Généralités**

- Doit comporter une introduction,
- La structuration reste la même du plus global au plus précis,
- Les modalités des réponses sont définies, à l'exception des questions ouvertes,
- Les variables sociodémographiques se situent à la fin.

- **Types de questions**

- **fermées** (close-ended question) : les réponses aux questions sont fixées, il n'est pas possible de rajouter des modalités de réponses
- **numériques**
 - Combien de voitures possédez-vous ?
- **semi-ouvertes** (semi-open question): questions pré codées, avec l'option "autre" qui permet d'ajouter des items de réponse.
 - Quels sont vos sports favoris :
 - Course à pied
 - Ski
 - Natation
 - Autres précisez
- **ouvertes (open-ended question)** : l'enquêté répond librement, pas de propositions de réponses a priori.

Questionnaire 2/2

- Réponse unique sur une échelle de réponse (scale) :
 - Echelle de Likert (Likert scale) La personne interrogée exprime son degré d'accord ou de désaccord vis-à-vis d'une affirmation. L'échelle contient en général cinq ou sept choix de réponse qui permettent de nuancer le degré d'accord. Pas du tout d'accord, Pas d'accord, Ni en désaccord ni d'accord, D'accord, Tout à fait d'accord. Pour les échelles impaires, le niveau central permet de n'exprimer aucun avis, tandis que les échelles paires (par exemple à quatre modalités) sont dites « à choix forcé ». (5)
- Réponse à choix multiples (multiple-choice questions) :
 - Choix de plusieurs modalités de réponse : (exemple j'utilise mon portable pour 1- Téléphoner 2 - Envoyer des SMS 3- Pour tenir mon agenda, 4-Pour me réveiller, 5- Autres)
 - À opposer aux questions à choix unique, pas de possibilité de rajouter une modalité de réponse
- Réponse avec une demande de classement :
 - les réponses sont ordonnées selon une préférence ou un ordre d'importance. :
 - *Pour le produit XYZ vous classerez les 3 caractéristiques suivantes en partant de ce vous considérez comme son point le plus fort vers son point le plus faible :*
 - *Facilité d'utilisation*
 - *Fiabilité d'utilisation*
 - *Son prix*



Traces

- Un guide de production des traces
 - Liste des variables utiles pour l'analyse
 - Temporelle : Time stamp, **et** no d'ordre
 - Utilisateurs : Code enseignant, apprenant, ...
 - Actions avec l'outil : nom de l'action, ...
 - Domaine de validité de ces variables
 - Age compris entre 18 et 25 ans
 - Heure de connexion entre 14h et 18h



- Rédiger quelques questions pour un guide d'entretien sur votre cas ?
- Rédiger quelques questions pour un questionnaire ?

Tracer l'activité des focus groups avec des activités des utilisateurs

No d'étape : Indiquer un numéro d'étape

Objectif de l'étape : Objectifs généraux de l'étape

Hypothèses et questions en lien avec cette étape : Indiquer les questions ou les hypothèses expérimentales abordées dans cette étape. Ces questions et hypothèses correspondent à celles posées dans le protocole expérimental.

Activités réalisées : Lister activités faites pendant cette étape. (p.ex., présentation du modèle, tests utilisateur, lecture du dictionnaire)

Responsables de la passation : Nom de la personne en charge de faire la passation de cette étape

Acteurs concernés: Chercheur/Méthodologue/Développeur/Utilisateur/autres préciser

Composant(s) activable(s) : Indiquer le(s) composant(s) activable(s) concernés par l'étape.

Format du composant : Indiquer l'état d'avancement de(s) composants : dynamique/statique/ non manipulable

Responsables du (des) composant(s) : Noms des personnes en charge de faire le(s) composant(s)

Matériel expérimental : Liste des matériels expérimentaux nécessaires à cette étape (p.ex., guide d'entretien, d'annotation). Indiquer le rôle de ces documents (présentation, recueil).

Responsable du matériel expérimental : Noms des personnes en charge de faire le(s) matériels expérimentaux

Durée (mn) prévue Indiquer une durée approximative de l'activité

Heure début L'heure de début de l'étape

Heure fin L'heure de fin de l'étape

Durée cumulée prévue (hh:mm): Cumul de la durée des activités



Processus d'évaluation longitudinal, (Mandran, et al 2019, EIAH)

- Les mesures de la littérature
 - Utilisabilité (SUS Brooke)
 - Student centered Learning (Hannafin 2012, Freeman 2014)
- Des spécifiques
 - Satisfaction, attentes, pratiques
- Des outils de mesures

Mesures quantitatives	Mesures qualitatives
Questionnaires baromètres (QB) Questionnaires de satisfaction (QS) Traces d'activité	Entretiens semi-directifs Analyse des missions Débriefings des équipes pédagogiques

- Le processus

	Avant utilisation	Année 1	Entre années 1 & 2	Année 2	Entre années 2 & 3	Année 3	Après utilisation
Responsables d'UE	QB0 Entretien initial		QB1 Analyse missions 1		QB2 Analyse missions 2		QB3 Analyse missions 3 Entretien final
Autres enseignants	QB0		QB1 Débriefing 1		QB2 Débriefing 2		QB3 Débriefing 3
Etudiants		Traces 1	QS1	Traces 2	QS2	Traces 3	QS3





Après la production, l'analyse et la rédaction





Analyse Qualitative

Entretien : Analyse thématique simplifiée

Méthode en Sociologie :

Exemple

Retranscrire les entretiens (1h => 6h)

A partir de la problématique et de la grille d'entretien, créer une grille thématique,

A partir de la **retranscription des entretiens**, identifier les thèmes qui correspondent à la problématique et aussi les thèmes qui émergent

Les thèmes sont larges, il convient de créer des sous-thèmes

Chacun des sous-thèmes est illustré par des verbatims (extraits du discours associé au code du sujet ayant prononcé de discours)

Synthèse pour chaque thème

Thème « Utilisation du numérique dans l'enseignement »

Thème nouveau « Utilisation que du numérique »

Sous- thèmes de « utilisation du numérique »

« utilisation volontaire »

« utilisation contrainte par le responsable »

Verbatim du sous- thème « utilisation contrainte par le responsable »

« c'est le responsable de cette UE cette année qui a choisit de mettre cela en place ans les cours, je ne suis pas certain que ce soit utile et que cela change quelque chose »(S1) »



Analyse Qualitative

Schéma

Créer une grille d'analyse pour savoir ce qui doit être repéré par rapport à la problématique

