

# Pourquoi et comment évaluer ?

André Tricot

ESPE de l'Académie de Toulouse &

Laboratoire Cognition, Langues, Langage, Ergonomie

Équipe « Apprentissages, motivation, métacognition »

UMR 5263 CNRS, EPHE & Université Toulouse 2



« Most teachers view tests as  
necessary evils »

Roediger, Putnam & Smith (2011)

# Plan

1. Introduction
2. Perceptions et effets secondaires de l'évaluation
3. Les limites de l'évaluation
4. Les fonctions de l'évaluation
5. Conclusion

# Plan

1. Introduction
2. Perceptions et effets secondaires de l'évaluation
3. Les limites de l'évaluation
4. Les fonctions de l'évaluation
5. Conclusion

# Pourquoi évaluer ?

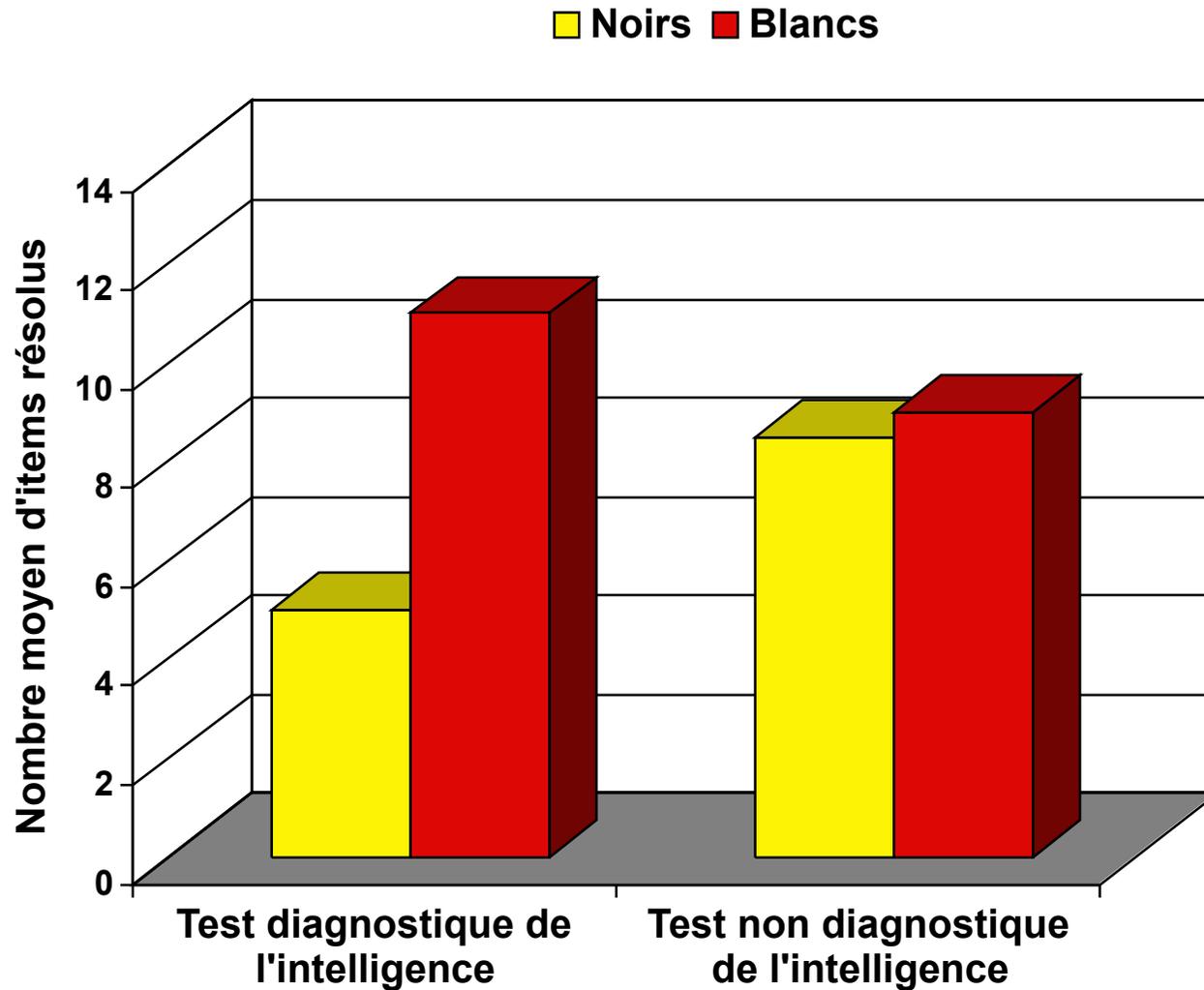
- Le pari rationnel
  - Étant donné un objectif à atteindre
  - Des moyens ont été mis en œuvre
  - Cet objectif a-t-il été atteint ?
  - Si non, comment améliorer les moyens ?

# Plan

1. Introduction
2. Perceptions et effets secondaires de l'évaluation
3. Les limites de l'évaluation
4. Les fonctions de l'évaluation
5. Conclusion

