

# Faut-il avoir peur de l'évaluation?

*Charles Hadji*

*Professeur émérite*

*Univ. Grenoble Alpes*

*Laboratoire des Sciences de l'Education*

4<sup>ème</sup> Colloque ADIUT

Lyon, vendredi 22 mars 2013

# AVANT PROPOS

1. Le vent de l'évaluation souffle fort et partout: une activité sociale omniprésente, à laquelle rien ne semble pouvoir échapper
2. Une activité de base dans l'agir humain sous 3 formes (cf. J.M. Barbier, 1985):
  - Implicite (ni instrumentée, ni formulée)
  - Spontanée (formulée, sans instrumentation)
  - Instituée (explicite: formulée et instrumentée)
3. Mais une activité sujette à toutes les dérives, et qui peut même rendre fou.

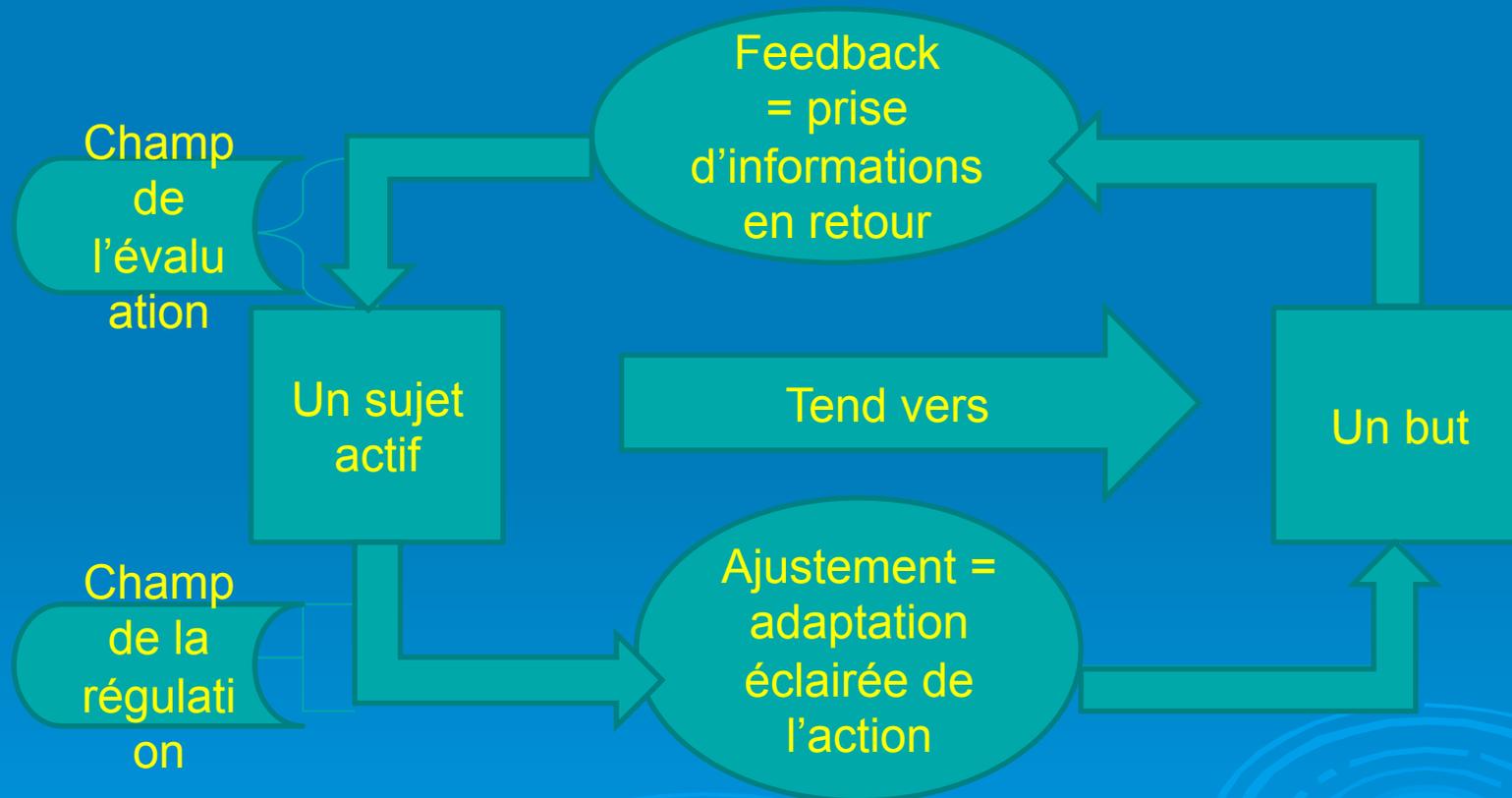
# Plan de l'exposé

1. Evaluer, qu'est-ce à dire?
  2. Une pratique à la recherche de sa légitimité méthodologique
  3. Une pratique à la recherche de sa légitimité sociale et éthique
- 

