

Test de Concordance des Scripts

JP. Fournier

Département de Pédagogie Médicale

Faculté de Médecine de Nice Sophia Antipolis

TCS, ECNi, et SIDES

Perspect Med Educ (2012) 1:162–171
DOI 10.1007/s40037-012-0024-1

ORIGINAL ARTICLE

Preferred question types for computer-based assessment of clinical reasoning: a literature study

**Lisette van Bruggen •
Margreet Manrique-van Woudenbergh •
Emely Spierenburg • Jacqueline Vos**

SCT and CIP yield higher discriminatory values and therefore better distinguish between prepared and unprepared students. Higher discriminatory values of questions increase the reliability of the test where fewer questions are needed. However, in the literature studied on SCT there is no mention of assessment of undergraduate students using this question type. The need for an expert panel for SCTs is labour intensive and therefore does not reduce the workload of teachers. It is not yet possible to organize an expert panel in CBA. SCTs cannot be recommended for use in CBA.

Conférence des Doyens, septembre 2013

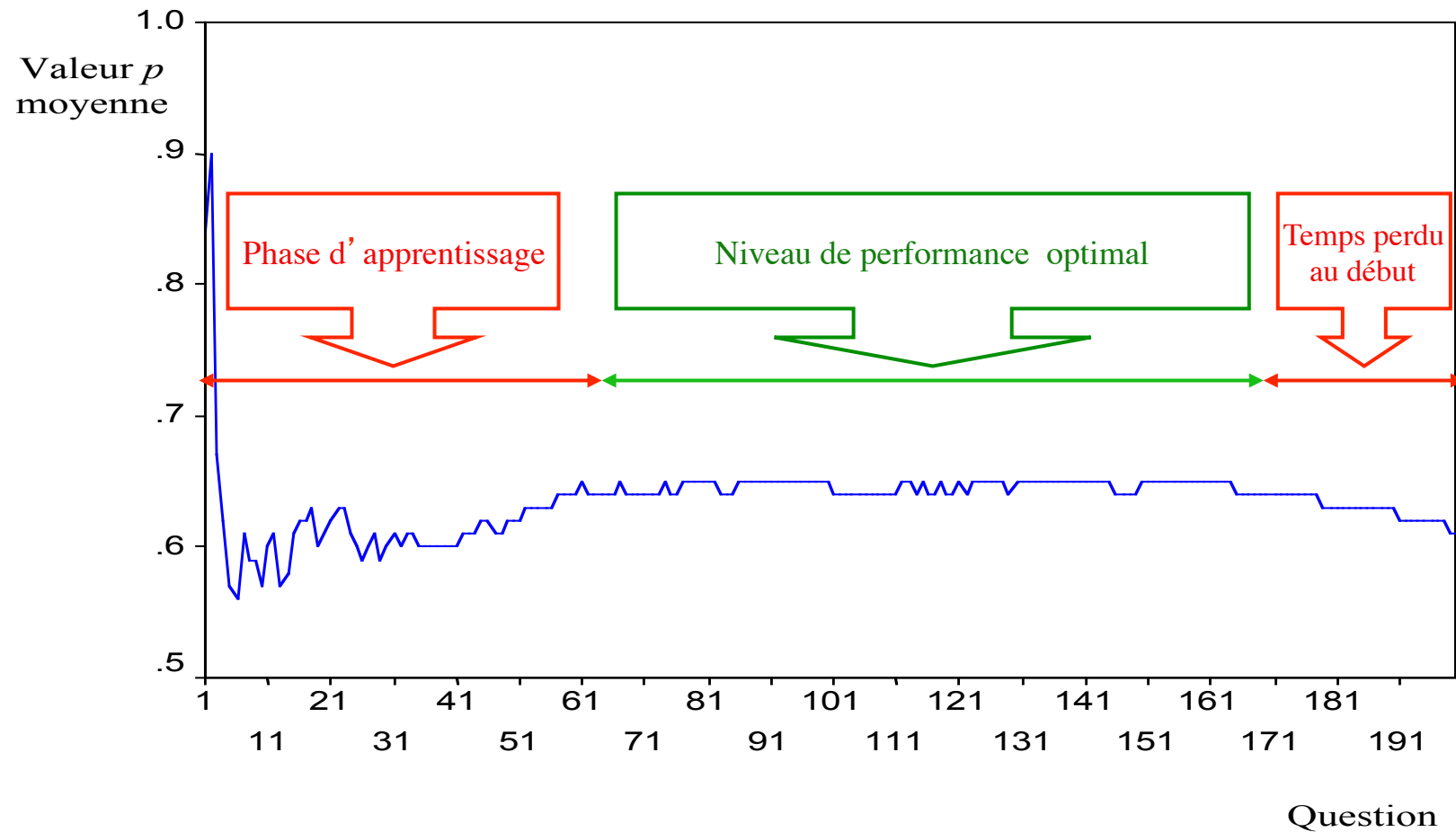
ECNi :

- ❖ Mai 2016 ;
- ❖ Dans chaque Faculté, sur tablettes numériques ;
- ❖ 3 jours d'épreuve :
 - ✓ 24 dossiers cliniques progressifs (4 demi-journées) ;
 - ✓ 1 demi-journée de questions isolées ;
 - ✓ 1 demi-journée de LCA (articles en Français) ;
 - ✓ TCS : 5 p. cent.