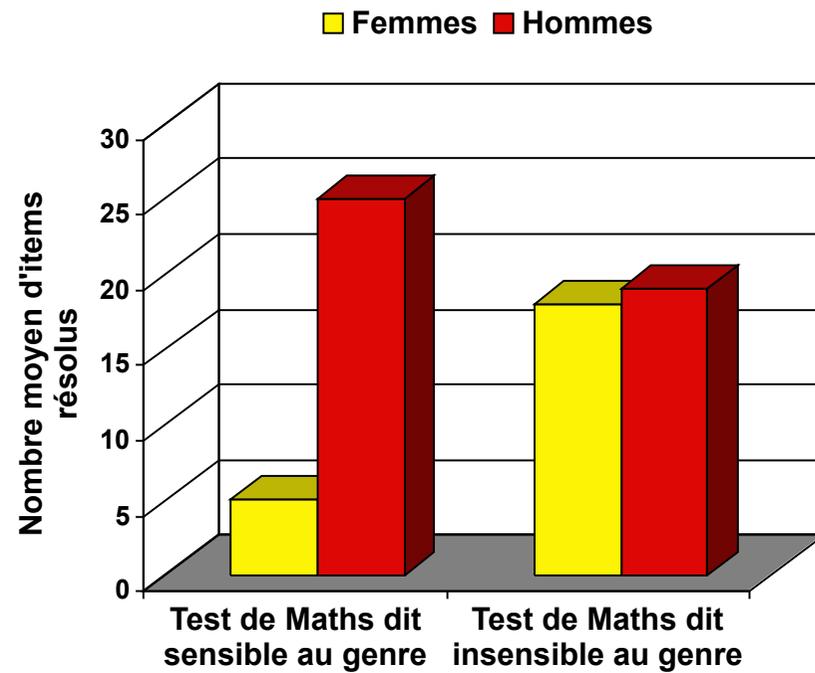
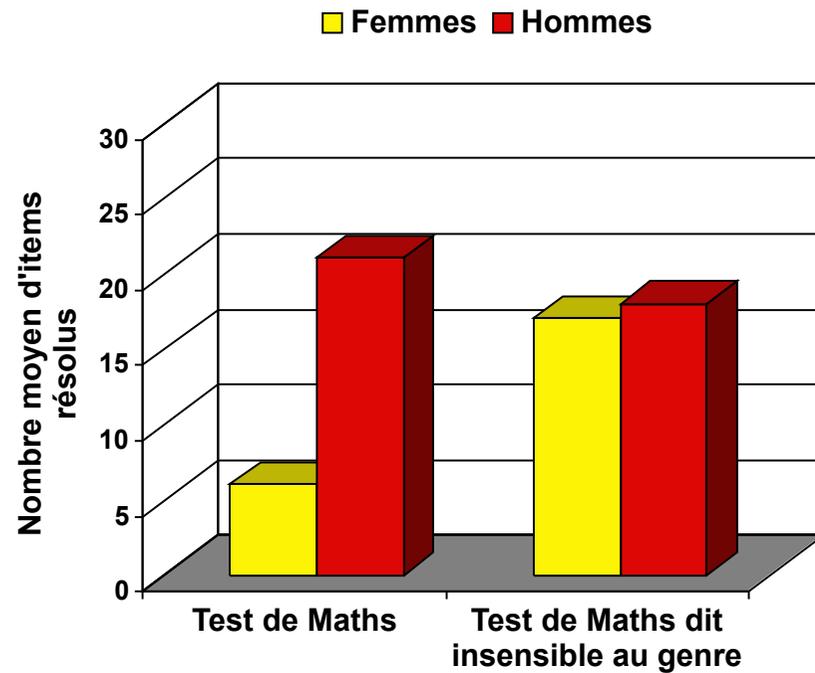
# La menace du stéréotype

(Steele & Aronson, 1995)