Photographier chacune des productions

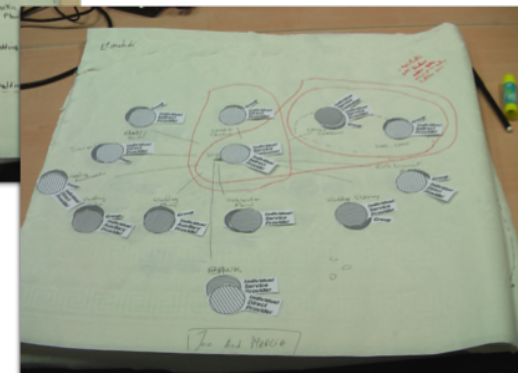
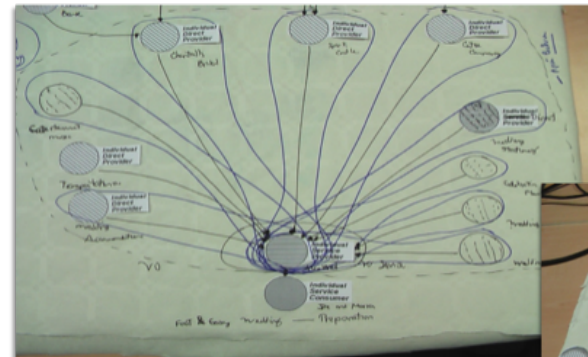
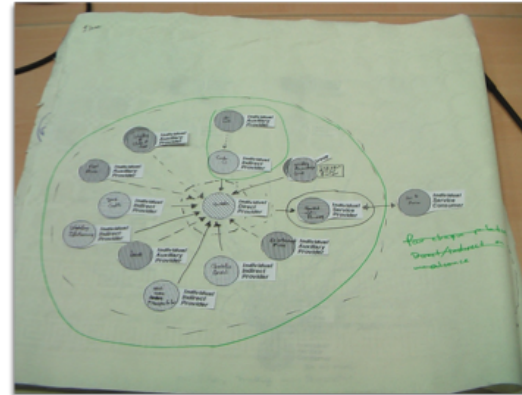
Repérer les ressemblances/les différences entre les représentations

Repérer ce qui confirme les objectifs de l'étude

Repérer les apports de la représentation des sujets par rapport à la proposition

Visionner et écouter les films

Conduire une « analyse thématique simplifiée »





Analyse Qualitative

Annotation Audio ou Video

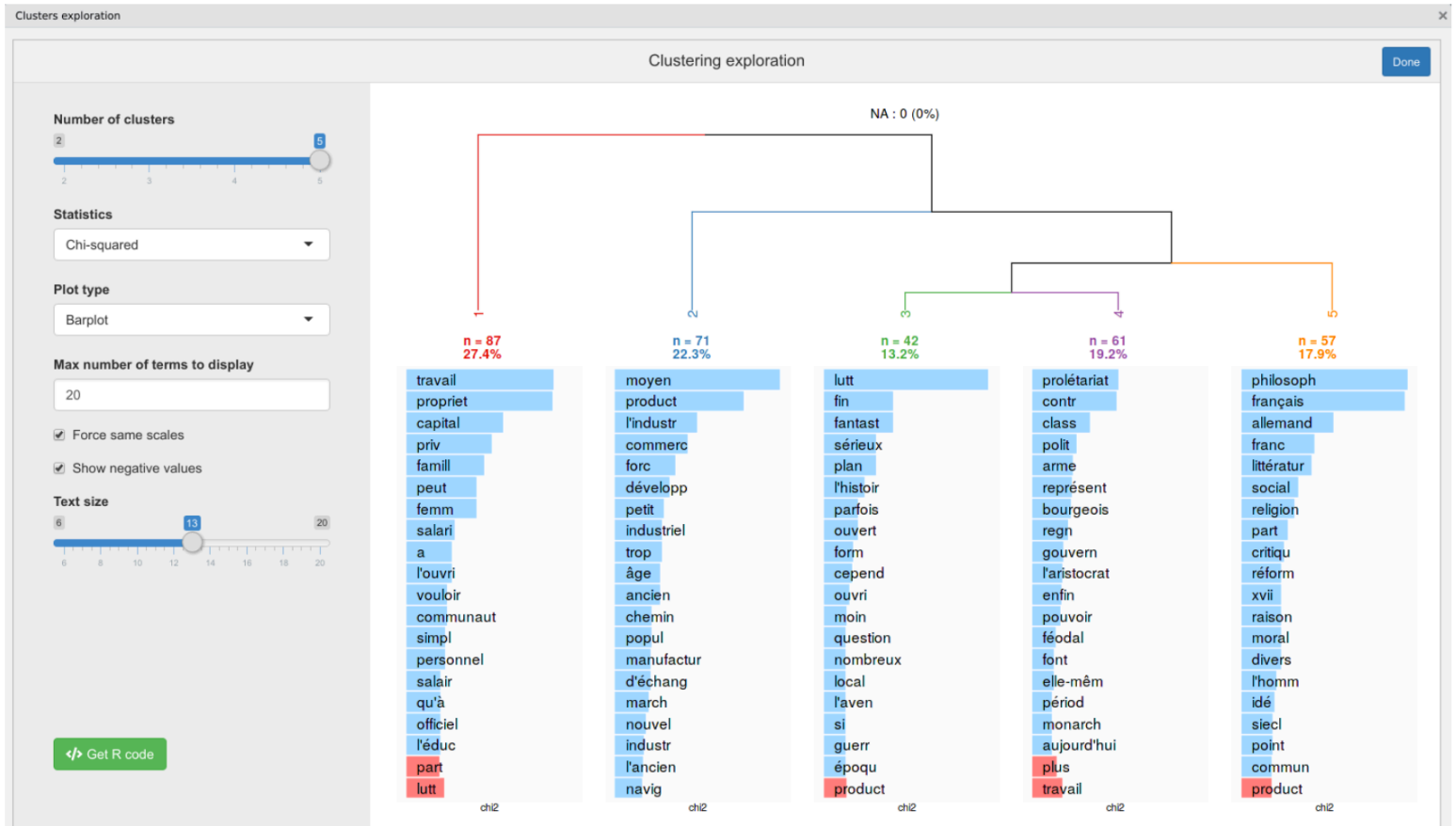
- Créer une grille d'analyse pour savoir ce qui doit être repéré par rapport à la problématique.
- Utiliser des logiciels d'annotations de video
- Analyser les annotations
 - Statistiques descriptives
 - Extraction de verbatims
- Faire une synthèse

The screenshot displays the ELAN 3.9.1 software interface. The main window shows a video of two people sitting and talking. The interface includes a menu bar (File, Edit, Annotation, Tier, Type, Search, View, Options, Window, Help) and a toolbar with various playback and editing controls. The 'Audio Recognizer' tab is active, showing parameters for 'Tag vowels (volume peaks of voiced timespans)'. The 'Parameters' section includes sliders for 'Pitch ceiling [Hz]' (set to 604.8) and 'Intensity change [dB] to start/end a peak (0.5 - 5.0)' (set to 2.0). The 'Progress' bar indicates the recognizer is 'Ready'. Below the video, there are several tracks: a pitch contour plot, an intensity plot, and a transcription tier. The transcription tier shows the following text: 'and so he climbs up a tree and he starts with the ladder | and he starts picking pears off the tree | and he puts the pears into an apron'. Below the transcription, there are gesture annotations labeled 'gestu', 'gestu 4', 'gestu', 'gesture 6', 'gesture 7', 'gesture', 'gesture 9', and 'gesture 1', with corresponding hand positions (R, B, B, B).



