Evaluer les compétences

Marc Braun

Faculté de médecine Université de Lorraine à Nancy





PRINCIPES GENERAUX

- 1. TESTER LES COMPETENCES (Qualification médicale et habiletés à l'exercice)
- 2. OBLIGE À CONSTRUIRE UNE ÉVALUATION
 - 1. CONFORME AUX PROGRAMMES (Blueprint)
 - 2. VALIDE ET FIDÈLE
 - 3. DONT LE NIVEAU MINIMUM ACCEPTABLE DOIT ÊTRE DÉFINIT À L'AVANCE
 - 4. AYANT CLARIFIE LES VERSANTS SOMMATIF ET FORMATIF (très difficile)

CONTEXTE GENERAL

 Les étudiants se sentent surchargés de travail et répondent en étudiant seulement les enseignements qui sont évalués.

MAIS

- La demande sociétale,
- · l'accent mis sur la performance des médecins et
- l'assurance que les médecins soient compétents,
 - l'évaluation une fonction sommative.
- Les tests de compétences cliniques,
 - -qui permettent de prendre une décision d'aptitude ou non
- Cette demande est un DEFI pour tous les acteurs

CONTEXTE GENERAL

Sovez clair sur le hut du test

Les programmes d'évaluation doivent correspondre aux compétences apprises et les formats d'enseignement.

Sommatif / formatif

	Soyez ciaii sui le but du test
Planifier	Planifier le test en fonction des objectifs d'apprentissage, des cours ou des compétences essentielles de la spécialité.
Validité	Sélectionnez les formats de test appropriés aux compétences à tester. Cette action entraîne invariablement un examen composite.

Échantillon fiable de tests.

Compétences cliniques sont différentes selon les tâches.

La longueur du test est cruciale si les décisions de haut niveau sont attendues.

Utilisez autant de examinateurs que possible

Définir les standards minimum

Définir le but final de l'évaluation.

Définir le niveau minimum exigible l'avance.

Items par items (long mais indispensable)

PROBLEME DE LA NORME

- Une norme claire doit être définie
 - Médecin jugé inapte à pratiquer
 - et un étudiant à valider son examen
- La norme minimale est décidée avant l'évaluation.
- Ces normes sont définies en référence à de critères. (Justifications)
- La variation de la difficulté du test
- Les normes définies pour chaque test d'évaluation, item par item

FIABILITE vs VALIDITE

- La fiabilité est une mesure de la reproductibilité ou la cohérence d'un test,
- Est influencée par de nombreux facteurs
 - jugements des examinateurs,
 - les cas utilisés
 - La nervosité des candidats et les
 - conditions générales de l'examen.
- Deux aspects de la fiabilité :
 - inter-évaluateur et
 - inter-cas (candidats)



Fiabilité inter-évaluateur

- <u>Définition</u>: mesure la cohérence de la notation de la performance par différents examinateurs.
- Plusieurs examinateurs et différents cas améliore la fiabilité inter-évaluateur.
- Dans un examen oral, le jugement moyen de dix examinateurs, chacun évaluant le candidat sur une question, produit un <u>test plus fiable</u> que celui d'un examinateur poser dix questions.

Swanson DB. A measurement framework for performance based tests. In: Hart IR, Harden RM, eds. Further developments in assessing clinical competence. Montreal: Can-Heal, 1987: 13–45.