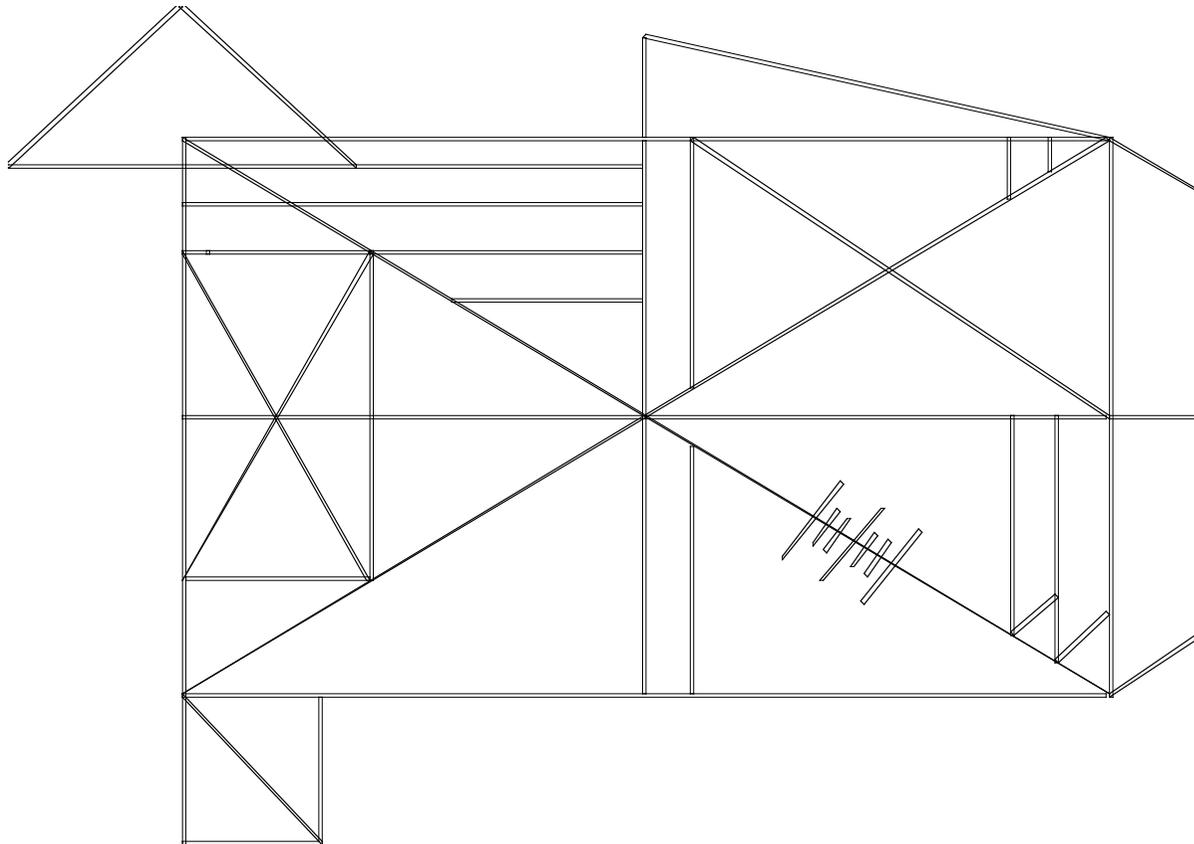


# Menace du stéréotype

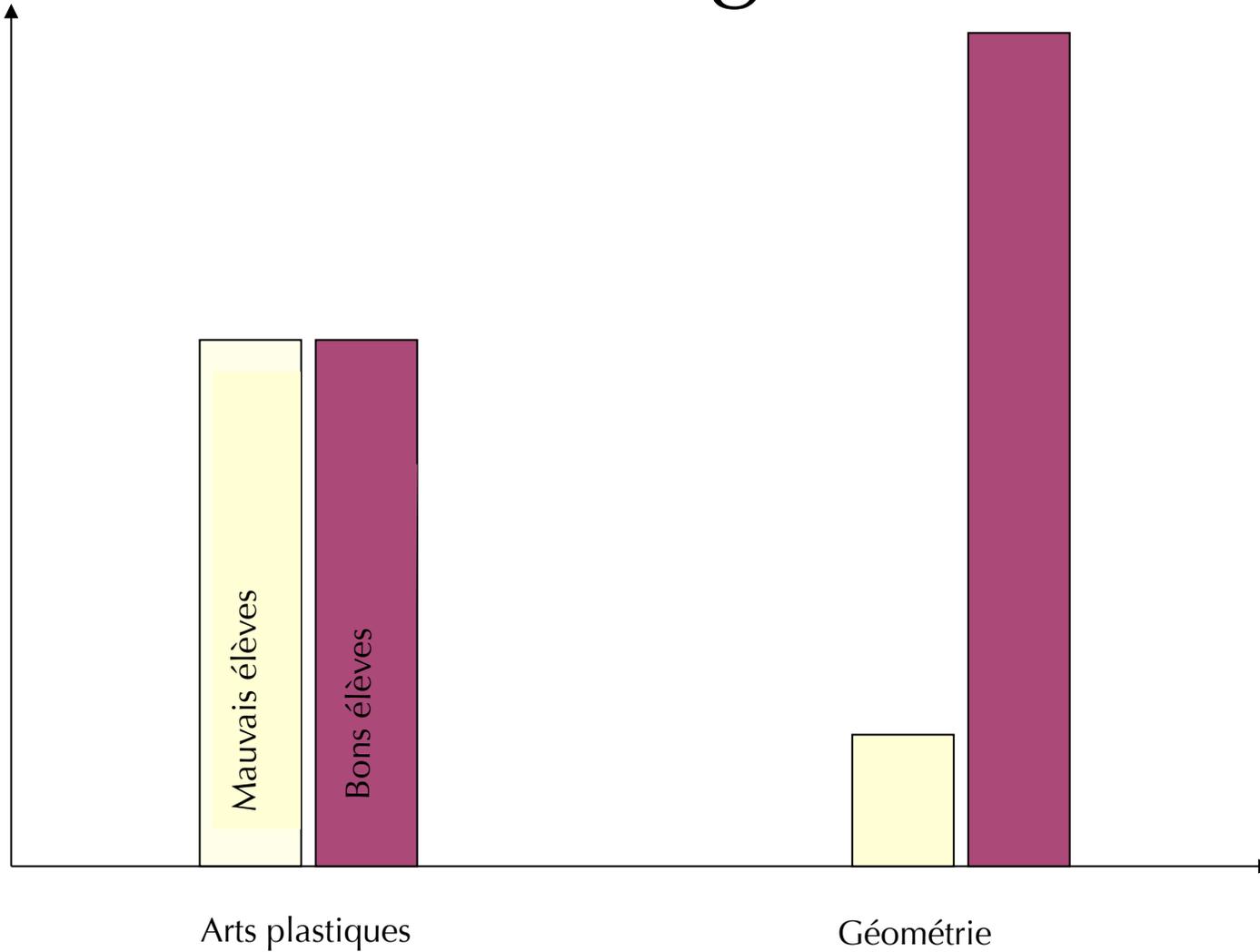
Spencer et al. (1999)



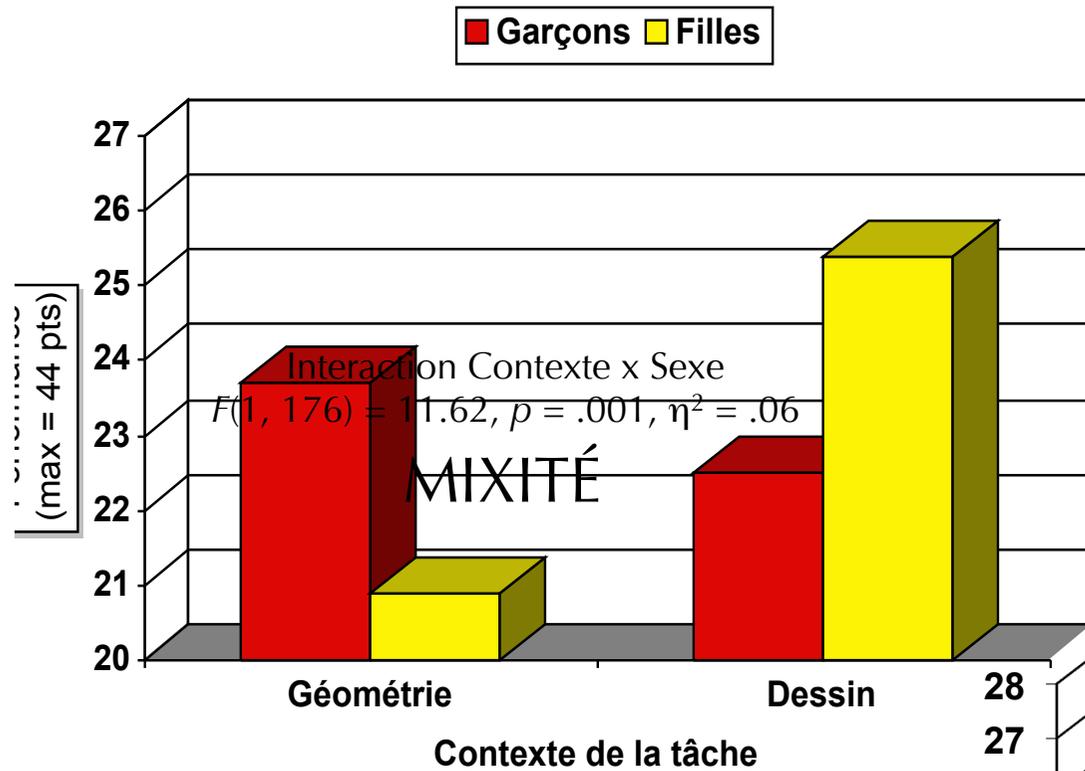
# Adaptée de la Figure Complexe de Rey-Osterrieth