Analyse textuelle

Alceste
RThemis
Rainette





Analyse Qualitative

Points forts et Points faibles

Outil : Grille qui liste les éléments relatifs à une proposition et qui à deux colonnes une pour noter les points forts et l'autre les points faibles

Exemple :

Table Number :

First and last name :

<i>List the strengths and weaknesses on the following</i>	Strengths	Weaknesses
the work performed with the method	+	-
the collaborative work enabled with the method	+	-
The creativity enabled with the method	+	-
the serious gaming of the method	+	-
The software compared to card game	+	-



Analyse Qualitative

Points forts et Points faibles

Méthode :
Saisir l'ensemble des réponses : Ligne=participants, Colonne = réponse aux questions, 2 colonnes par éléments

Experimentation menée en 2011 lors de l'école d'été sur les GBL. LIG/METAH Pernin et al							
				The collaborative work enabled with the method		The method ability to give new ideas	
				What you like +	What you don't like -	What you like +	What you don't like -
saisie	su	tab	ses				
CM	1	1	2	It's easy to work collaboratively	Difficult to agree in the selected cards	very intuitive to create many different ideas	
CM	2	1	2	immediate reviewing	maybe it can be difficult to make the game with people with different backgrounds	the fact of having one expert facing newbies on the subject	
CM	3	1	2	The cards are good support to communicate with other members of the team. Rapid modification / iteration	If you move the table everything moves and you have to replace the cards / jewels	the existing game principles and cards help us to set up some foundations	can also avoid some possible activities or scenarios because of the formalism. If / esle ? Random ?
CM	1	2	2	great for triggering discussion. The terminology might need to be tightened. Capitalize might not work that well.	Potential for one person to dominate the interaction physically (but not completely)	it was very helpful. I have a lot of new ideas I may follow up. Perhaps more creative game scenarios in activity 1 would be useful for inspiration.	need more time afterwards for reflection + exploring ideas that arise
CM	2	2	2	we shared many different ideas	we should take some roles so we could be more organized	we produced a lot of ideas	
CM	3	2	2	cards facilitate greatly the collaborative approach.		this method brings new ideas because students start on groupware gaming activities and their synchronisation and communication issues.	the method should be extendible ; introduce profiles.
CM	4	2	2	I like combination of different opinions of participants		i like to see differents point of view	
CM	1	2	3	échanges nombreux ; matériel (forte flexibilité dans la manipulation des différents éléments) a aidé à la collaboration	auto-régulation => en cas de désaccord, pas de moyen prévu pour aider à trouver des compromis	grandes variétés de choix; Prévoit différentes méthodes de travail, etc.	cadre tout de même très normatif trop contraignant, il est difficile de se conformer aux cartes dans la phase de conception globale
CM	2	2	3	la répartition des rôles ; le retour collectif	pas assez détaillé dans la répartition des rôles	l'encadrement proposé	
CM	3	2	3	visualisation de la progression, des itérations	n'aide pas forcément à la collaboration quand les mécaniques de jeu sont machées dès la début	on essaie de partir des ressorts de jeu	donner des exemples de jeu pour faciliter les idées





Analyse Qualitative

Points forts et Points faibles

The collaborative work enabled with the method

Les points positifs

3 cases vides (dont 2 vides aussi côté points négatifs)

Méthode :
Faire une « analyse
thématique simplifiée » et une
synthèse

1. Concernant la méthode		
La méthode facilite et stimule la collaboration des formateurs	great for triggering discussion The method promotes collaboration. Everyone participating. Good to develop team spirit Easy, nice, good way	14
Elle permet de tous participer, de partager des idées différentes et d'être créatifs collectivement	very good- helped to see each participants point of view and thus come up with a collective work Effective debating, compromising, selection and implementation We all participate in the creation process by bringing ideas and say what we think is not good at other ideas emergence of the solution by speaking and interacting confront, exchange, share ideas, créativité, débat, opinion, brainstorm, open discussion friendly	17



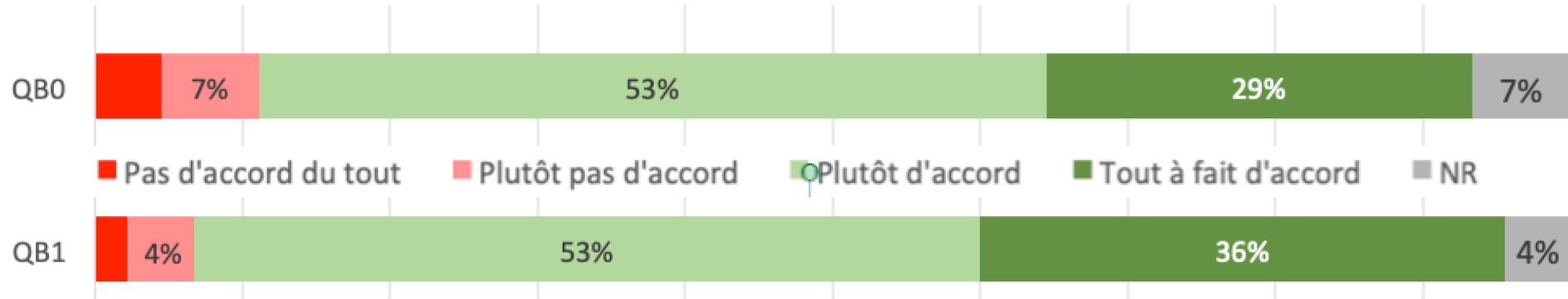
Analyse Quantitative

- Cycle de vie des données => importance du pré-traitement
- Echantillon
- Type de variables => quali/quant/ordinales
- Rôles des variables => explicatives et à expliquer
- Règles de décision sur le choix de la méthode



Analyse Quantitative

"Je pense que LabNbook sera (est) utile pour mes pratiques d'enseignement"

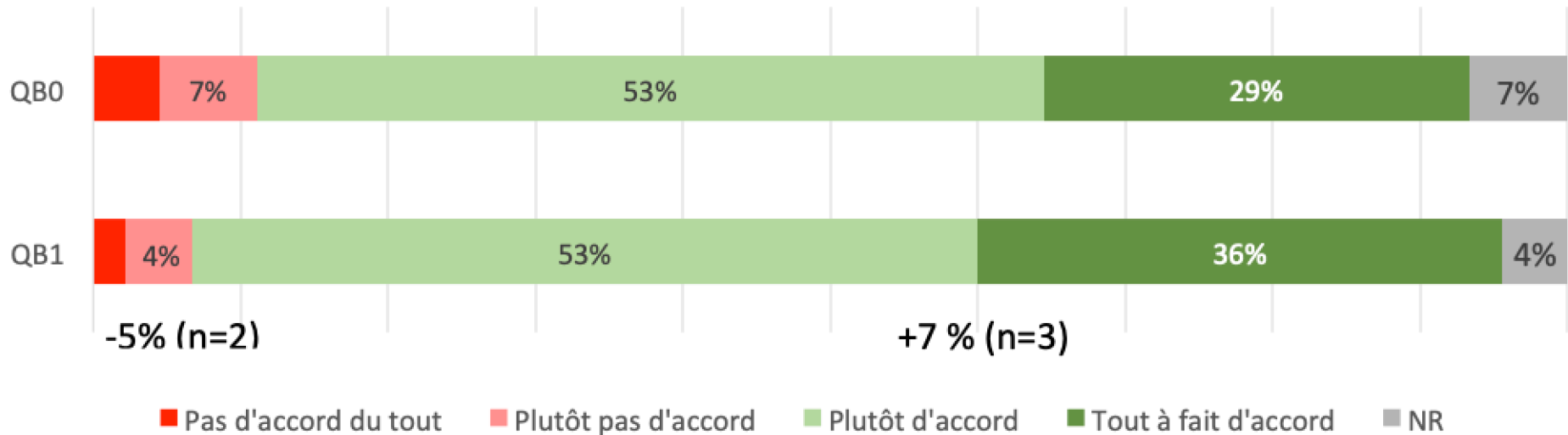


Votre Analyse ??