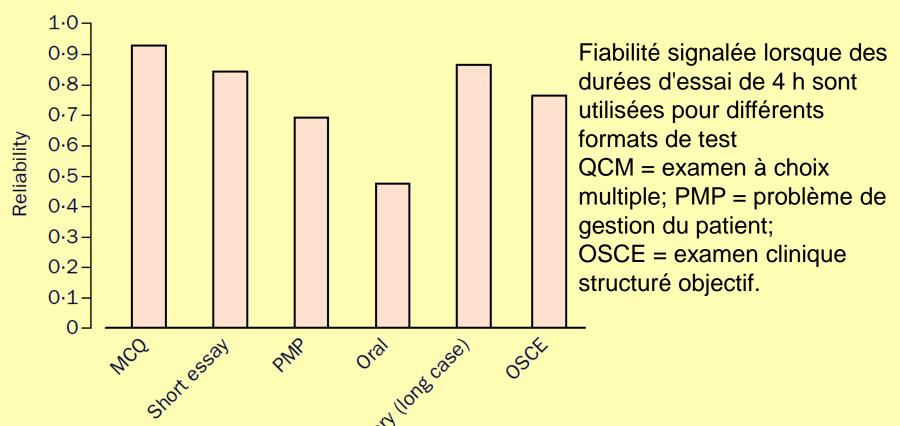
Fiabilité inter-étudiants

- Question la plus importante : La cohérence de la performance des candidats à propos d'un cas (fiabilité d'inter-individuelle)
- Un spectre large de cas est essentiel pour évaluer la compétence clinique fiable.
- Différences d'expériences individuelles rencontrées pendant la formation et la pratique,
- Remise en cause de l'approche traditionnelle par laquelle la compétence des candidats a été évaluée sur un seul cas.

GAGES DE FIABILITE

- Tests de compétences cliniques sont développés dans le format de <u>cas multiples</u> avec le développement de l'examen clinique objectif structuré (ECOS),
- consistant en une série de tâches et de rencontres (stations).
- De <u>nombreuses stations et un temps</u> de test suffisant sont essentiels pour atteindre une fiabilité d'interindividuelle adéquate pour le test.
- Quel que soit le format du test, <u>la longueur est</u> décisive pour la fiabilité de tout test de compétence clinique.

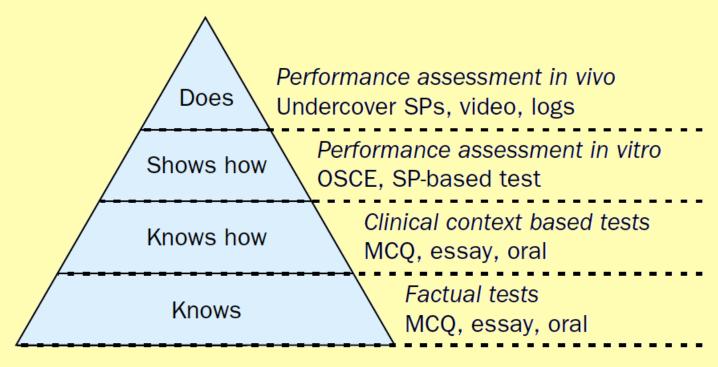
GAGES DE FIABILITE

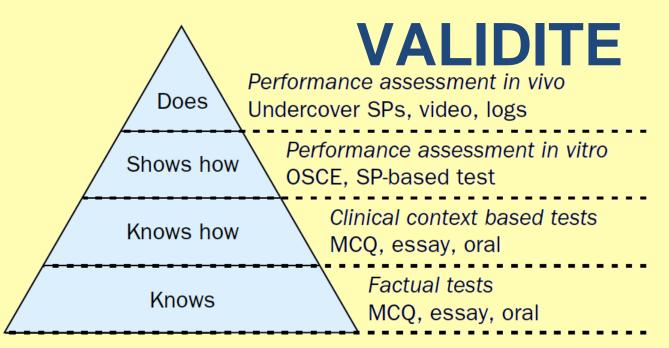


- Norcini JJ, Swanson DB, Grosso LJ, Shea JA, Webster D. Reliability, validity and efficiency of multiple choice questions and patient management problem items formats in the assessment of physician competence. Med Educ 1985; 19: 238–47.
- Stalenhoef-Halling BF, van der Vleuten CPM, Jaspers TAM, Fiolet JFBM. The feasibility, acceptability and reliability of openended questions in a problem based learning curriculum. In: Bender W, Hiemstra RJ, Scherpbier AJJA, Zwiestra RP, eds. Teaching and assessing clinical competence. Groningen: Boekwerk,1990: 1020–31.
- Wass V, Jones R, van der Vleuten CPM. Standardised or real patients to test clinical competence? The long case revisited. Med Educ 2001, 35(4)-321-325.
- Newble DI, Swanson DB. Psychometric characteristics of the objective structured clinical examination. Med Educ 1996; 22: 325–34.

