

Le testing adaptatif...

... en quelques bits (36624992)

[Christophe Lalanne, CIEP/DEC]

Sommaire

- estimation de l'habileté des candidats
- objectifs du testing adaptatif
- quelques solutions...

Concept hypothétique

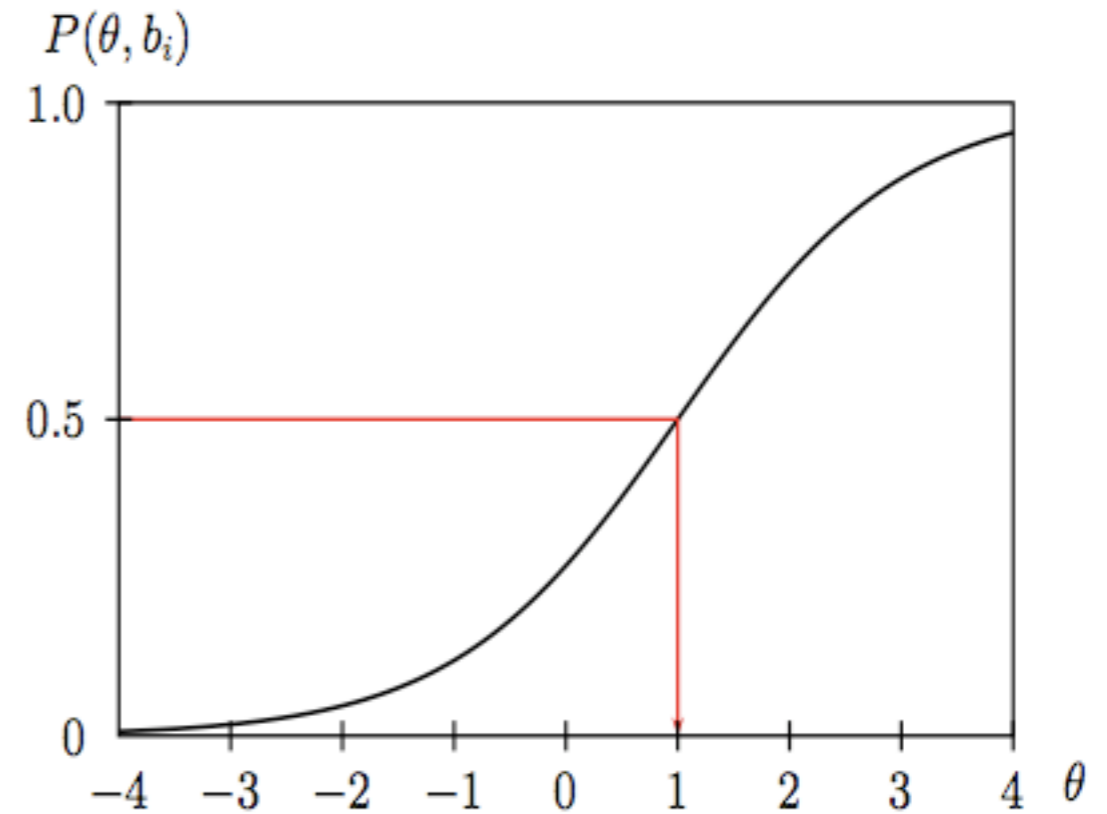
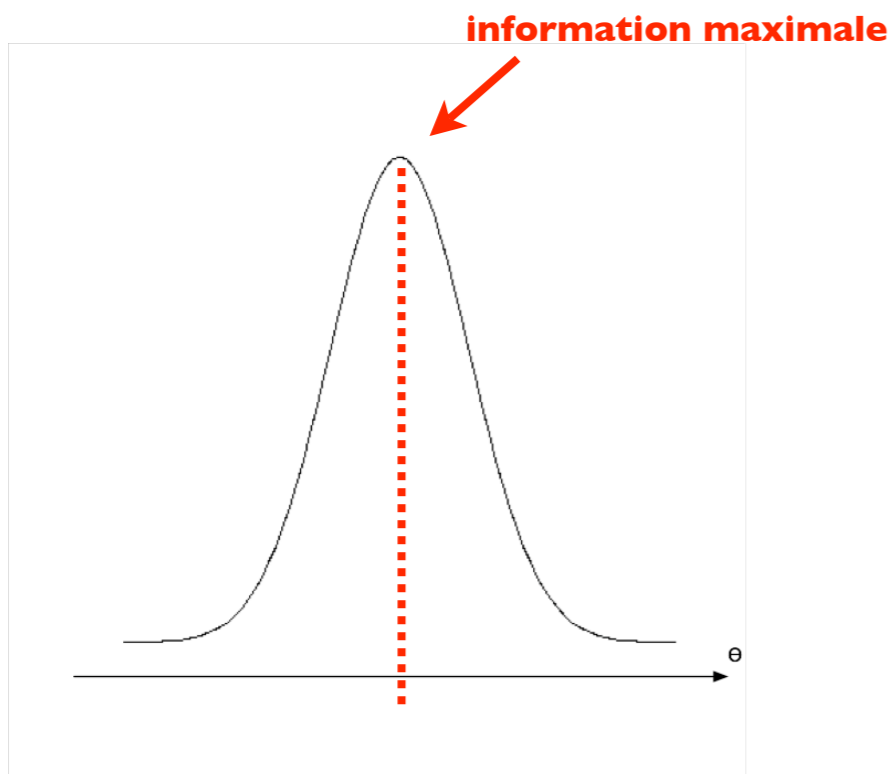
C2	Supérieur avancé Excellente maîtrise de la langue. La personne comprend sans effort pratiquement tout ce qu'elle lit ou entend et peut tout résumer de façon cohérente. Elle s'exprime très couramment et de façon différenciée et nuancée sur des sujets complexes.	C2
C1	Supérieur Bonne maîtrise de la langue. La personne peut comprendre une grande gamme de textes longs et exigeants comportant des contenus implicites. Elle s'exprime couramment et de façon bien structurée sur sa vie sociale, professionnelle ou académique et sur des sujets complexes.	C1
B2	Intermédiaire avancé Maîtrise générale et spontanée de la langue. La personne peut comprendre l'essentiel d'un texte complexe. Elle peut participer à une conversation sur un sujet général ou professionnel de façon claire et détaillée en donnant des avis argumentés.	B2
B1	Intermédiaire Maîtrise efficace mais limitée de la langue. La personne comprend un langage clair et standard s'il s'agit d'un domaine familier. Elle peut se débrouiller en voyage, parler de ses centres d'intérêt et donner de brèves explications sur un projet ou une idée.	B1
A2	Élémentaire avancé Maîtrise élémentaire de la langue. La personne peut comprendre des phrases isolées portant sur des domaines familiers. Elle peut communiquer dans des situations courantes, et évoquer avec des moyens simples des questions qui la concernent.	A2
A1	Élémentaire Maîtrise de base du français. La personne est capable de comprendre des situations simples et concrètes se rapportant à la vie quotidienne. Elle peut communiquer de façon simple si l'interlocuteur parle lentement.	A1