Analyse Quantitative

"Je pense que LabNbook sera (est) utile pour mes pratiques d'enseignement"



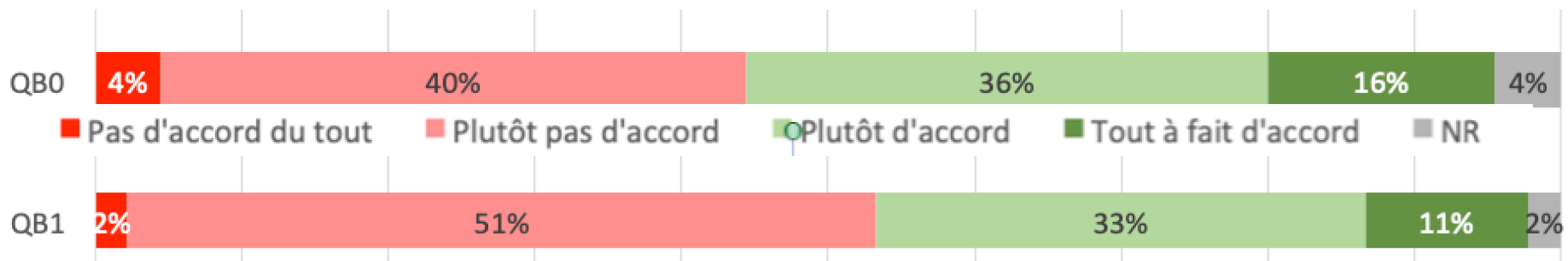
- Evolution positive concernant l'utilité de LabNbook : + 7% d'enseignants sont « *tout à fait d'accord* » avec la proposition

Les chiffres sont affichés en % pour une meilleure lisibilité, mais l'effectif est de 45 personnes, c'est pourquoi nous indiquons les effectifs à coté des %



Analyse Quantitative

« L'utilisation de LabNbook me demandera (me demande) un investissement important en terme de formation »

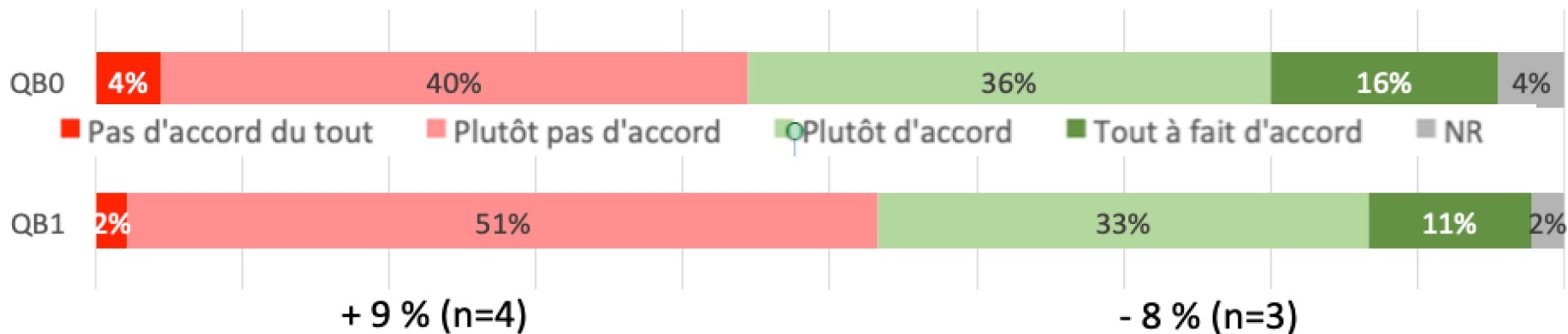


Votre Analyse ??



Analyse Quantitative

« L'utilisation de LabNbook me demandera (me demande) un investissement important en terme de formation »

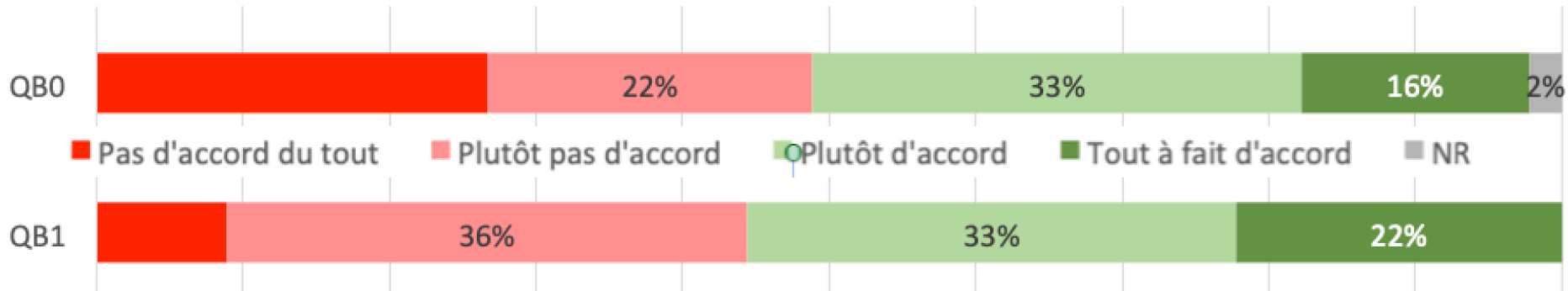


- Evolution positive également concernant le coût en terme de formation : après un semestre d'utilisation, + 9 % d'enseignants pensent que LabNbook ne nécessite pas un investissement important en terme de formations



Analyse Quantitative

« Pour cette formation, je permets (j'ai permis) aux étudiants de travailler de manière itérative »



Votre Analyse ??



Et en bonus, quelques de mots de stats ...

Méthodes statistiques en sciences humaines, Howell DC : 6^{ème} édition, 2015. (traduit de l'anglais) Edition DeBoeck.

*De l'échantillonnage, **Statistiques descriptives, Visualisations, Corrélations, Statistiques Inférentielles, Modèle linéaire généralisé (ANOVA, régression, ...)**,*

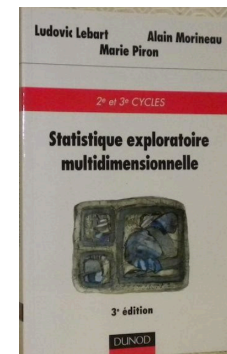
Analyse factorielles simples et multiples. B.Escoffier & J. Pages. 6^{ème} édition. 2016. Edition Dunod.