VALIDITE

- <u>Définition</u>: Explore le potentiel d'un test à effectivement tester les compétences pour lequel il est conçu
- La pyramide de la compétence (Miller) est un modèle conceptuel simple, qui décrit les problèmes impliqués lors de l'analyse de la validité.





- La base représente la connaissance composante de la compétence: « sait » (connaissances de base) suivi par « sait comment faire » (connaissance appliquée).
- la facette la plus importante de compétence requise par un médecin qualifié, c.-à-d. « démontre comment faire ».
- Cependant, le but ultime d'une évaluation valide la compétence clinique est de tester ce que le médecin « fait » dans le lieu de travail.

EVALUER l'étape« Montre comment faire »

PLUSIEURS FORMATS TRADITIONNELS POSSIBLES:

- 1 CAS LONG & 3 CAS COURTS DE PATIENTS.

Peu de généralisation, trop hétérogène et aléatoire, trop peu de cas pour avoir une pertinence

- •Variante OSLER (Objective Structured Long Examination Record) d'observations prolongées de patients
- •Newble DI. The observed long case in clinical assessment. Med Educ 1994; 25: 369–73.
- •Price J, Byrne GJA. The direct clinical examination: an alternative method for the assessment of clinical psychiatric skills in undergraduate medical students. Med Educ 1994; 28: 120–25.
- Variante cas plus raccourcis et plus nombreux
- •Shatzer JH, Wardrop JL, Williams RC, Hatch TF. The generalisability of performance of different station length standardised patient cases. Teach Learn Med 1994; 6: 54–58.

TOUS SONT PEU FIABLES et NON REPRODUCTIBLES

Mini Clinique (Mini-CEX)



Étudiant :	Service:							
Diagnostic:								
Catégorie de problème clinique : Douleur Dyspnée Psychiatrie – comportement	Dyspnée Circulatoire				☐ Neurologie ☐ Digestif			
Objectif de l'évaluation : Interrogatoire Diagnostic			□P	rise en ch	arge	Expli	cations	
	patient a été vu par l'étudiant auparavant : Patient : 1 à 4							
Difficulté du cas : Limite Basse	Moye	nne	ΔĘ	llevée				
Score	1	2	3	4	5	6	Non évaluable	
			Limite	Atteint le niveau attendu		ssus du attendu	(pas d'opportunité d'observation)	
1. Interrogatoire								
2. Examen physique								
3. Communication - professionnalisme								
Jugement clinique								
5. Information								
5. Organisation - efficacité								
7. Compétence globale								
Points forts	Comr	mentaires		Poi	nts faibles	,		
Toms fores				101	nts ratoles	,		
Suggestions d	u tuteur po	ur amélior	er la mini c	clinique				
Évaluateur :		PU-PH Interne Autre :		ACU-PH urveillant		□ CCA □ IDE		
Nombre de mini cliniques évaluées	s: [0 🗆	1 2	3	<u>4</u>	5-9	> 9	
Durée de l'observation :		Durée du feed-back :						
Signature:				Date:		JP F	Fournier Nice	÷

Norcini J et V Burch. Workplace-based assessment as an educational tool: AMEE Guide No. 31.

Med Teach. 2007 Nov;29(9):855-71.