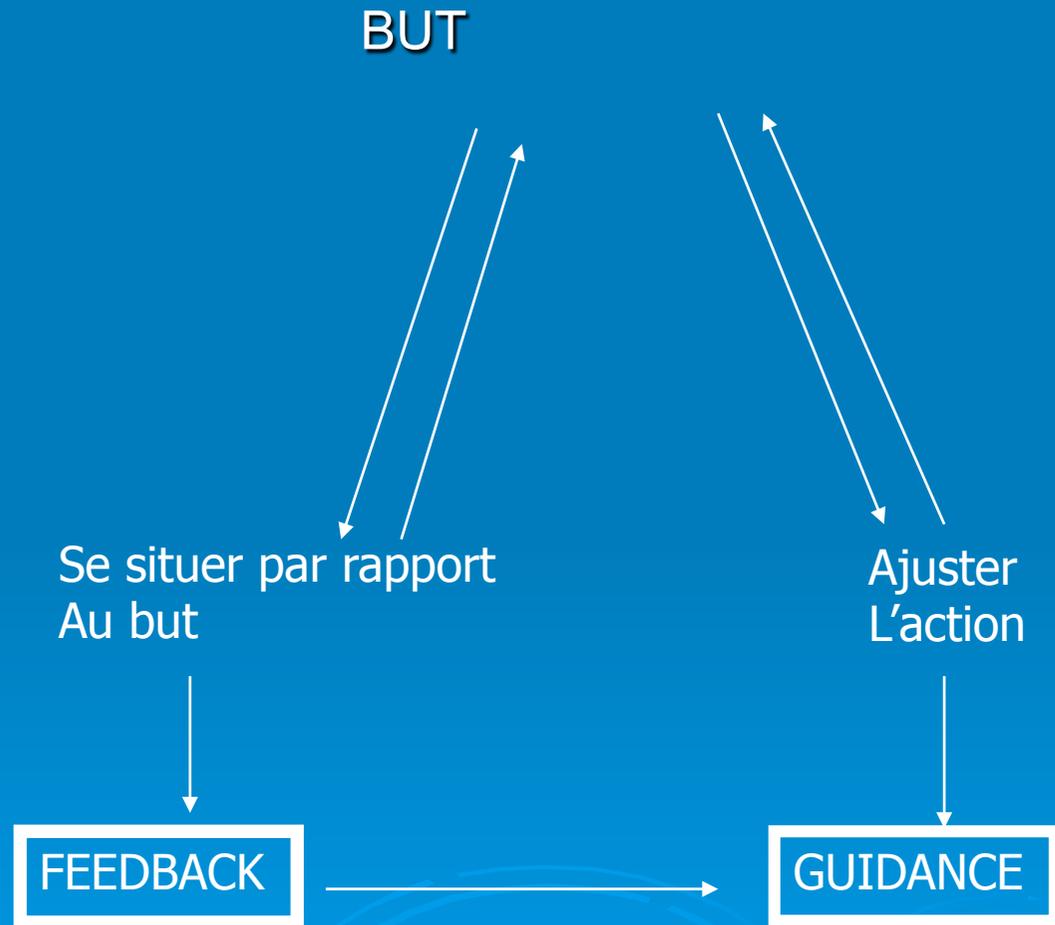
# Point 1: Evaluer, qu'est-ce à dire?

1. Une activité fondamentalement utile
  2. ... Qui produit un jugement d'acceptabilité
  3. ... Mais en étant écartelée entre estimation et appréciation
- 

# 11 Une pratique qui participe à la conduite éclairée de l'action



# L'activité de régulation: le modèle feedback-→guidance



*(D'après Linda Allall, 1988)*

# 12 Une pratique qui produit un jugement d'« acceptabilité »...

- Ex: dans le champ éducatif, évaluer c'est apprécier la qualité d'une action éducative, dans l'axe d'une question générale:
- *Quel est le « service » (éducatif, formatif, développemental) rendu?*