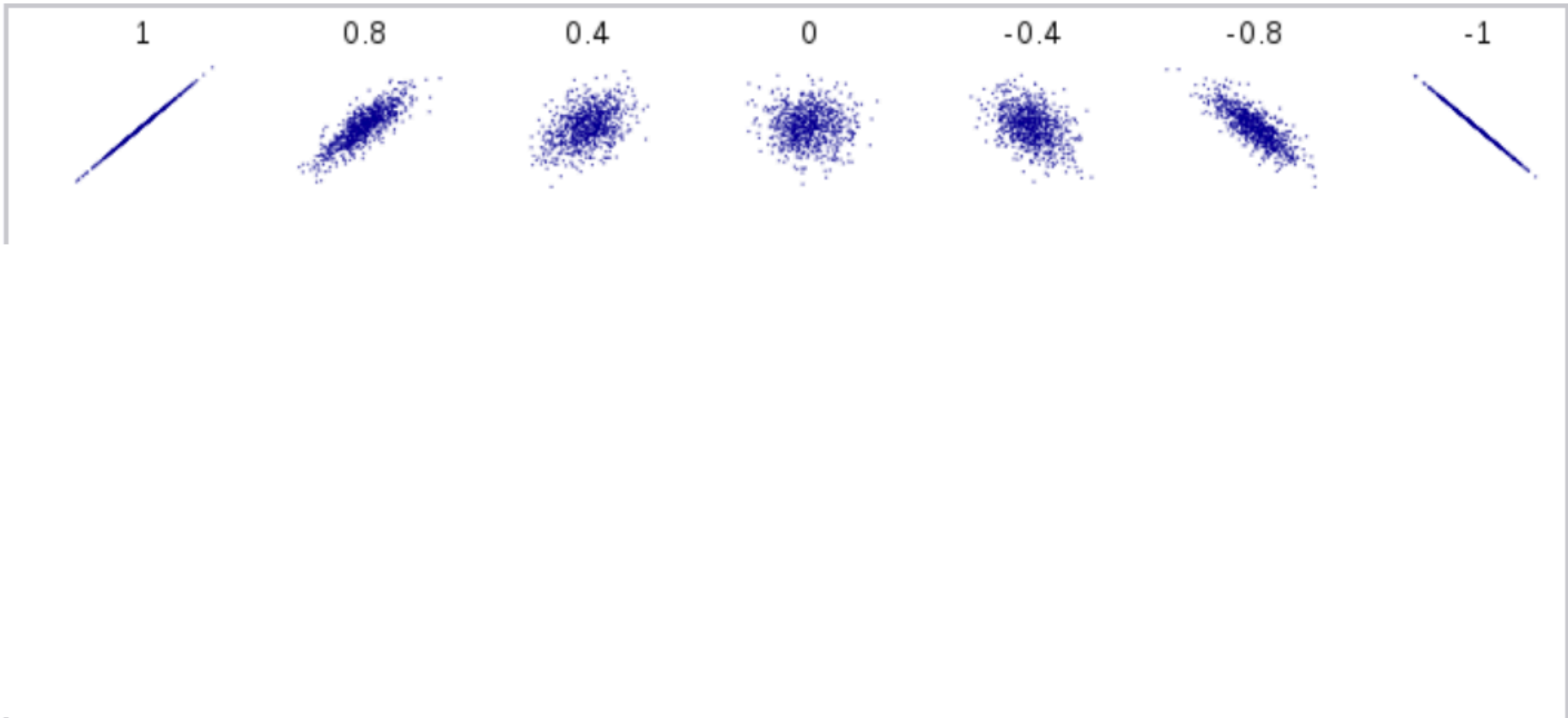
Statistiques exploratoires multidimensionnelles. L.Lebart, A.Morineau, B.Piron. Edition Dunod.