Les modèles de réponse à l'item

- objectif : permettre l'estimation du niveau d'habileté d'une personne sur une échelle de compétence définie a priori
- méthode : modèle statistique de réponse à l'item
- utilisation d'un instrument de mesure : permet d'opérationnaliser le concept hypothétique

Approche psychométrique

Un item possède un niveau de difficulté propre, défini à partir d'un modèle probabiliste :
la difficulté est définie telle que la probabilité que l'individu réponde correctement vaut 1/2.
(maximisation de l'information)

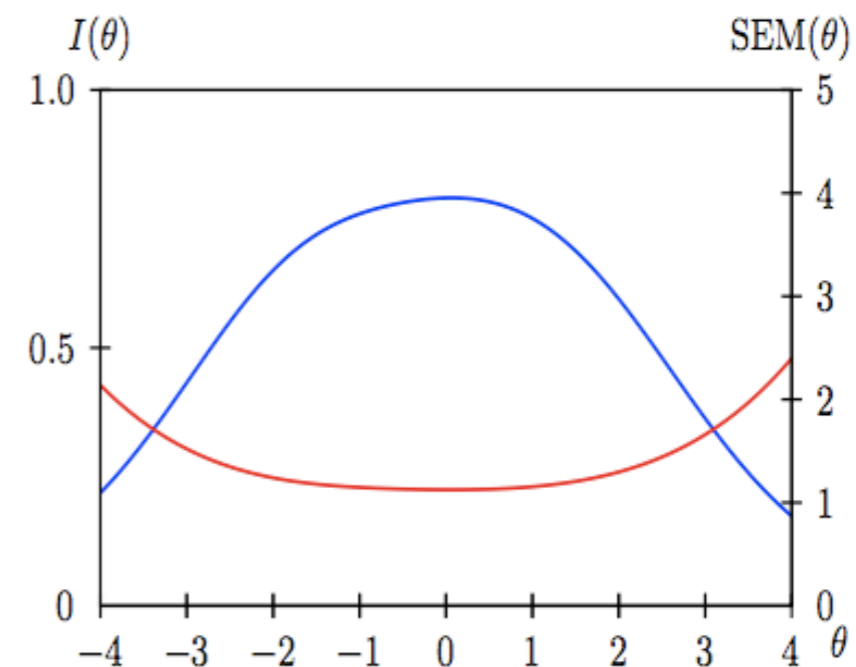
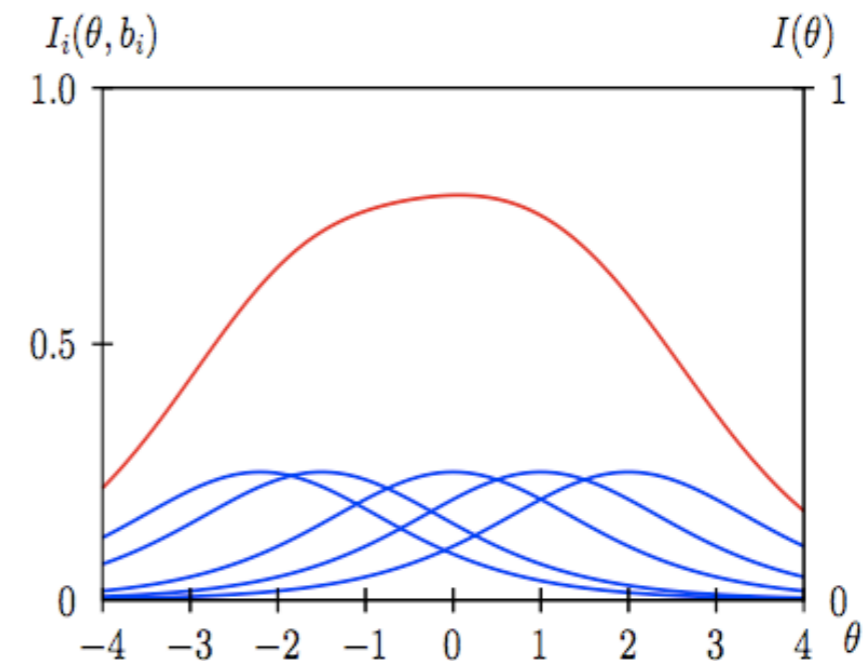
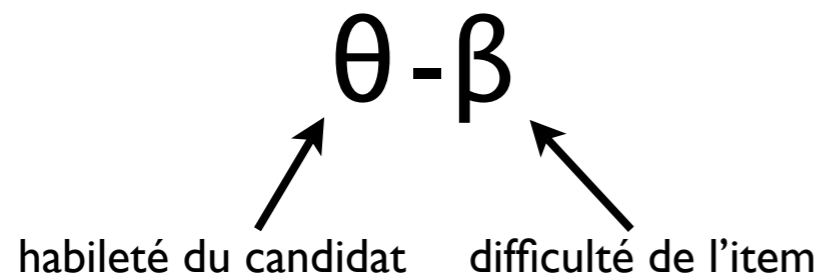


Approche psychométrique

Un test est composé d'un ensemble d'items, chaque item apportant sa contribution à l'information totale du test.

La qualité des mesures dépend en grande partie de l'adéquation du niveau du test avec celui des candidats.