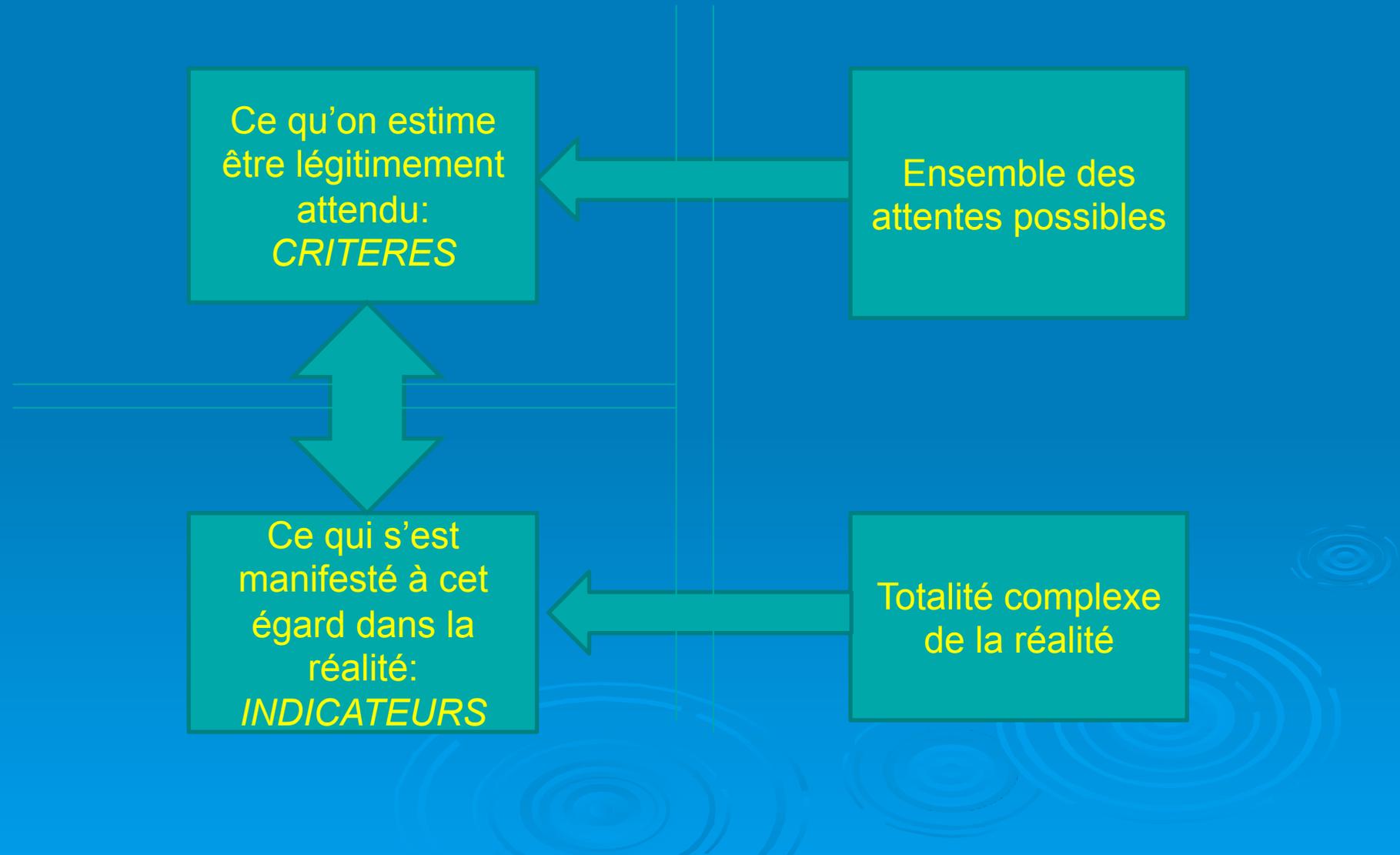
# ...par une mise en rapport SR/ SA

- Ce service rendu (SR) ne peut être apprécié que par rapport à un service attendu (SA)
- → évaluer une action éducative ou formative revient donc à comparer SR et SA, i.e. à apprécier le service qu'elle a rendu par référence au service qu'on en attendait

# Pour produire un jugement évaluatif, il faut donc savoir:

- Ce que l'estime être légitimement en droit d'attendre de la réalité évaluée (*l'attendu*)
  - définition nécessaire des modèles, cibles ou objectifs (qui permettront de spécifier des **CRITERES**)
- Dans quelle mesure la réalité évaluée (*le manifeste*) est en adéquation avec ce que l'on attend d'elle
  - identification nécessaire d'éléments significatifs (qui joueront le rôle d'**INDICATEURS**)

# La double articulation dans le jugement d'acceptabilité



# 13 Mais une pratique écartelée entre 2 projets

- **Projet 1: estimer l'importance (l'ampleur) d'un phénomène**
  - → **évaluation estimative**, « par défaut de mesure »: centrée sur l'appréciation de la grandeur d'un phénomène, et n'ayant pas pour fin la production d'un jugement de valeur
- **Projet 2: débattre de la valeur d'une réalité**
  - → **évaluation appréciative**: centrée sur l'appréciation de la qualité d'une réalité, et ayant pour fin la production d'un jugement de valeur.

# Exemples

## ➤ 1: évaluation estimative:

- -évaluer les dégâts occasionnés par le tremblement de terre
- -évaluer la menace de paludisme
- -évaluer la dette du pays
- -évaluer le prix d'un appartement
- -évaluer les inégalités socio-économiques
- -évaluer le chômage, la pauvreté, la fraude fiscale....