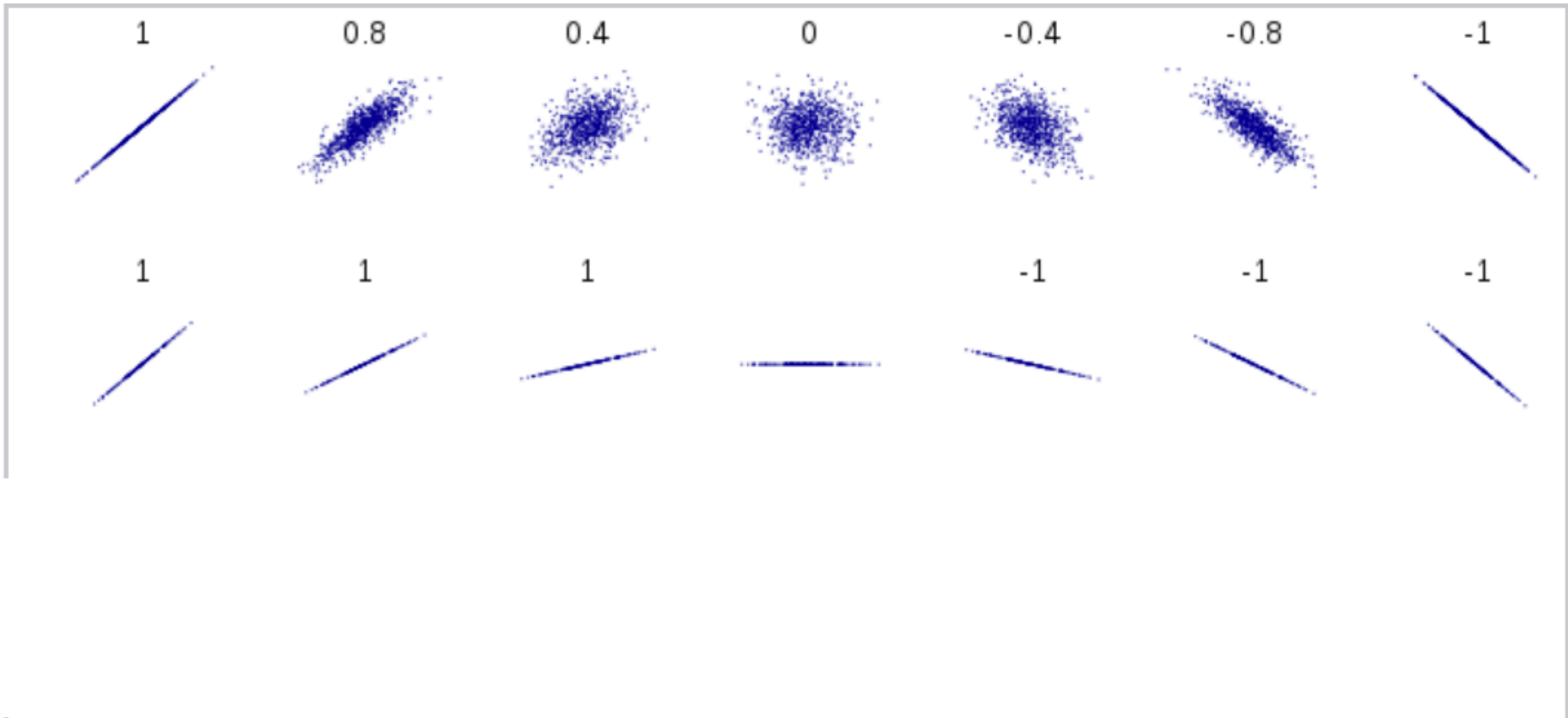
Masters
Écoles d'ingénieurs

DUNOD

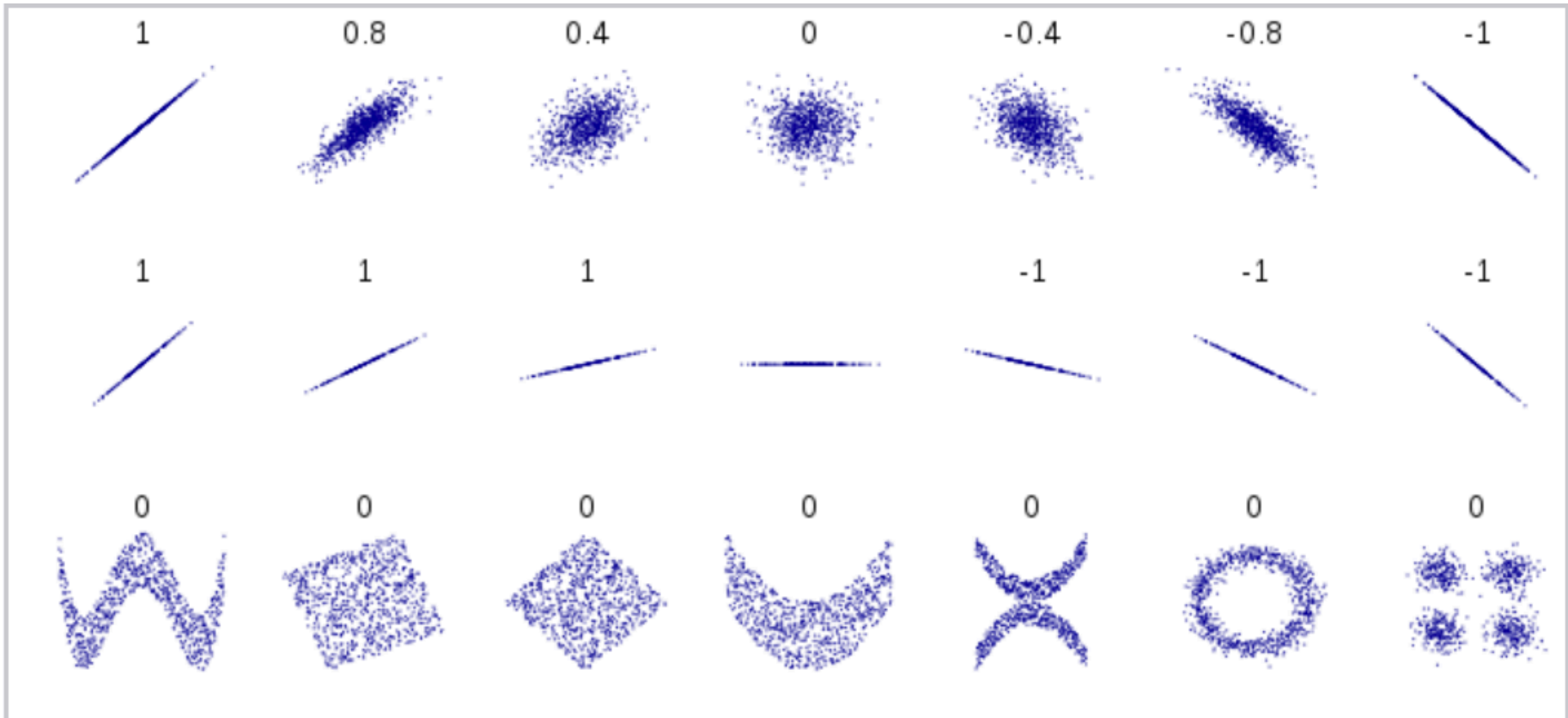




Exemples de coefficients de corrélation : corrélation linéaire pour les deux premières lignes, non linéaire pour la troisième ligne.



Exemples de coefficients de corrélation : corrélation linéaire pour les deux premières lignes, non linéaire pour la troisième ligne.



Exemples de coefficients de corrélation : corrélation linéaire pour les deux premières lignes, non linéaire pour la troisième ligne.



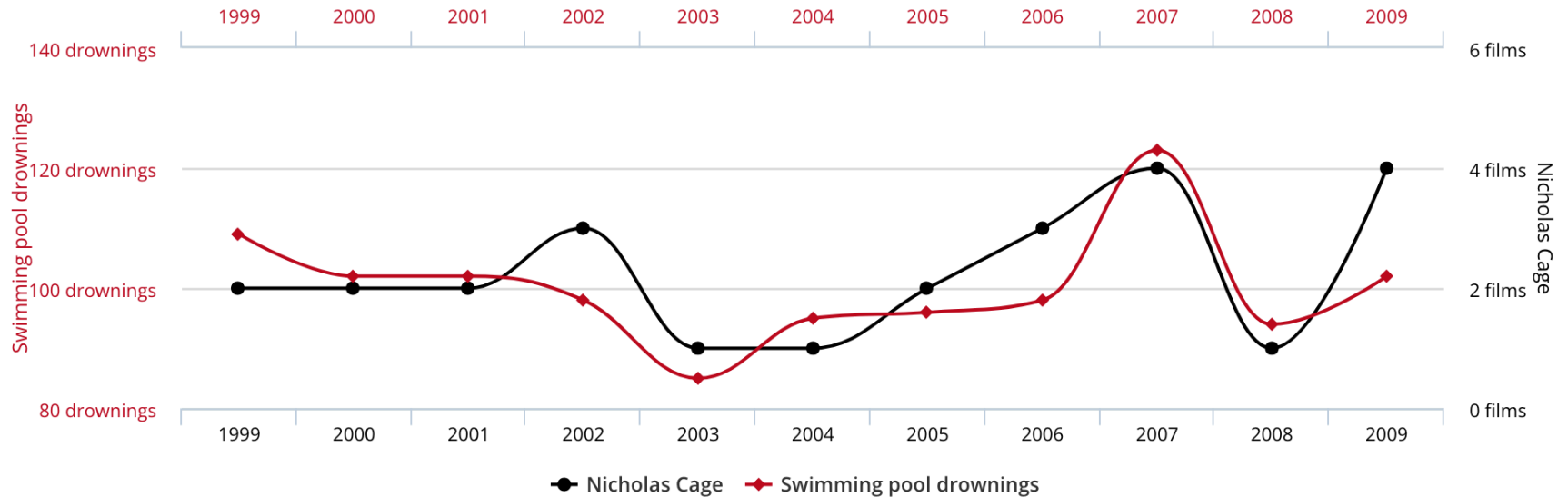
Corrélation et causalité

- Deux variables/mesures sont **corrélées** si
 - Elles évoluent dans le même sens (corrélation positive : quand l'un augmente, l'autre aussi)
 - Ou dans un sens opposé (corrélation négative : quand l'un augmente, l'autre diminue).
- Mais pas forcément de causalité entre ces deux variables/mesures



Number of people who drowned by falling into a pool correlates with Films Nicolas Cage appeared in

Correlation: 66.6% (r=0.666004)



Data sources: Centers for Disease Control & Prevention and Internet Movie Database