# Monteil et Huguet (1991)

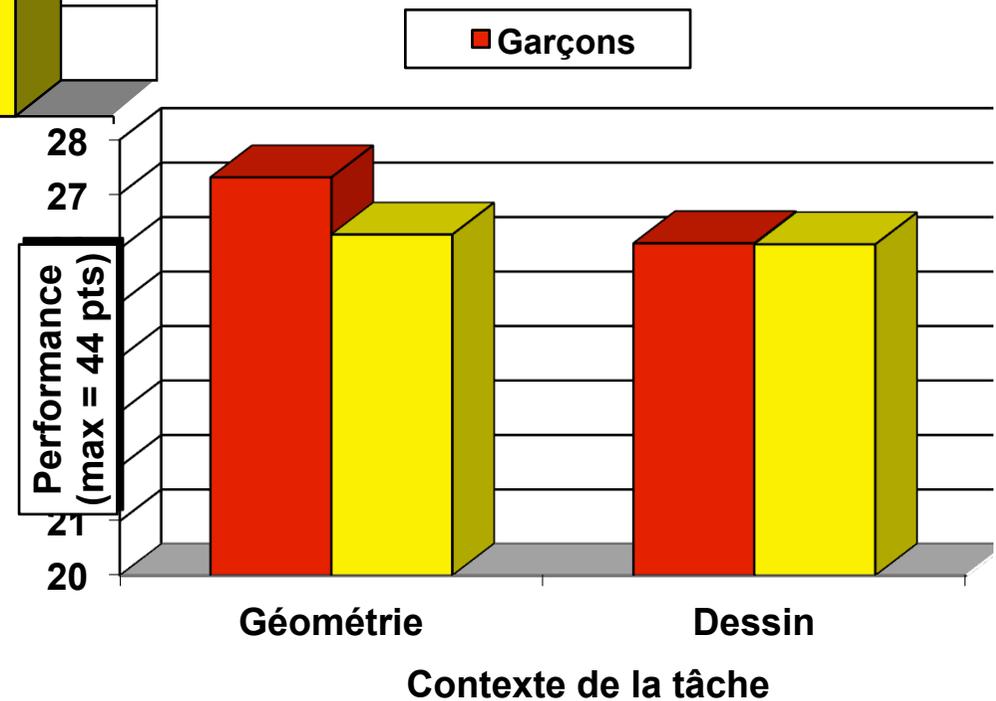


# Huguet & Régner (2007)



## NON-MIXITÉ

Interaction Contexte x Sexe  
 $F(1, 266) < 1$



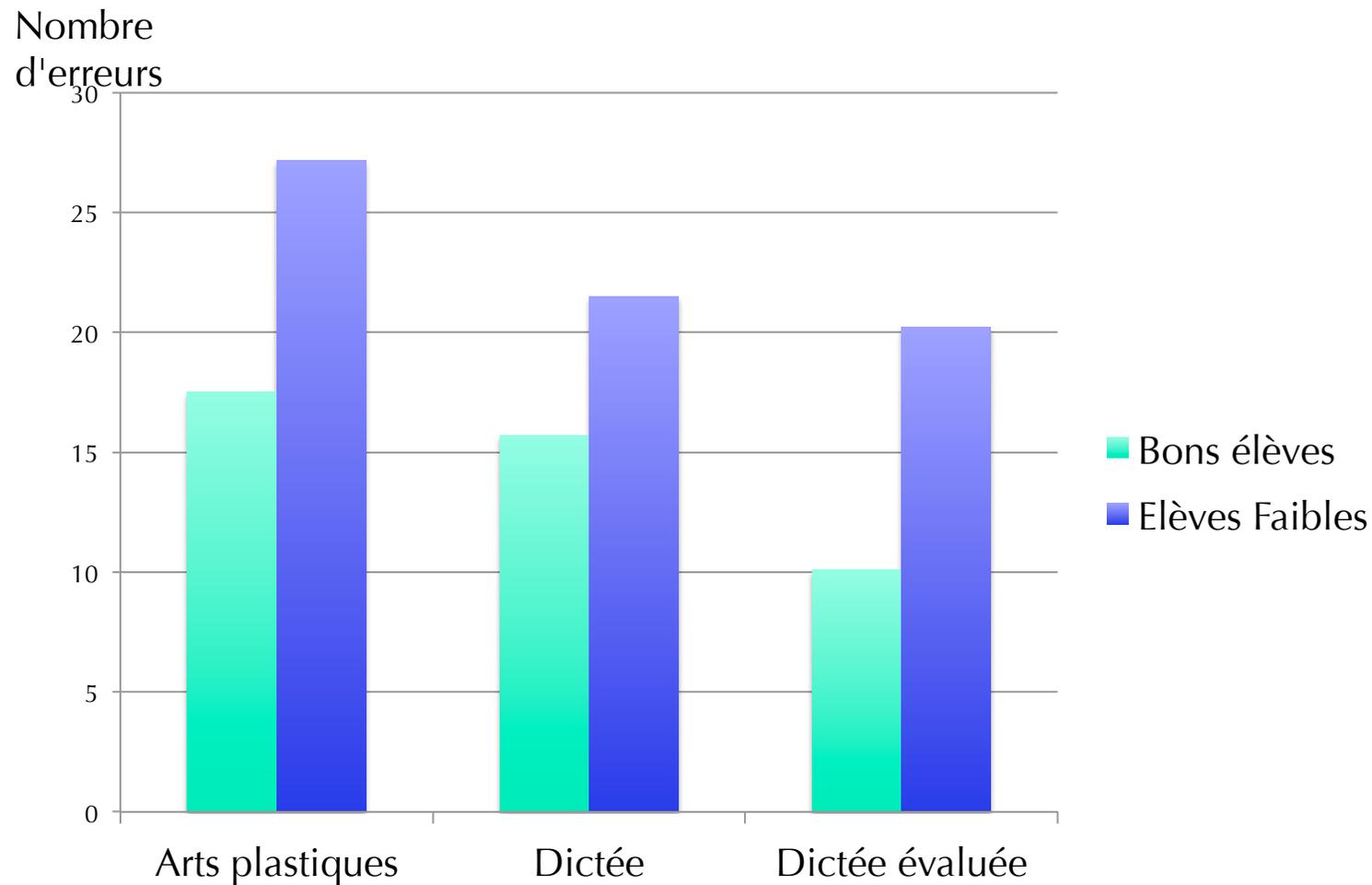
# *Le testing effect*

Roediger, Putnam & Smith (2011)

- L'évaluation permet d'identifier les lacunes en termes de connaissances,
- conduit les élèves à apprendre plus la fois suivante,
- améliore l'organisation des connaissances,
- améliore le transfert des connaissances,
- permet de mobiliser des connaissances qui n'ont pas été préalablement évaluées,
- améliore le contrôle métacognitif,
- prévient l'interférence avec les contenus préalables quand on aborde un nouveau contenu,
- fournit un retour aux enseignants,
- fréquente encourage les élèves à apprendre.