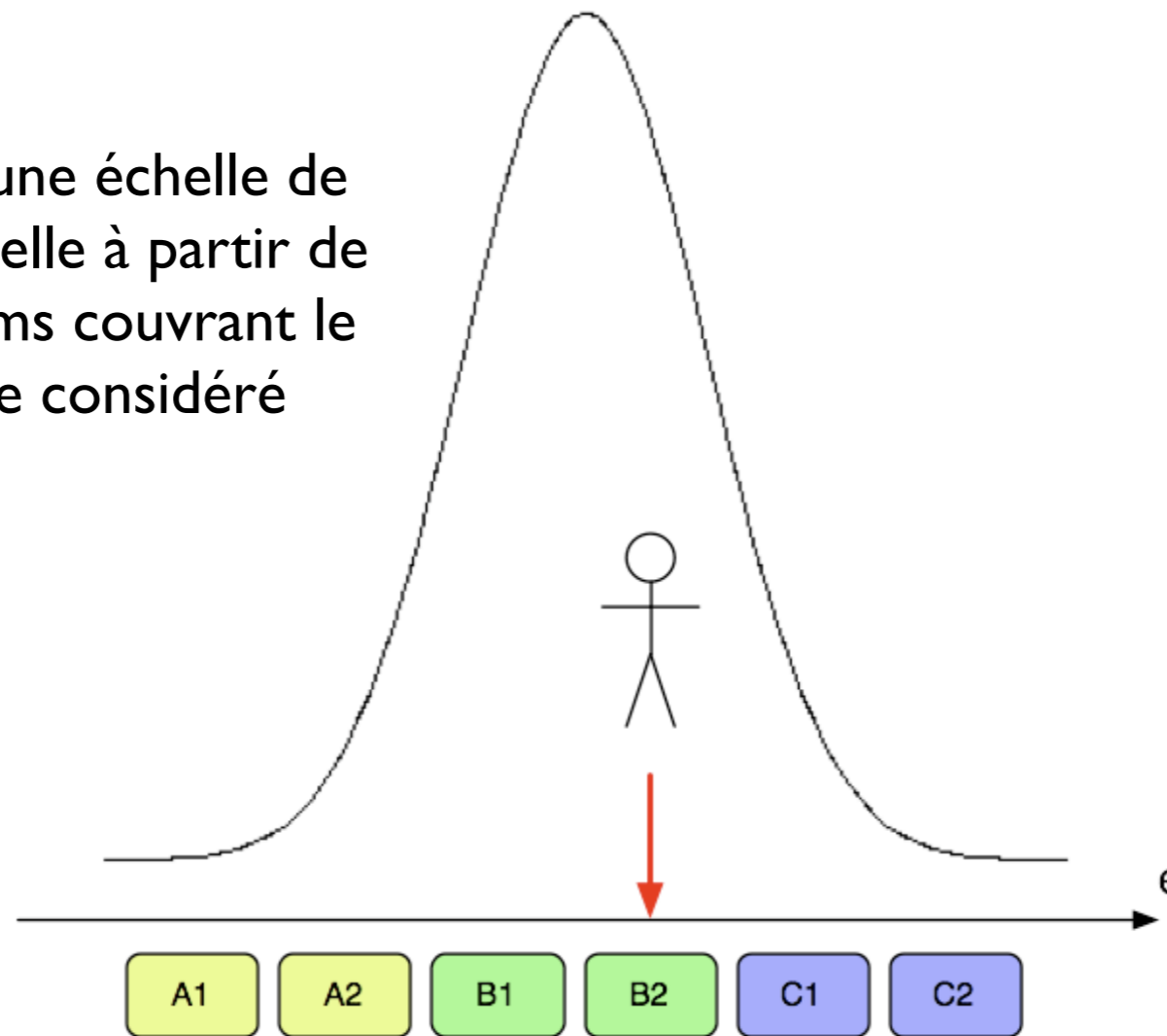
position du candidat déterminée par un MRI (p. ex. Rasch) :



Idée de base