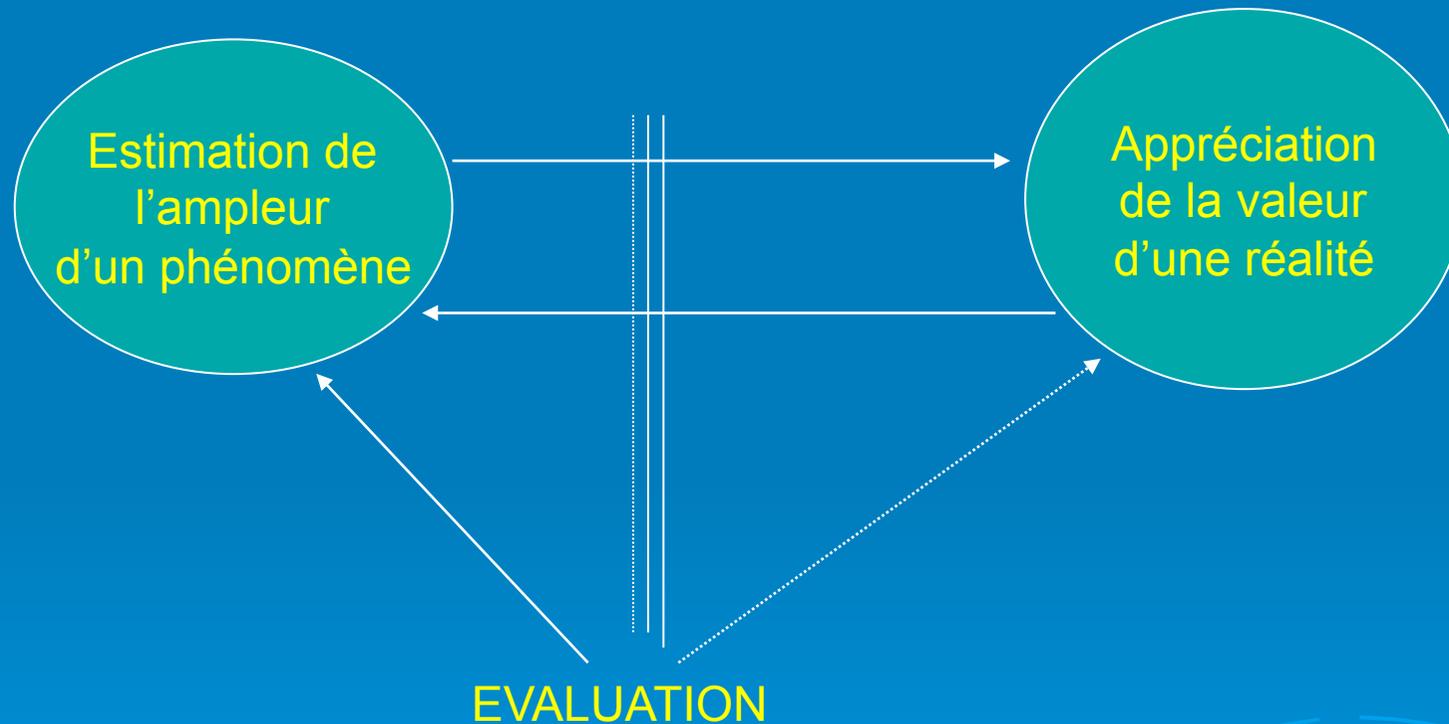
## ➤ 2: évaluation appréciative:

- -évaluer la « performance » des dirigeants d'entreprise
- -évaluer la « performance » d'un service
- -évaluer la politique de lutte contre la pauvreté
- -évaluer la qualité de l'enseignement universitaire
- -évaluer un ministre

# Les 2 pôles de l'activité évaluative

	Évaluation estimative	Évaluation appréciative
caractéristiques	Constat (de fait)  Quantitatif  Continuité des chiffres	Jugement (de valeur)  Qualitatif  Discontinuité des seuils
Exigences spécifiques	Précision  Fiabilité  Justesse	Pertinence    Justice
Risques	Valorisation subreptice	Insuffisante prise en compte du réel

# 2 projets dont aucun ne peut ignorer totalement l'autre



# EN CONSEQUENCE:

1. Il y a toujours une part d'estimation, voire de mesure, dans l'évaluation appréciative: le jugement se fonde sur un constat qu'il intègre
2. L'estimation court toujours le risque de la valorisation subreptice et induite :présence de valeurs sous-jacentes à la mise en place d'un indicateur (ex: nombre d'actes, taux de mortalité pour un hôpital); choix orienté des seuils et des limites de catégories (ex: le seuil de pauvreté).
3. De par son étymologie, le terme d'évaluation devrait être réservé à l'évaluation appréciative

## Point 2: Une pratique à la recherche de sa légitimité méthodologique, car:

- 21: risque de déni de la dimension de jugement
- 22 : risque d'oubli du questionnement porteur
- 23 : risque de soumission à la religion du chiffre

## 21 Le risque de déni du jugement: l'évaluation est une lecture appréciative, qui doit s'assumer comme telle

- « Il n'y aura évaluation qu'au moment ... où émerge le qualitatif dans le quantitatif. En ce sens, il n'existe d'évaluation que qualitative...
- La chaîne quantitative n'est jamais qu'un descriptif organisé de la réalité et ... l'évaluation consiste justement à briser la continuité de cette chaîne »

(Jacques Ardoino, Guy Berger, 1986)

# le processus dans sa dynamique

- **1: analyse des attentes** (voilà à quoi doit servir l'institution, Voilà ce que doit devenir le formé)
- **2: spécification des critères** (voilà quels sont les effets escomptés, domaine par domaine)
- **3: observation analytique du réel** (voilà les informations que j'ai pu produire pour mon évaluation)
- **4: identification et saisie de « signes » parlants** (voilà à quoi je vois que mes attentes sont satisfaites ou non)
- **5: formulation d'un « jugement d'acceptabilité »** (voilà ce que je pense de la qualité de...)