# Dictée notée ou non-notée ?

Toczek, Fayol & Dutrevis (2012)



# Plan

1. Introduction
2. Perceptions et effets secondaires de l'évaluation
3. Les limites de l'évaluation
4. Les fonctions de l'évaluation
5. Conclusion

- La place de la copie (premier tiers ; copie avant)
- L'effet Pygmalion, le niveau scolaire de l'élève (les élèves faibles sont sous-notés, les bons sur-notés)
- Age, origine sociale de l'élève, aspect physique (la bonne opinion sur l'élève augmente sa note)
- Sexe (filles +); Sexe / discipline (filles - dans disciplines scientifiques corrigées par des hommes)
- Age, expérience de l'enseignant
- Type d'établissement ; réputation de l'établissement d'origine (on surnote les élèves issus d'établissements réputés ; on sous note les autres)
- Discipline scolaire
- La précision des critères
- La constante macabre

# Que peut-on évaluer ?

- Les compétences ?
- Les connaissances ?
- Les activités ?
  - L'atteinte du but ?
  - La façon d'atteindre le but ?
- Des connaissances primaires ?

# La boule de pétanque et un cochonnet

- J'ai acheté une boule de pétanque et un cochonnet. J'ai payé le tout 1,10 €
- La boule coûte 1€ de plus que le cochonnet
- Combien coûte le cochonnet ?

# Un problème de carte

Johnson-Laird (2006)

- Un jeu de Black Jack

Si on tire un Roi, alors il y a un As juste après

*Ou bien*

Si on ne tire pas un Roi, alors il y a un As juste après

Je tire un Roi, qu'est-ce qu'il y a après ?

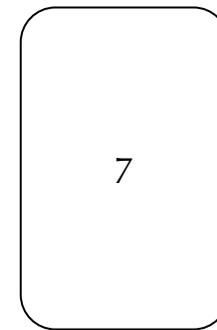
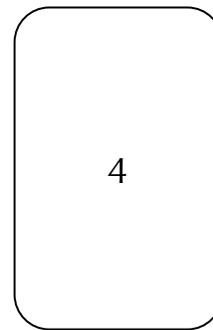
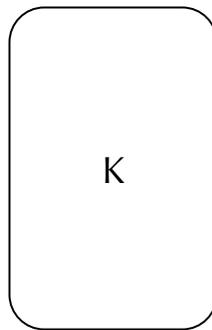
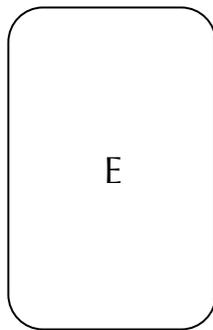
# Un problème de corde

- Soient deux points distants de 200 mètres l'un de l'autre. A chacun de ces deux points, on fixe une corde. Une fois fixée, la corde fait 200,04 mètres ; de sorte qu'elle est trop grande de 4 centimètres, elle flotte un peu. Maintenant, on va au milieu des deux points, à 100 mètres de chacun donc, et on soulève la corde pour qu'elle soit bien tendue. De combien peut-on soulever la corde ? à peu près 2 cm ? à peu près 20 cm ? à peu près 2 m ?

# Un autre problème de carte

Wason (1966)

Si il y a une voyelle sur une face il doit y avoir un nombre pair de l'autre



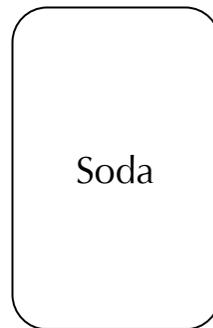
Quelle(s) carte(s) est-il nécessaire de retourner pour vérifier si la règle a été respectée?

- de 10% de bonnes réponses

# Un autre problème de carte

Wason (1966)

Pour boire de la bière dans les pubs il faut avoir plus de 18 ans



Quelle(s) situation(s) est-il nécessaire de vérifier pour voir si la règle a été respectée?