tylervigen.com

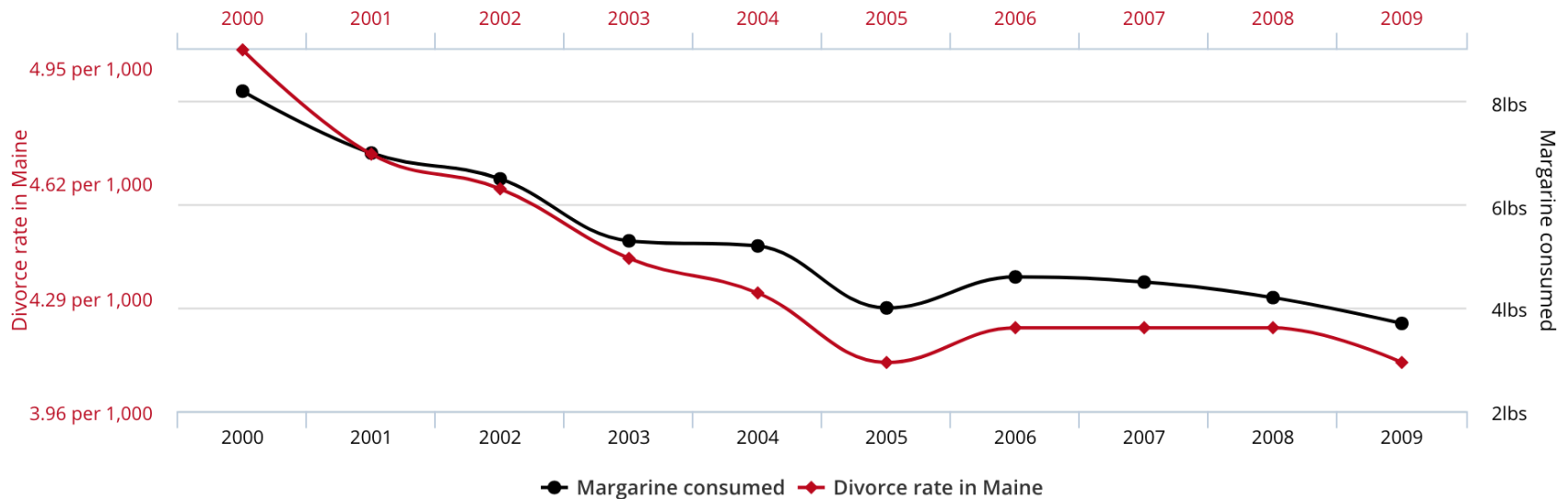


<http://tylervigen.com/spurious-correlations>



Divorce rate in Maine correlates with Per capita consumption of margarine

Correlation: 99.26% (r=0.992558)



tylervigen.com

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



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



Corrélation et causalité

- Deux variables/mesures sont **corrélées** si
 - Elles évoluent dans le même sens (corrélation positive : quand l'un augmente, l'autre aussi)
 - Ou dans un sens opposé (corrélation négative : quand l'un augmente, l'autre diminue).
- Mais pas forcément de causalité entre ces deux variables/mesures
- La **corrélation** est donc un lien entre deux variables/mesures , sans que pour autant la variation de l'une soit **la cause** de la variation de l'autre.



Variable latente

Espérance de vie et consommation de viande en 2014 par pays

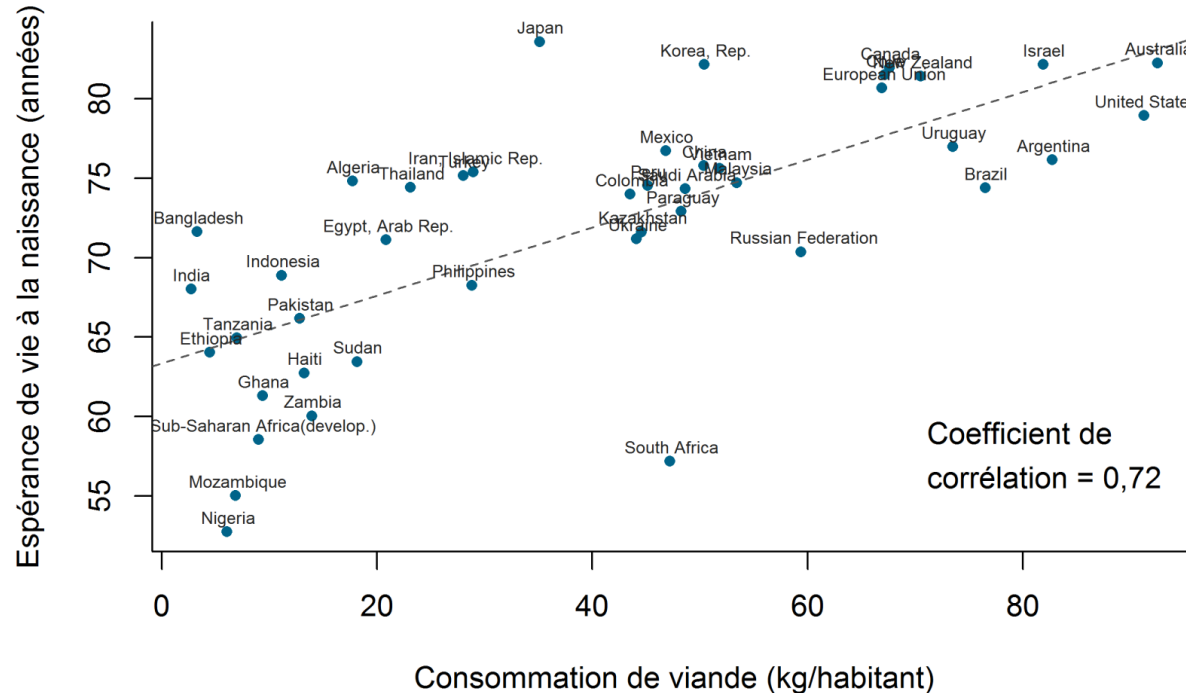


Figure 3 : Espérance de vie à la naissance et consommation de viande en 2014 dans certains pays du monde. Sources : OECD-FAO Agricultural Outlook (Edition 2015) et The World Bank, World Development Indicators.

<https://soepidemio.com/2016/11/14/correlation-statistique-prudence-a->





Corrélation et dépendances

- Taille et poids ?
 - Quanti*Quanti => R2
- Couleur des yeux et couleurs des cheveux ?
 - Quali*Quali => Chi2
- Couleur des cheveux et taille ?
 - Quali*Quanti =>
 - recoder le quanti en quali (classe de taille)
 - Faire des tests, comparaison de moyennes

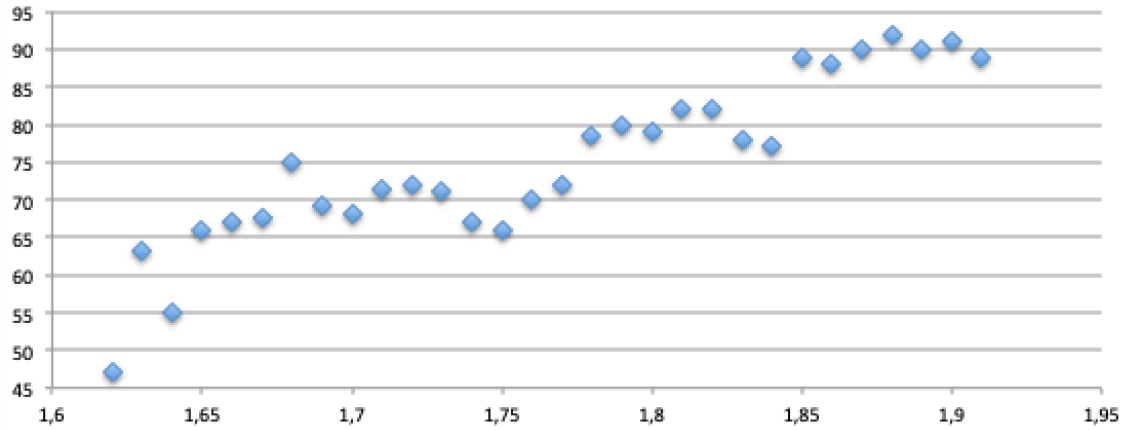


Corrélation et Dépendance

- Corrélation : Variables quantitatives/ordinales
 - Graphes XY
 - R^2 , R de Pearson \Rightarrow variables normales
 - Coefficient de Spearman \Rightarrow non normalité, indicateurs non paramétriques
 - Exemple : Taille et poids
- Dépendance : Variables Qualitatives/nominales
 - Tableau de contingence
 - Calcul du χ^2
 - Exemple : Couleur des yeux et couleur des cheveux