# En condensé: 3 temps forts, autour de 3 questions

- **Question 1:** qu'est-on essentiellement en droit d'attendre de l'« objet » évalué? Quel est le service attendu (SA)?  
→ assigner des **attentes prioritaires**
- **Question 2:** que faudra-t-il alors aller voir (observer)?  
→ identifier des **espaces d'observations**, i.e. des « lignes de lecture » de l'objet
- **Question 3:** et à quoi verra-t-on que l'attente est satisfaite ou non?  
→ définir des **indicateurs de réussite**, i.e. des signes d'appréciation du service rendu (SR) par la réalité évaluée

## 22 : Le risque de l'oubli du questionnement porteur: évaluer, c'est (tenter de) répondre à des questions

- Il paraît possible de distinguer à cet égard des « **projets de questionnement** » se situant à des niveaux différents (du plus fondamental au plus particulier)

# Questionnement de niveau 1: sur la réalisation de la mission fondamentale du système d'action évalué

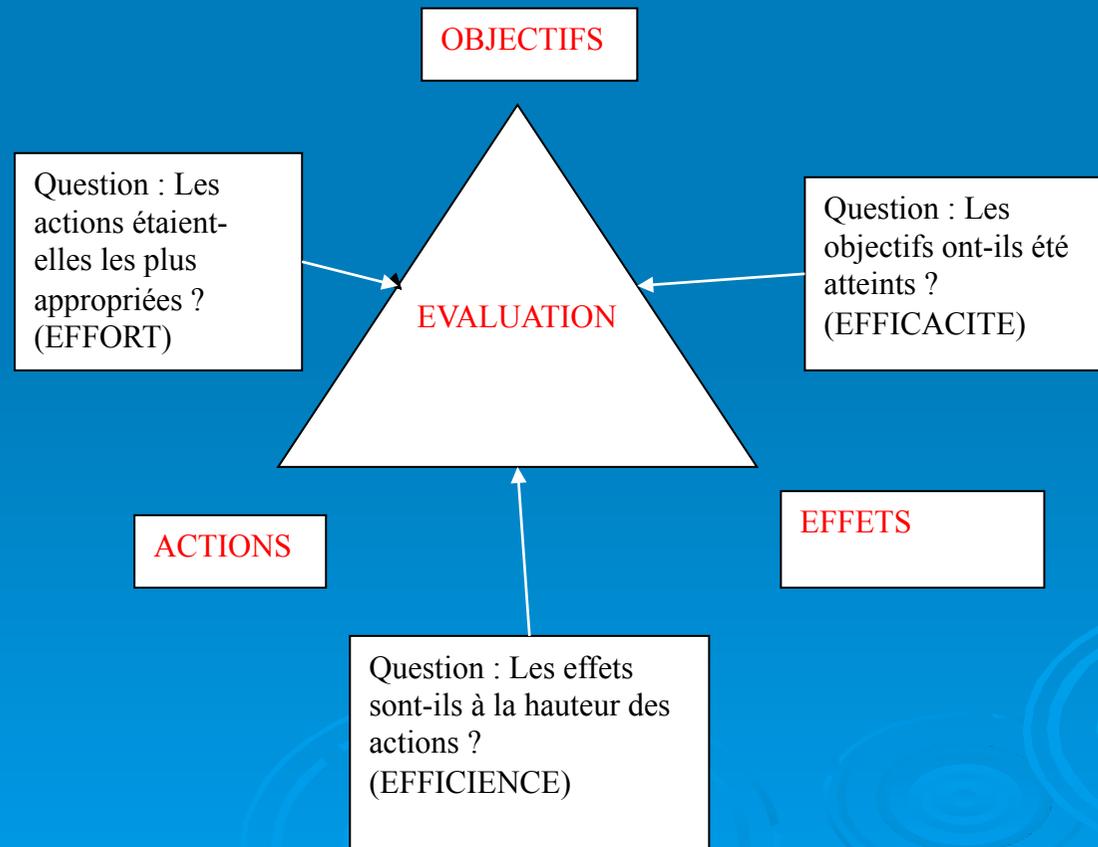
- Quelle est la valeur du système d'action comme solution aux « problèmes posés par l'action collective » (Crozier et Friedberg, 1977) qu'il était censé résoudre?
- Exemple:
  - L'école éduque-t-elle?
  - L'hôpital soigne-t-il?
  - La police sécurise-t-elle?
  - Le gouvernement permet-il au pays de progresser (d'aller mieux)?

# Questionnement de niveau 2: sur la qualité de l'action selon certaines dimensions fondamentales

- Le questionnement porte alors sur des qualités transversales exigibles de tout système d'action sociale. Exemple 1:
  1. **Pertinence** (adaptation aux objectifs)
  2. **Efficacité** (réalisation des objectifs)
  3. **Équité** ( juste « traitement » des situations, des individus ou des groupes).

# Deuxième exemple: le triangle évaluatif (effort, efficacité, efficacité)

LE TRIANGLE GENERAL DE L'EVALUATION DES ACTIONS



# Questionnement de niveau 3: sur la satisfaction d'attentes particulières

- Soit parce qu'elles sont « étroites » en tant qu'elles correspondent aux besoins d'un public restreint.  
Exemples:
  - satisfaction d'un besoin de reconnaissance (telle action sociale a-t-elle contribué à mieux faire connaître les élus?)
  - satisfaction d'un besoin local (les habitants de tel quartier ont-ils trouvé leur compte?)
- Soit parce qu'elles ne concernent qu'un objet particulier.  
Exemples:
  - M. X est-il un bon professeur?
  - que vaut ce ministre?