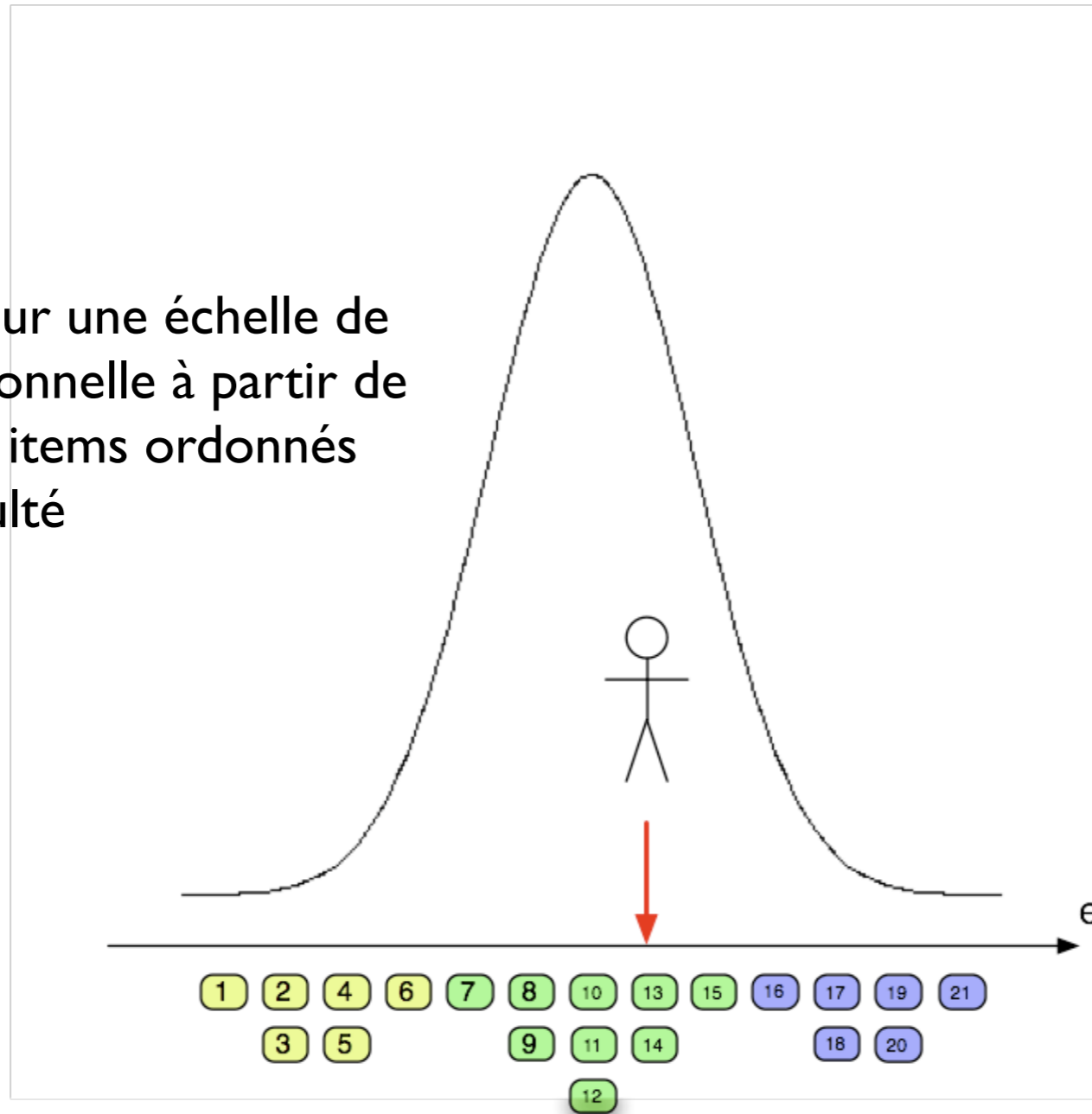
$$P_i(\theta) = P(\alpha_i^*, \beta_i, \theta) = \Psi(Z_i) = \frac{e^{Z_i}}{1 + e^{Z_i}} = \frac{1}{1 + e^{-Z_i}}$$