+ de 80% de bonnes réponses

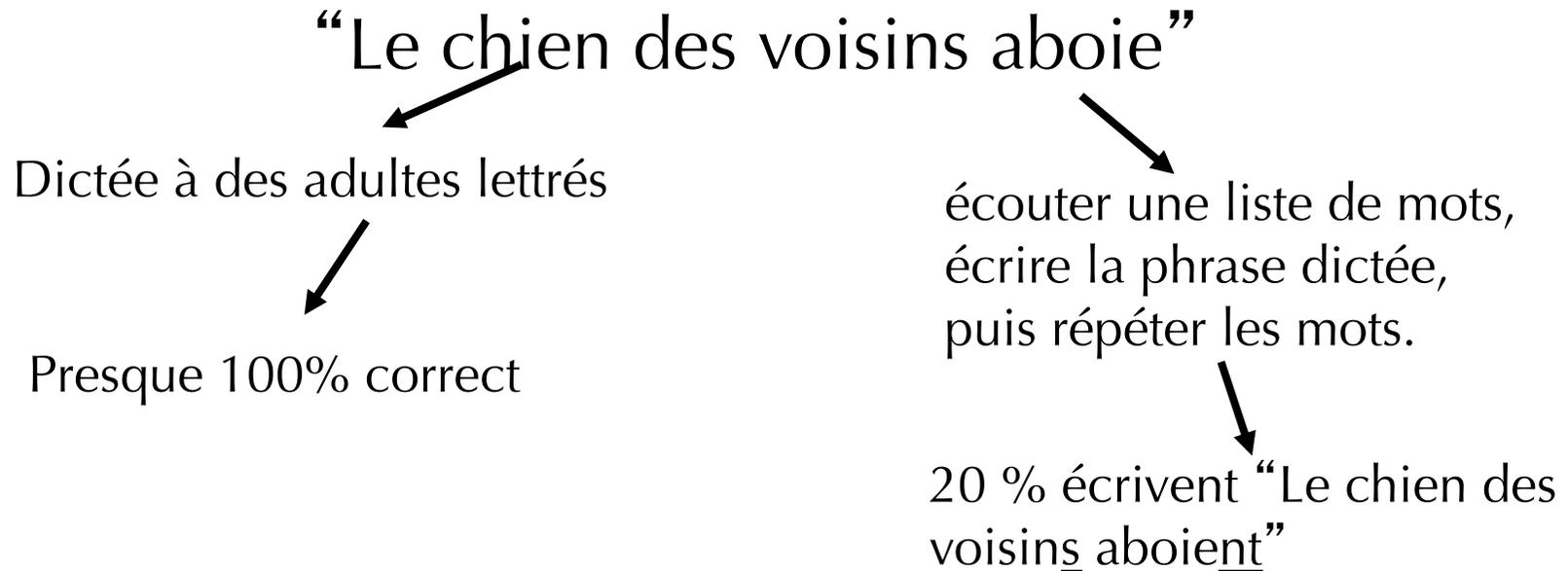
# Exemple dû à R. Brissiaud

- En début de CE1
  - Combien y a-t-il de gâteaux dans 3 paquets de 10 gâteaux ?
    - taux de réussite 48%
  - Combien y a-t-il de gâteaux dans 10 paquets de 3 gâteaux ?
    - taux de réussite 17%
- En fin de CE1
  - 73%
  - 53%

# Exemple dû à R. Brissiaud

- En fin de CE1
  - Dans sa tirelire Leila a 27 euros. Elle y ajoute d'autres euros et après elle a 31 euros. Combien a-t-elle ajouté d'euros ?
    - taux de réussite de 68%
  - Dans sa tirelire, Leila a 31 euros. Elle en sort 27 euros pour s'acheter un jouet. Combien lui reste-t-il d'euros dans sa tirelire ?
    - taux de réussite de 38%

# L'orthographe chez les étudiants (Largy, Fayol & Lemaire, 1996)



# 1ère conclusion

- Si un élève ne réussit pas une tâche qui implique la connaissance A, cela ne permet pas de conclure qu'il n'a pas la connaissance A
- Si un élève réussit une tâche qui implique la connaissance A, cela ne permet pas de conclure qu'il a la connaissance A
- Seule la multiplication des tâches et des modalités de présentation de la tâche permet une approximation du fait que l'élève maîtrise ou pas la connaissance A

# 1ère conclusion (suite)

- Deux élèves peuvent réussir (ou rater) la même tâche pour des raisons différentes
- Intérêt d'évaluer par les compétences, mais avec une définition stricte de la notion de compétence
- Le piège de la réponse attendue

# Les électroniciens

(Besnard et Bastien-Toniazzo, 1999)

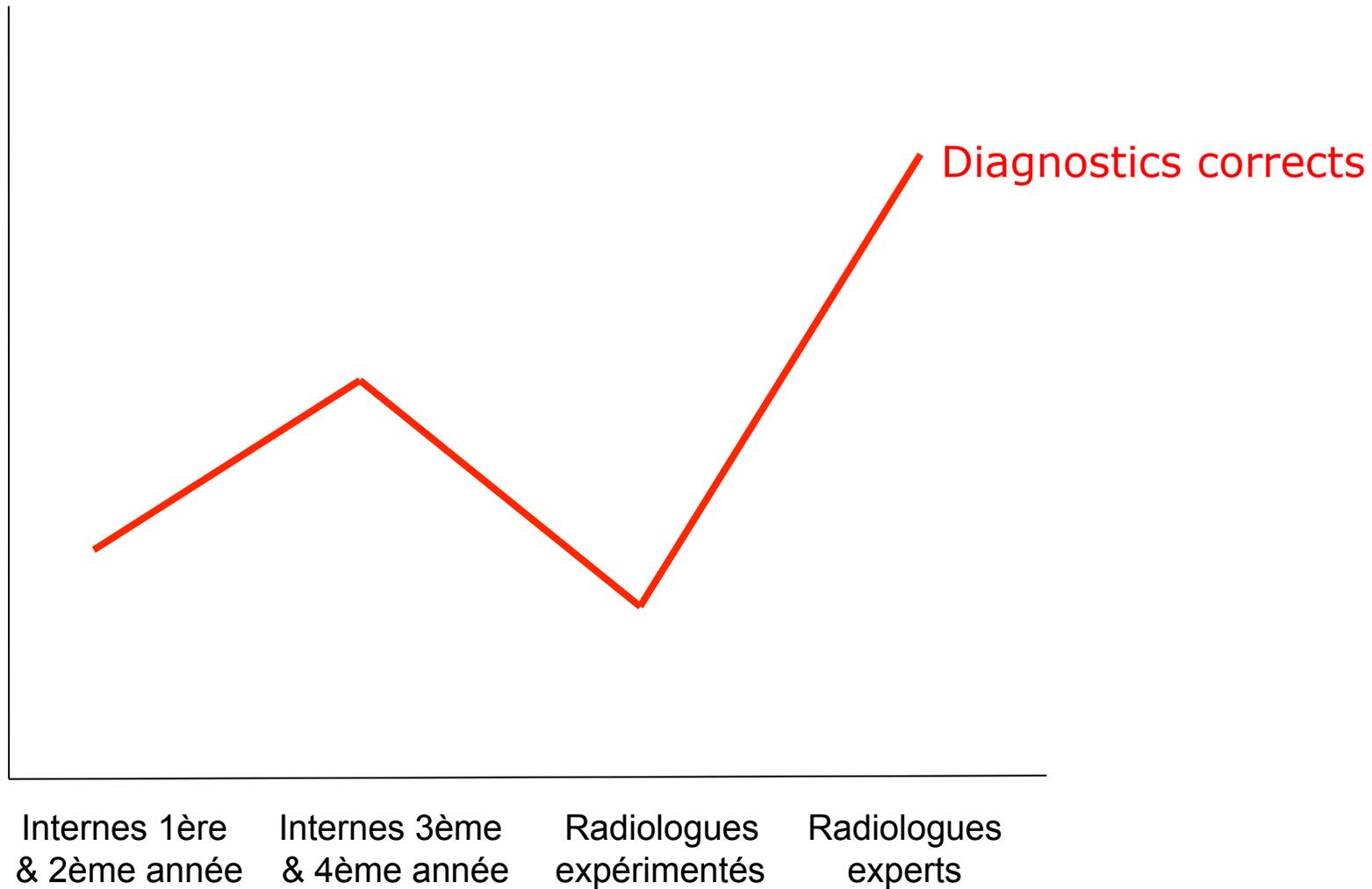
- Tâche de diagnostic de panne d'un dispositif d'alarme électronique à des électroniciens experts (professionnels, 20 ans d'expérience) et novices (étudiants, 2ème année d'IUT).
- Panne introduite artificiellement : court-circuit sur un condensateur placé sous le dispositif, donc invisible.
- Or il existe une panne extrêmement fréquente avec ce type d'alarme électronique : mettant en cause un circuit intégré.
- Les résultats montrent que
  - Les experts testent en moyenne 18 fois le circuit intégré qui n'est pas défaillant ;
  - Les novices 2 fois seulement.
  - Les experts éprouvent beaucoup plus de difficultés que les novices pour trouver la panne, parce qu'ils n'arrivent pas à prendre conscience que leur connaissance de la panne fréquente n'est pas la bonne.