# 23 Le risque de soumission à la religion du chiffre: l'impérialisme du quantitatif

- On tombe trop souvent dans les facilités des « politiques du chiffre »
    - Exemple 1: les critères du classement de Shanghai
    - Exemple 2: les indicateurs bibliométriques
- (Alors que toute évaluation est « bipolaire »)

# Ex 1: les critères du classement de Shanghai (Université Jiao Tong, 2003)

➤ R1: les critères retenus répondent eux-mêmes à 2 critères:

- C1: être bien visibles
- C2: être immédiatement et facilement quantifiables

R2: si bien qu'ont été retenus:

1. des critères de notoriété, valorisant des indicateurs de réputation
2. En nombre limité → simplification excessive de l'objet
3. Ne prenant en compte que les « performances » dans le domaine de la recherche, critères qui alors:
  - figent le classement dans le temps, du fait du poids accordé aux « Prix Nobel »
  - sont trop sensibles à l'effet « taille »
  - favorisent les universités les plus riches
  - privilégient les publications en anglais

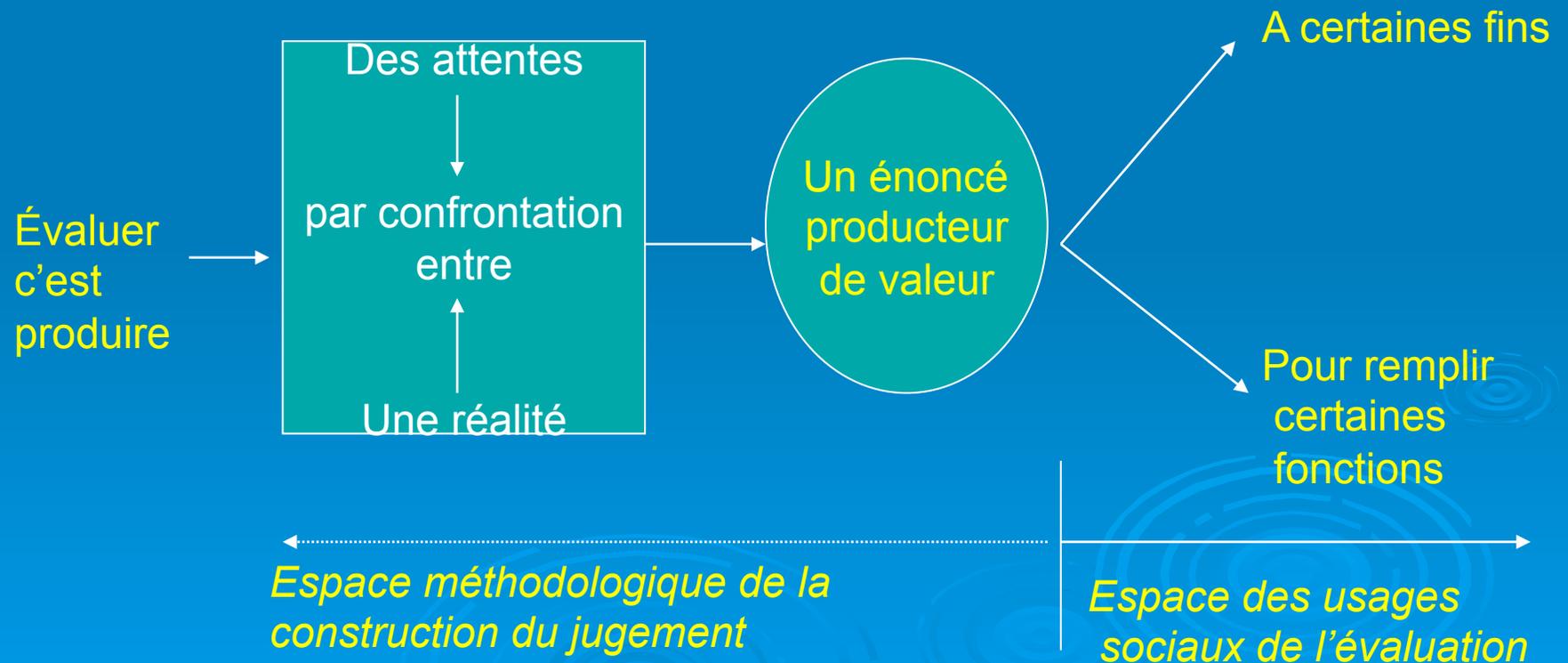
# La réduction méthodologique: l'exemple de la bibliométrie: biais, tentations et oublis

- Effet de résonance du thème
- Effet de variabilité de la base des données
- Effet d'impact propre au support de publication
- Effet d'appartenance à des réseaux
  - Tentation de « saucissonnage »
  - Tentation de citation d'amis
  - Tentation d'autocitation
    - Non prise en compte des cultures d'évaluation spécifiques
    - Non prise en compte de la qualité/attractivité du projet de laboratoire
    - Non prise en compte de l'intérêt et de l'importance réels des articles

### 3 Une pratique à la recherche de sa légitimité sociale et éthique

- 31 Qui doit lutter contre le **risque de réduction fonctionnelle**
- 32 Qui doit lutter contre le **risque d'asservissement idéologique**
- Qui doit lutter contre le **risque d'oubli des exigences éthiques**

# 31 Une pratique orientée par des valeurs, visant des fins, et remplissant des fonctions



# Quelles sont les fins les plus légitimes?

Plus légitimes?	Moins légitimes?
connaître	Mettre sous pression
éclairer	Procéder à un modelage idéologique
aider	Mettre en concurrence
accompagner	Servir une « autorité »

# Une réduction fonctionnelle: l'addiction aux classements

- Alors que plusieurs grandes fonctions sont possibles pour l'évaluation, l'obsession évaluative fait du classement (dans une évaluation « normative ») la fonction privilégiée

# Quelques grandes fonctions possibles pour l'évaluation scolaire

- Faciliter les apprentissages (= évaluation *formative*).
- Vérifier la maîtrise d'un savoir ou d'une compétence (contrôle des acquis = évaluation *sommative*).
- Attester la maîtrise d'un acquis (= évaluation *certificative*).
- Classer (= évaluation *normative*).
- Repérer des niveaux collectifs (= évaluation *normative de groupe*).
- Apprécier un écart à une cible (= évaluation *critériée*).
- Faire un diagnostic forces vs faiblesses (= évaluation *diagnostique*).