Corrélation Taille Poids



H0 : R=0

H1 : R≠0

R=0,91 - p value <0,05

	blonds	bruns	noirs	roux	Total
bleus	25	9	3	7	44
gris	13	17	10	7	47
marrons	7	13	8	5	33
Total	45	39	21	19	124

Répartition couleur cheveux/yeux

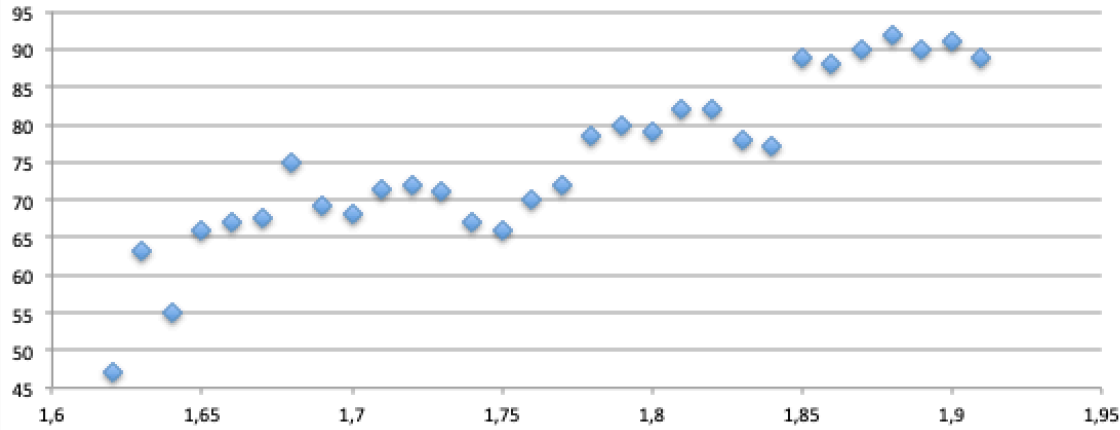
H0 : Variable A est indépendante de la variable B

H1 : Variable A est dépendante de la variable 2

Chi2 =15,1, P value < 0,01



Corrélation Taille Poids



R=0,91 - P value <0,05

H0 : R est égal à 0

H1 : R est significativement différent de 0

	blonds	bruns	noirs	roux	Total
bleus	25	9	3	7	44
gris	13	17	10	7	47
marrons	7	13	8	5	33
Total	45	39	21	19	124

Répartition couleur cheveux/yeux

Chi2 =15,1, P value < 0,01

H0 : la couleur des yeux est indépendante de la couleur des cheveux

H1 : H0 : la couleur des yeux est dépendante de la couleur des cheveux



<https://statistics.ead-minerve.fr>



Matrice de corrélations

<http://www.sthda.com/french/wiki/matrice-de-correlation-avec-r-analyse-et-visualisation>

```
library(Hmisc)
rcorr(as.matrix(mtcars[,1:7]))
```

```
      mpg   cyl  disp    hp  drat    wt  qsec
mpg   1.00 -0.85 -0.85 -0.78  0.68 -0.87  0.42
cyl  -0.85  1.00  0.90  0.83 -0.70  0.78 -0.59
disp -0.85  0.90  1.00  0.79 -0.71  0.89 -0.43
hp   -0.78  0.83  0.79  1.00 -0.45  0.66 -0.71
drat  0.68 -0.70 -0.71 -0.45  1.00 -0.71  0.09
wt   -0.87  0.78  0.89  0.66 -0.71  1.00 -0.17
qsec  0.42 -0.59 -0.43 -0.71  0.09 -0.17  1.00
```

n= 32

P

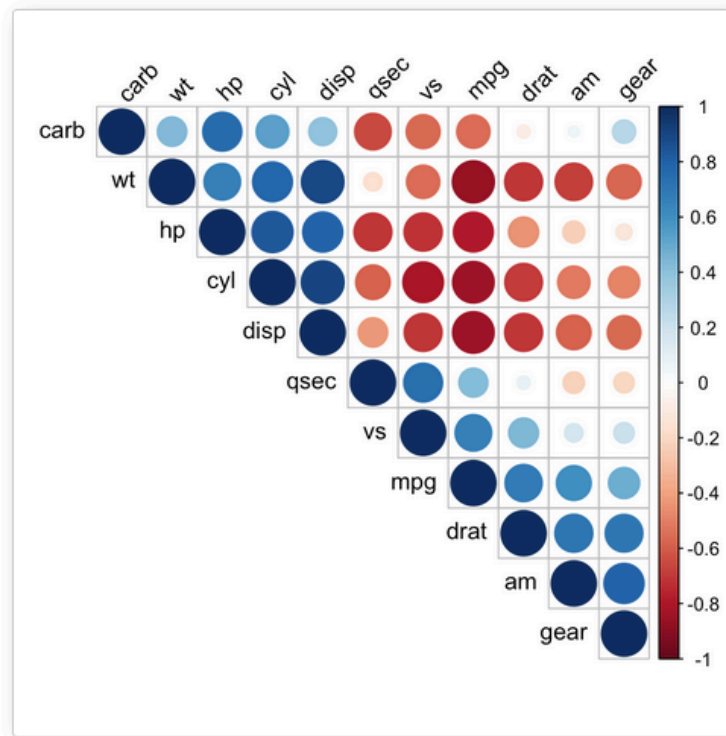
```
      mpg   cyl   disp   hp   drat   wt   qsec
mpg   0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0171
cyl  0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0004
disp 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0131
hp   0.0000 0.0000 0.0000 0.0000 0.0100 0.0000 0.0000
drat 0.0000 0.0000 0.0000 0.0100 0.0000 0.0000 0.6196
wt   0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.3389
qsec 0.0171 0.0004 0.0131 0.0000 0.6196 0.3389
```



Matrice de corrélations

<http://www.sthda.com/french/wiki/matrice-de-correlation-avec-r-analyse-et-visualisation>

```
library(corrplot)  
corrplot(mcor, type="upper", order="hclust", tl.col="black", tl.srt=45)
```



✓ Les **corrélations positives** sont affichées en bleu et les **corrélations négatives** en rouge. L'intensité de la couleur et la taille des cercles sont proportionnelles aux **coefficients de corrélation**. A droite du **corrélogramme**, la légende de couleurs montre les **coefficients de corrélation** et les couleurs correspondantes.



Tableaux des fréquences

<i>fréquence</i>	blond	bruns	noir	roux	Total
bleu	25	9	3	7	44
gris	13	17	10	7	47
marron	7	13	8	5	33
Total	45	39	21	19	124

<i>%total</i>	blond	bruns	noir	roux	Total
bleu	20,2	7,3	2,4	5,6	35,5
gris	10,5	13,7	8,1	5,6	37,9
marron	5,6	10,5	6,5	4,0	26,6
Total	36,3	31,5	16,9	15,3	100,0

<i>%ligne</i>	blond	bruns	noir	roux	Total
bleu	56,8	20,5	6,8	15,9	100,0
gris	27,7	36,2	21,3	14,9	100,0
marron	21,2	39,4	24,2	15,2	100,0
Total	36,3	31,5	16,9	15,3	100,0