# Les techniciens de maintenance automobile

- Même protocole : une panne artificielle mais plausible (obturation visible d'un échappement, juste à la sortie du cylindre)
- Même comparaison : experts (armée) et novices (étudiants)
- Même résultat : les novices sont meilleurs que les experts

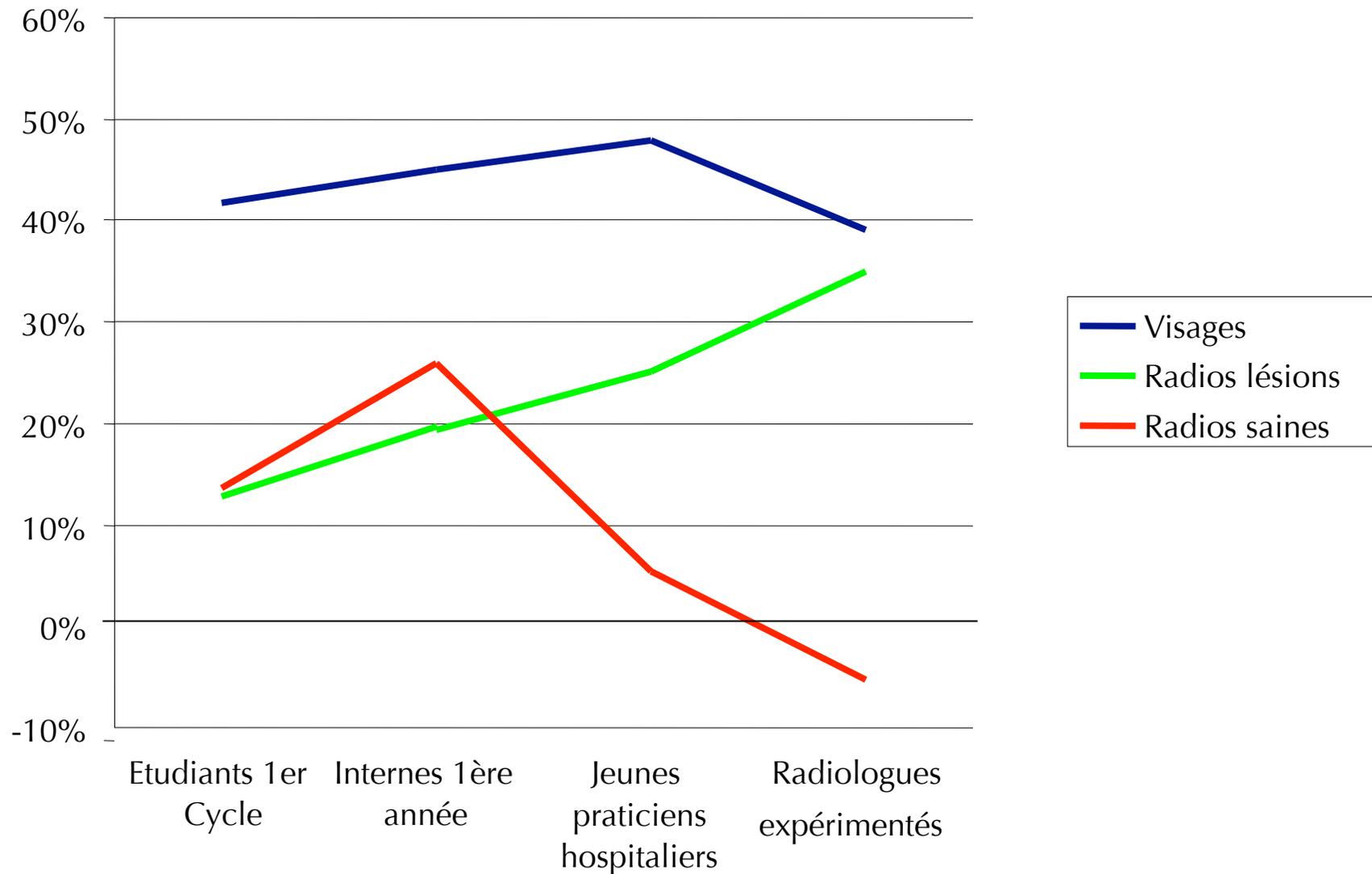
# Les radiologues

(Raufaste, Eyrolle & Mariné, 1998)