## 32 Le risque d'asservissement idéologique: un « cousinage » thématique avec l'idéologie dominante

- L'extension du domaine de l'évaluation s'est effectuée dans un contexte idéologique marqué, celui de l'idéologie néo-libérale actuellement dominante, où est méconnue, voire bafouée, l'exigence du respect de la dignité de la personne humaine, volontiers réduite à sa dimension de marchandise.

# Mais qu'est-ce qu'une idéologie?

- Le terme d'idéologie désigne un système explicatif, fait d'idées, de doctrines et de croyances propres à une époque, une société ou une classe particulières, et qui constitue une grille dont la fonction est de rendre intelligibles les phénomènes humains.
- Se présentant comme un corps de savoirs, l'idéologie est valorisée dans toute la société où elle domine, alors qu'elle n'exprime que des convictions utiles au groupe qui l'a forgée, et à la scientificité hautement discutabile.

**Les promesses de  
l'idéologie  
(un modèle idyllique)**

**L'envers du tableau  
(une réalité  
calamiteuse)**

Efficacité et résultats

Politique du chiffre

Saine concurrence

Élimination des faibles

Compétitivité

Enrichissement des plus  
cupides

Performances

Excès et dérives

Rentabilité

Inégalités et  
précarisations

Mérite

Bénédiction du fait

Excellence

Dévalorisation

Marché autorégulateur

Crises et krachs

# Une idéologie dominante où s'exprime une vision réductrice de la valeur

- 1) Réduction de la réalité humaine à sa dimension quantitative (je possède, donc je suis); substitution du désir de posséder au désir de comprendre.
- 2) Réduction de la régulation au contrôle social
- 3) Réduction du travail scolaire à la compétition concurrentielle, et de la réussite à la performance.

# Dans ce contexte, l'évaluation a tendance à devenir servile

1. **Asservissement à des croyances:** une évaluation soumise aux dogmes néo-libéraux, et aux prétendues lois de l'économie (sacralisation des notions d'excellence et de concurrence)
2. **Asservissement au court terme:** une évaluation soumise à la dictature de l'immédiat (tournée vers ce qui est le plus facile à estimer, et/car le plus visible à court terme)
3. **Asservissement au pouvoir:** une évaluation soumise aux politiques (tentation de la bienveillance à l'égard des puissants)
4. **Asservissement à l'argent :** une évaluation soumise à la finance (prégnance des logiques comptables).

# 33 Le risque d'oublier les exigences éthiques

- Evaluer, c'est prendre le risque de commettre des abus de pouvoir, et d'humilier autrui.
  - → nécessité d'une « éthique de l'agir évaluationnel »

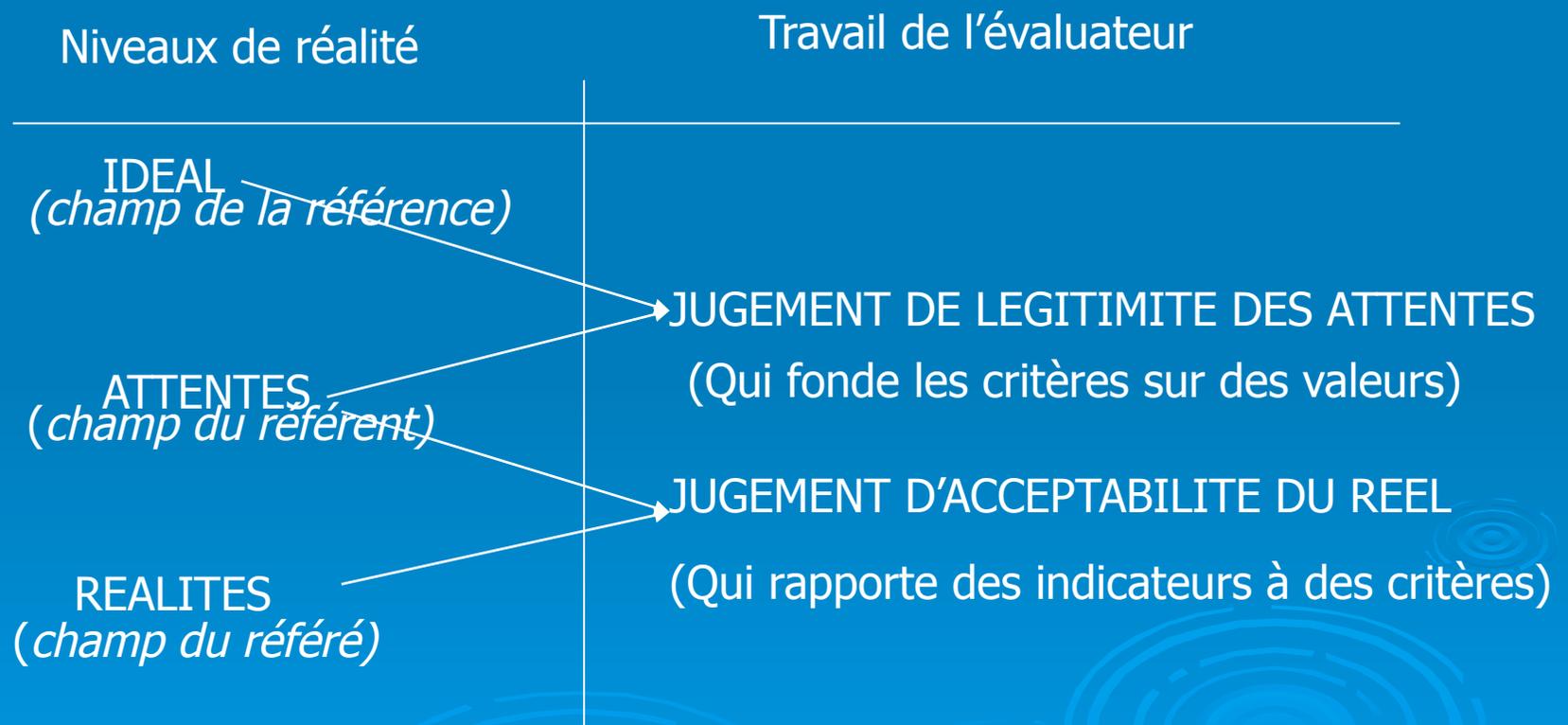
# Or l'évaluation est concernée par l'éthique de 2 grandes façons