Situer un individu sur une échelle de mesure unidimensionnelle à partir de ses réponses à des items couvrant le spectre de compétence considéré



Idée de base (2)

Situer un individu sur une échelle de mesure unidimensionnelle à partir de ses réponses à des items ordonnés par degré de difficulté



Idée de base (3)

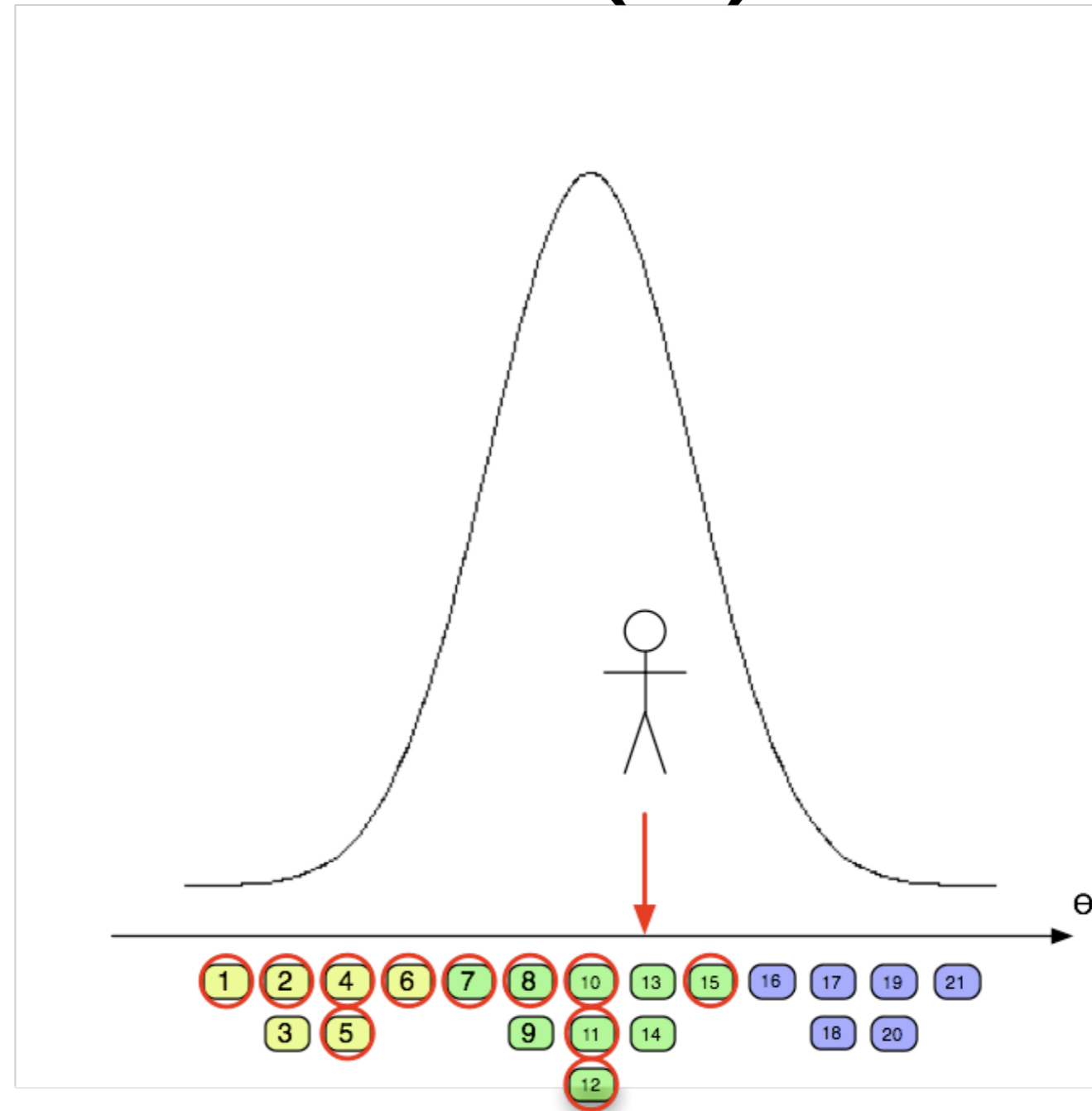
11/21 réponses
correctes

TCT

- difficulté des items
- pouvoir discriminant
- consistance interne

TRI

- difficulté des items
- habileté des candidats
- adéquation du modèle
- biais



Facteurs influençant l'estimation

- au niveau de la **calibration** :
 - ✓ pouvoir discriminant élevé
 - ✓ absence de fonctionnement différentiel
- au niveau du **test** : adéquation du niveau de difficulté des items avec le niveau des candidats (effets plancher/plafond, décentrage)