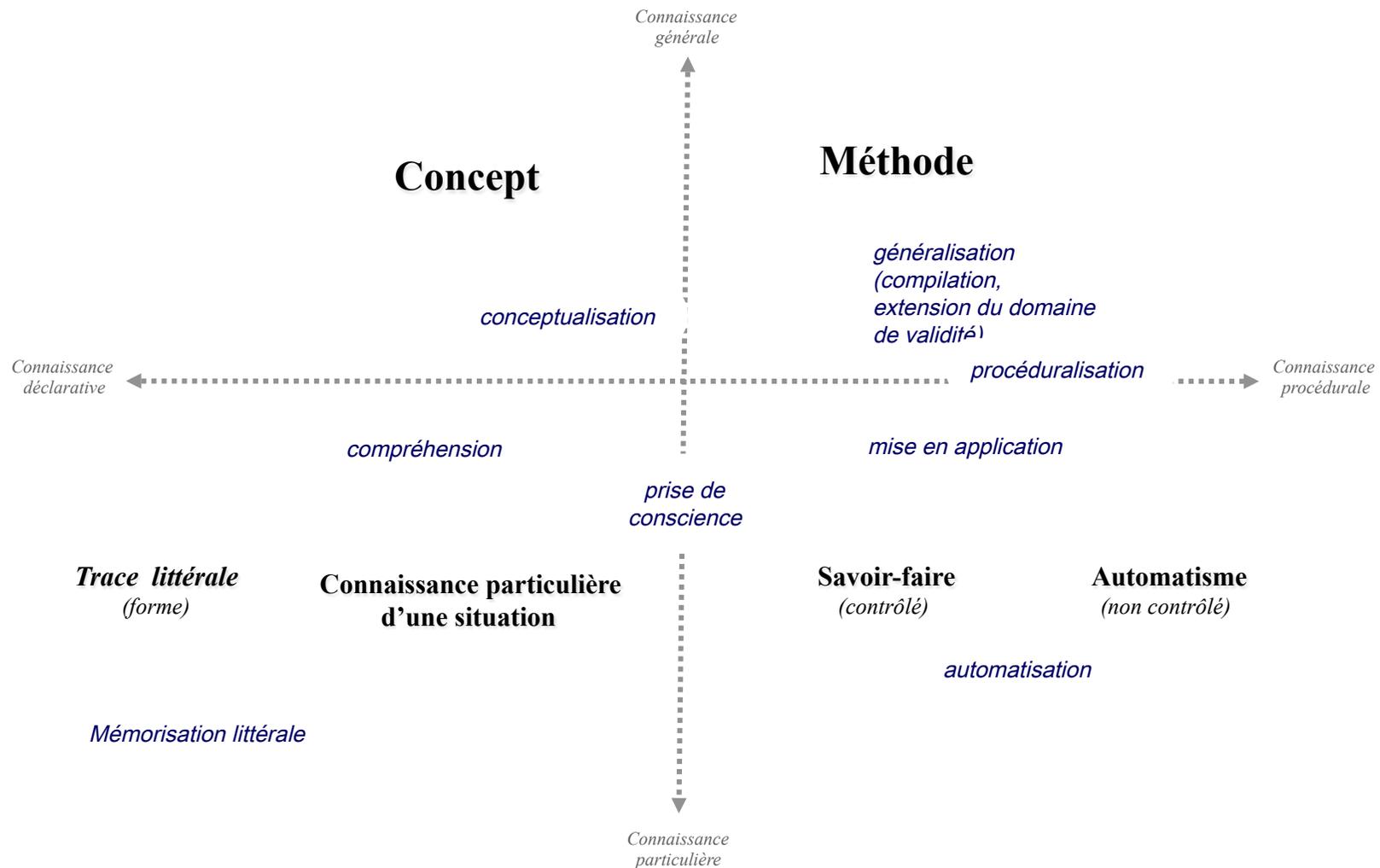
# Tâches de reconnaissance

(Myles-Worsley et al., 1988)



# Processus d'apprentissage et formats de connaissance

## Un outil pour concevoir des parcours d'apprentissage



(Musial, Pradère & Tricot, 2012)

# 2ème conclusion

- Importance de distinguer les formats de connaissances quand on évalue
  - Les savoir-faire automatiques (automatismes)
  - Les savoir-faire contrôlés (méthodes)
  - Les savoirs généraux (concepts)
  - Les savoirs locaux (représentations)
  - Les savoirs littéraux (mémoire)
- Attention aux tâches de transfert dans une évaluation !

# Plan

1. Introduction
2. Perceptions et effets secondaires de l'évaluation
3. Les limites de l'évaluation
4. Les fonctions de l'évaluation
5. Conclusion

# Evaluer = Situer un acte par rapport à une référence

- Il peut y avoir différents types de références
  - Situer un élève par rapport à lui-même (par ex. un progrès)
  - Le situer par rapport au groupe
  - Le situer par rapport aux attendus de son âge, de son cycle, du programme
  - Encore une fois intérêt des compétences :  
évaluation par rapport à la réalisation d'une tâche

# Pour qui on évalue ?

- Pour l'enseignant
- Pour le système
- Pour l'élève

# Quand évalue-t-on ?

- En fin d'apprentissage = évaluation
- En cours d'apprentissage = régulation

- De façon non intrusive
- De façon intrusive
- De façon standardisée
  
- De l'intérêt de l'autoévaluation
- De l'intérêt d'un contrat clair
  - Qu'est-ce qui va être évalué ?
  - Avec quelle activité ?
  - Selon quels critères ?

# Les fonctions de l'évaluation

- Pour sélectionner
- Pour valider
- Pour sanctionner
- Pour aider à apprendre
- Pour motiver
- Pour prendre conscience, réguler son propre apprentissage
- Pour dépister
- Pour évaluer le système, l'établissement, le dispositif, etc.

# Principaux diagnostics

- L'élève ne travaille pas
- Il est persuadé qu'il ne peut pas y arriver
- Il n'a pas la connaissance nécessaire à la réalisation de la tâche
- Il a la connaissance mais ne parvient pas à la mobiliser
- Il mobilise la bonne connaissance mais fait une erreur de mise en œuvre
- Il évalue mal l'atteinte du but

# Pour aider les élèves

- Centration sur le progrès et la maîtrise personnels
- Faire des évaluations privées, non publiques
- Valoriser l'effort personnel
- Encourager la conception de l'erreur comme inhérente à l'apprentissage

# Plan

1. Introduction
2. Perceptions et effets secondaires de l'évaluation
3. Les limites de l'évaluation
4. Les fonctions de l'évaluation
5. Conclusion

Pouvons-nous relever le pari  
de la rationalité ?

# Je me risque à quelques pistes...

- Être prudent avec les évaluations en général et les notes en particulier
- Bien connaître les biais
- Ménager une place aux évaluations non-intrusives, par observation
- Contractualiser clairement les objectifs et les modalités des évaluations, notamment sommatives
- Multiplier les formes d'évaluation pour véritablement diagnostiquer
- Être clair sur les formats de connaissance visés par l'évaluation
- Utiliser l'évaluation comme moteur de progrès et de motivation

# Evolution possible ???

- Faire du socle le cœur de l'évaluation
- Evaluer les progrès
- Souligner les réussites plutôt que stigmatiser les échecs
- Réduire **beaucoup** la place des évaluations intrusives
- Mettre à disposition des équipes des évaluations standardisées
- Concevoir collectivement l'évaluation

Merci pour votre attention !

[Andre.Tricot@univ-tlse2.fr](mailto:Andre.Tricot@univ-tlse2.fr)