EVALUER l'étape« Montre comment faire »

- STANDARDISER LES CAS ET ECHANTILLON CONTRÔLE : ECOS (OSCE) Harden RM, Gleeson FA. ASME medical educational booklet no 8: assessment of medical competence using an objective structured clinical examination (OSCE). J Med Educ 1979; 13: 41–54.
- Les candidats passent par une série de stations évaluant des compétences cliniques appliquées dans divers contextes.
- Large échantillonnage de cas et évaluation structurée
- Cher et laborieux.
- Améliore la fiabilité
- · La validité peut être perdue car
- les <u>compétences complexes</u>, nécessitent un jugement professionnel intégré MAIS fragmenté par la brieveté de la station (5-10 min).

ECOS (2)

- L'évaluation de la communication et des attitudes de l'OSCE.
 - spécifique à chaque cas
 - faible généralisation
- Pour évaluer l'empathie de manière fiable, 37 scénarios différents Colliver JA, Willis MS, Robbs RS, Cohen DS, Swartz MH. Assessment of empathy in a standardized-patient examination. Teach Learn Med 1998; 10: 8–11.
- la perception par les patients des aptitudes à communiquer des médecins sont <u>fiables</u> et utilisés dans de nombreux patients normalisés
- Williams RG, McLaughlin MA, Eulenberg B, Hurm M, Nendaz MR. The patient findings questionnaire: one solution to an
- important standardized patient examination problem. Acad Med 1999; 74: 1118–24.

Qui devrait être le juge: le patient ou l'examinateur ?

- Evaluation faite par le patient ou l'examinateur n'a pas d'importance.
- van der Vleuten CPM, Swanson DB. Assessment of clinical skills with standardised paients: state of the art. Teach Learn Med 1990; 2: 58–76.
- Colliver JA, Verhulst SJ, Williams RG, Norcini JJ. Reliability of performance on standardised patient cases: a comparison of consistency measures based on generalizability theory. Teach Learn Med 1989; 1: 31–37.

ECOS (3)

- Cocher sur une liste des items n'est pas aussi objective
- Reznick RK, Regehr G, Yee G, Rothman A, Blackmore D, Dauphinee D. Process-rating forms versus task-specific checklists in an OSCE for medical licensure. Acad Med 1998; 73: S97–99.
- Les évaluations globales, par les médecins, sont aussi fiables que les checklists.
- Regehr G, MacRae H, Reznick R, Szalay D. Comparing the psychometric properties of checklists and global rating scales for
- assessing performance on an OSCE-format examination. Acad Med 1998; 73: 993-97.
- Schwatz MH, Colliver JA, Bardes CL, Charon R, Fried ED, Moroff S. Global ratings of videotaped performance versus global ratings of actions recorded on checklists: a criterion for performance assessment with standardized patients. Acad Med 1999; 74: 1028–32.
- Une formation approfondie des juges est nécessaire pour assurer la cohérence.
- CONCLUSIONS Ce n'est pas un véritable «étalon-or»
- PATIENTS STANDARDISES (Normalisés)
 - SIMULATION OUTILS VALIDES

PLACE DE LA SIMULATION

(cf ATELIER 3 Jeudi 30)

- Identification d'objectifs d'apprentissage ;
- Identification de cas à résoudre ;
- Construction de scenari;
- Choix de la technique de simulation ;
- Chaque étudiant accomplit plusieurs tâches ;
- Check-list ou évaluation globale ; observateurs entraînés et en nombre suffisant ;
- Comparabilité du niveau de difficulté d'une tâche à l'autre ;
- Vérification de l'extrapolation en pratique réelle.

LE VRAI CHALLENGE LE SOMMET

- Evaluation de la performance réelle en garde et consultation.
- En formation post-universitaire,
- LA REVALIDATION
- CONSTAT DE L'INSUFFISANCE PROFESSIONNELLE (Atelier n° 9 30 Novembre)
- INTERET PUBLIC
- BALANCE RESPECTEE VALIDITÉ ET FIABILITÉ
- SIMULATION et PORTFOLIO (Atelier n° 11 Jeudi 30)

Shows how

Knows how

Knows