Test linéaire

- choix préconisé pour les tests certifiants
- intérêts : approche normative, standardisation et administration facilitées, banque d'items en élaboration continue
- inconvénient : précision de l'estimation généralement moins bonne (paradoxe d'atténuation)

Test adaptatif

- alternative au test linéaire qui présente plusieurs avantages dans un contexte de test de placement :
 - ✓ souplesse d'administration
 - ✓ obtention des résultats immédiate
 - ✓ précision de la mesure
- nombreuses contraintes additionnelles

Principe du test adaptatif

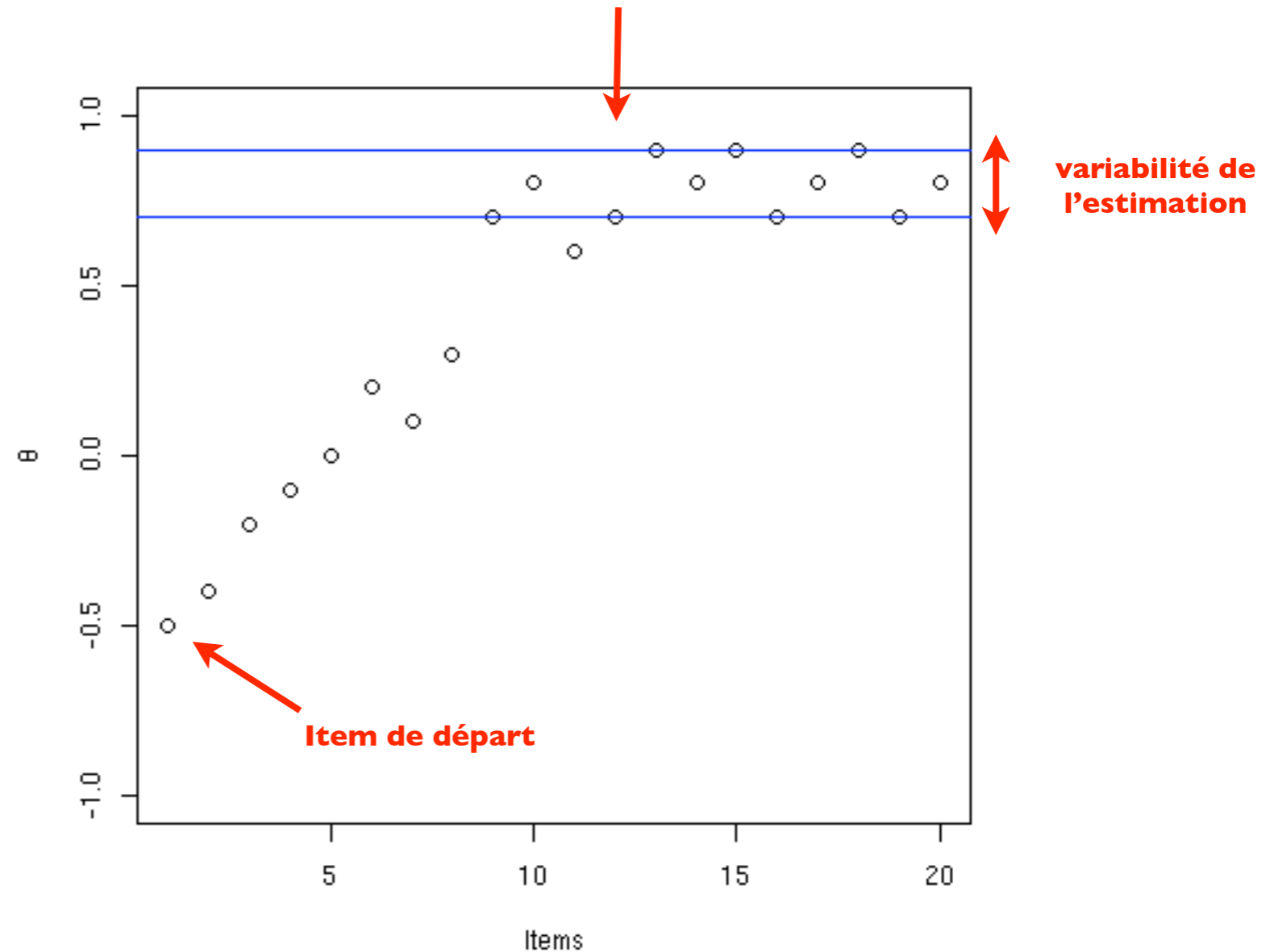
- évaluer dynamiquement le niveau du candidat (habileté) en sélectionnant les items les plus appropriés à son niveau
- optimiser à la fois le contenu (couverture du domaine) et la mesure (précision de l'estimation)