- Premièrement:

De façon évidente, dans la détermination des critères, qui disent ce qu'on est **LEGITIMEMENT** en droit d'attendre de...

- Comment donc savoir quelles sont les attentes légitimes?

# Le double jugement dans l'activité d'évaluation



# L'évaluation est concernée par l'éthique, deuxièmement...

- Dans le rapport à l'autre que crée la relation évaluative:
- Le « rapport de communication évaluative » (RCE)

# La double logique du Rapport de Communication Evaluative (RCE)

- Comme tout rapport de communication (cf. Watzlawick, 1972) le RCE fonctionne selon une double logique:
  - \_ Une logique de **transmission d'informations** (*aspect contenu, ou indice*)
  - \_ Une logique de **communication sur la communication** (*aspect relation, ou ordre*)

# Or le RCE peut être sain ou « malade »

- « Plus une relation est spontanée et saine, et plus l'aspect « relation » de la communication passe à l'arrière-plan .
- Inversement, des relations « malades » se caractérisent par un débat incessant sur la nature de la relation, et le « contenu » de la communication finit par perdre toute importance » (Watzlawick, 1972, p. 50)

# La relation évaluative est « malade » quand l'évaluateur:

- passe son temps à affirmer à l'évalué: « voilà comment tu dois me voir: comme le chef, et le maître »
- A pour ambition dominante de maintenir l'évalué en position basse
- Cède à la tentation du plaisir (coupable) de dominer
- Succombe à l'abus de pouvoir

# La relation évaluative est saine quand l'évaluateur:

- Se situe dans une logique d'information (« observation formative », Perrenoud, 1998)
- Se situe dans une logique d'aide (« un retour d'information multidirectionnel, s'adressant à {l'intéressé}, au lieu de porter sur {l'intéressé} », Cardinet, 1989)
- Se situe dans une logique de régulation (mieux encore: d'autorégulation)

# Conclusion: l'évaluation n'est pas à craindre si elle respecte 3 conditions

- Condition 1: se délivrer d'une triple obsession:
  - l'obsession de la sélection → contre la « constante macabre » (Antibi, 2003)
  - - l'obsession de la compétition → contre la débauche de classements
  - - l'obsession de la notation → contre la « folie » des notes

## Condition 2: Privilégier en toutes circonstances l'informativité de l'évaluation

- Il s'agit de revenir à l'essentiel: informer est la fonction « naturelle » de l'évaluation
- → exemple: pour ne pas sombrer dans les querelles de la notation, mettre en place des outils de diagnostic personnalisé, tels que les « échelles descriptives » de Scallon, (2000, 2007)
- → le problème étant: informer qui?, ce qui soulève la question des publics prioritaires

# Vouloir une évaluation démocratique, c'est vouloir que ceux sur qui porte l'information en soient aussi les destinataires

- Une évaluation sera de ce point de vue d'autant plus démocratique qu'elle fera du **public-objet** un **public-destinataire**, en ayant pour fin d'éclairer ce public sur sa situation quant aux attentes qui appelaient légitimement une évaluation.
- Les résultats de l'évaluation lui seront alors directement utiles, puisqu'il pourra en tirer profit pour son action, dont ils rendront possible une meilleure régulation.

# Condition 3: Faire un choix éthique fondamental: servir l'homme, plutôt que le pouvoir ou l'argent

- La première façon d'être utile (en évaluation, comme ailleurs) est de ne pas nuire!
- → les 2 règles d'une éthique « minimale »:
- - 1 **Neminem laede** = ne fais de mal à personne
- -2 **Imo omnes, quantum potes, juva** = mais bien plutôt vient en aide à tous, autant que tu peux...