Approche basique

Proposer au candidat des items de difficulté croissante ; lorsqu'il échoue, proposer un item moins difficile.
Itérer jusqu'à parvenir à une erreur de mesure acceptable...

Problèmes liés à cette approche :

- choix de l'item de départ
- longueur du test de taille non fixe
- erreur de mesure = non nécessairement le critère le plus pertinent comme critère d'arrêt



Solutions pour les tests contraints

- partition de la banque d'items selon différents attributs et sélection d'items représentatifs (algorithme minimax)
- méthode de déviation pondérée :
spécifications du test exprimées sous forme de contraintes numériques (sélection séquentielle des items par minimisation d'une fonction d'utilité)

Solutions pour les tests contraints (2)

- testing adaptatif reposant sur des blocs d'items (*testlets*) : à la place d'unités discrètes, on travaille avec des paquets d'items partageant des spécifications communes (en termes de contenu)
- testing hiérarchique (*multi-stage*) : utilisation de mini-tests cohérents, ordonnés par difficulté (globale) croissante

Solutions pour les tests contraints (3)

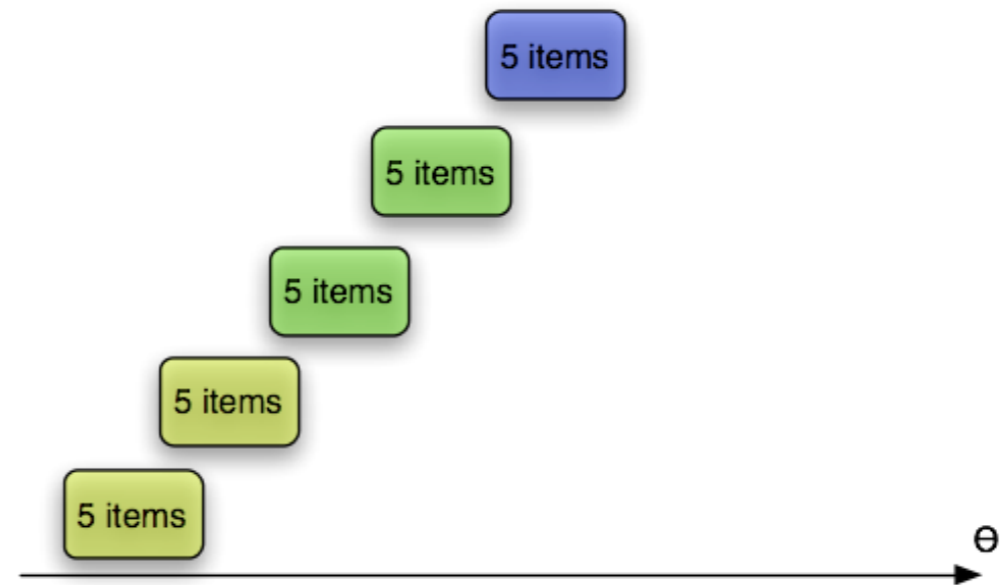
- méthode des tests virtuels (*shadow tests*) :
procédure itérative incluant la construction
de tests complets en vue d'utiliser un seul
item parmi l'ensemble (programmation
linéaire binaire)

Approches actuelles

- maximiser l'information apportée par une mesure ponctuelle (réponse à un item), en fonction des réponses antérieures :
contrainte de mesure
- procédure itérative de sélection des items, en respectant la couverture du domaine :
contrainte de contenu
- permet de garantir la validité du test

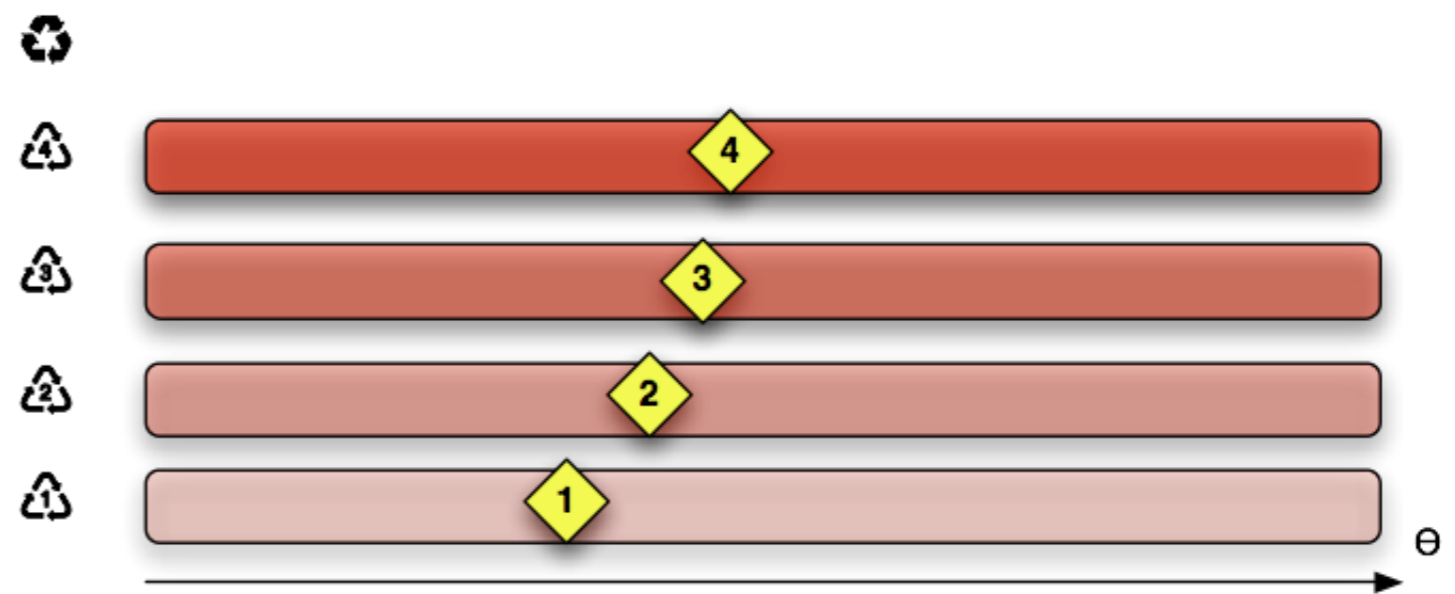
Méthode des blocs d'items

- Administration de différents blocs d'items insécables
- En fonction du degré de réussite du candidat (approche hiérarchique), la difficulté globale des blocs peut être augmentée ; sinon, administration linéaire des différents blocs



Méthode des tests virtuels

1. Initialisation de l'estimateur
2. Construction d'un test virtuel remplissant les contraintes et possédant l'information maximale
3. Administration de l'item de ce test possédant l'information maximale au voisinage de l'habileté estimée
4. Mise à jour de l'habileté estimée
5. Retour des items dans la banque
6. Ajustement des contraintes pour tenir compte de l'item utilisé
7. Itération des étapes 2 à 6



$$\max_i \sum_{i \in O_k} I_i(\hat{\theta}_{k-1}) x_i$$

Conclusions

- intérêt pour situer rapidement un individu sur l'échelle de compétence, avec une précision de mesure qui est fonction de la durée du test et du choix des items
- nécessite une banque d'items consistante et bien alimentée, pour éviter les problèmes de couverture et satisfaire les contraintes de sélection