CNCI (novembre 2014)

- ❖ TCS en 2017 ? ;
- ❖ ECNi sur 10 000 points ;
- ❖ LCA :
 - ✓ 3 heures ;
 - ✓ 2 Dossiers progressifs ;
 - ✓ 1 sur une étude clinique, 1 sur l'analyse d'un document scientifique ;
 - ✓ Articles en Anglais en 2017 ? ;
- ❖ En 2016 :
 - ✓ Dossier : 70 p. cent ?, 18 dossiers ;
 - ✓ QI : 20 p. cent ;
 - QRM à 5 options de réponse ;
 - Suppression des QRU ;
 - QROC supprimés ?
 - ✓ LCA : 10 p. cent.

L'enjeu, c'est...



De quoi parle-t-on ?

C'est un test utilisé dans les professions de santé pour mesurer un aspect particulier du raisonnement clinique : l'aptitude à interpréter des informations médicales dans un contexte d'incertitude.

Il présente un certain nombre de caractéristiques :

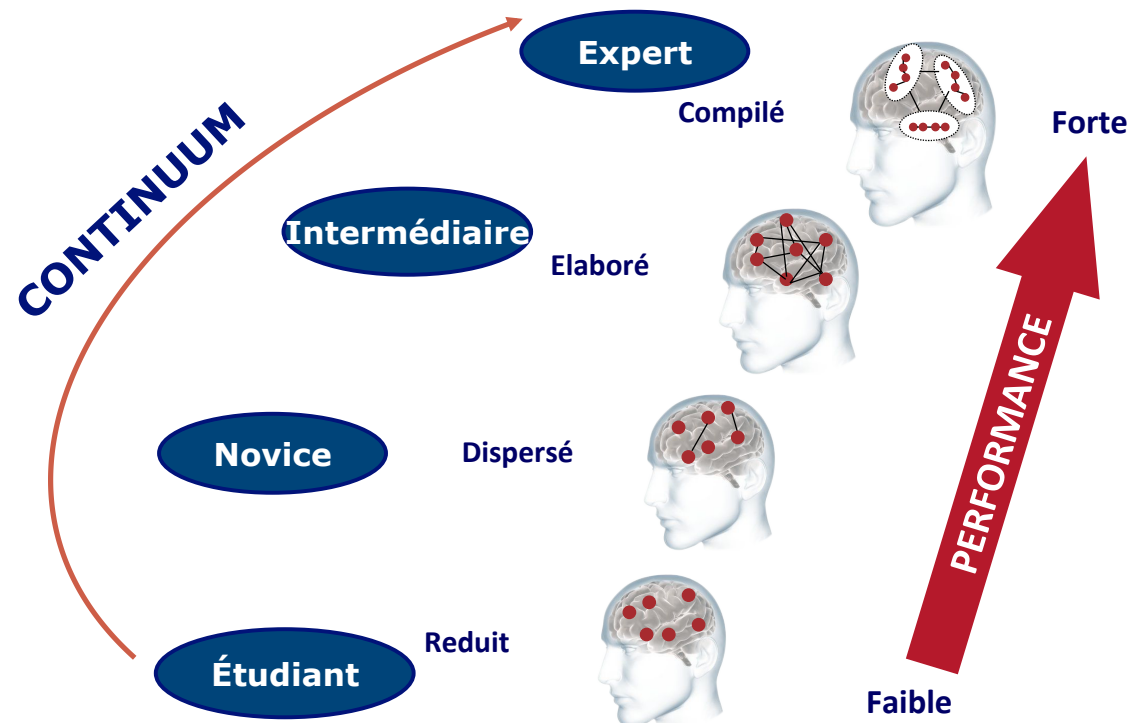
- ❖ Technique d'évaluation standardisée ;
- ❖ Ne s'applique qu'aux aspects cliniques ;
- ❖ Repose sur des bases conceptuelles solides ;
- ❖ Format docimologiquement correct ;
- ❖ Utilisable de la formation pré-graduée jusqu'au DPC.

Bases conceptuelles : 4 concepts majeurs

- ❖ Théorie des scripts ;
- ❖ Raisonnement clinique ;
- ❖ Scores combinés (*Agregate scoring*) ;
- ❖ *The reflective practitioner*.

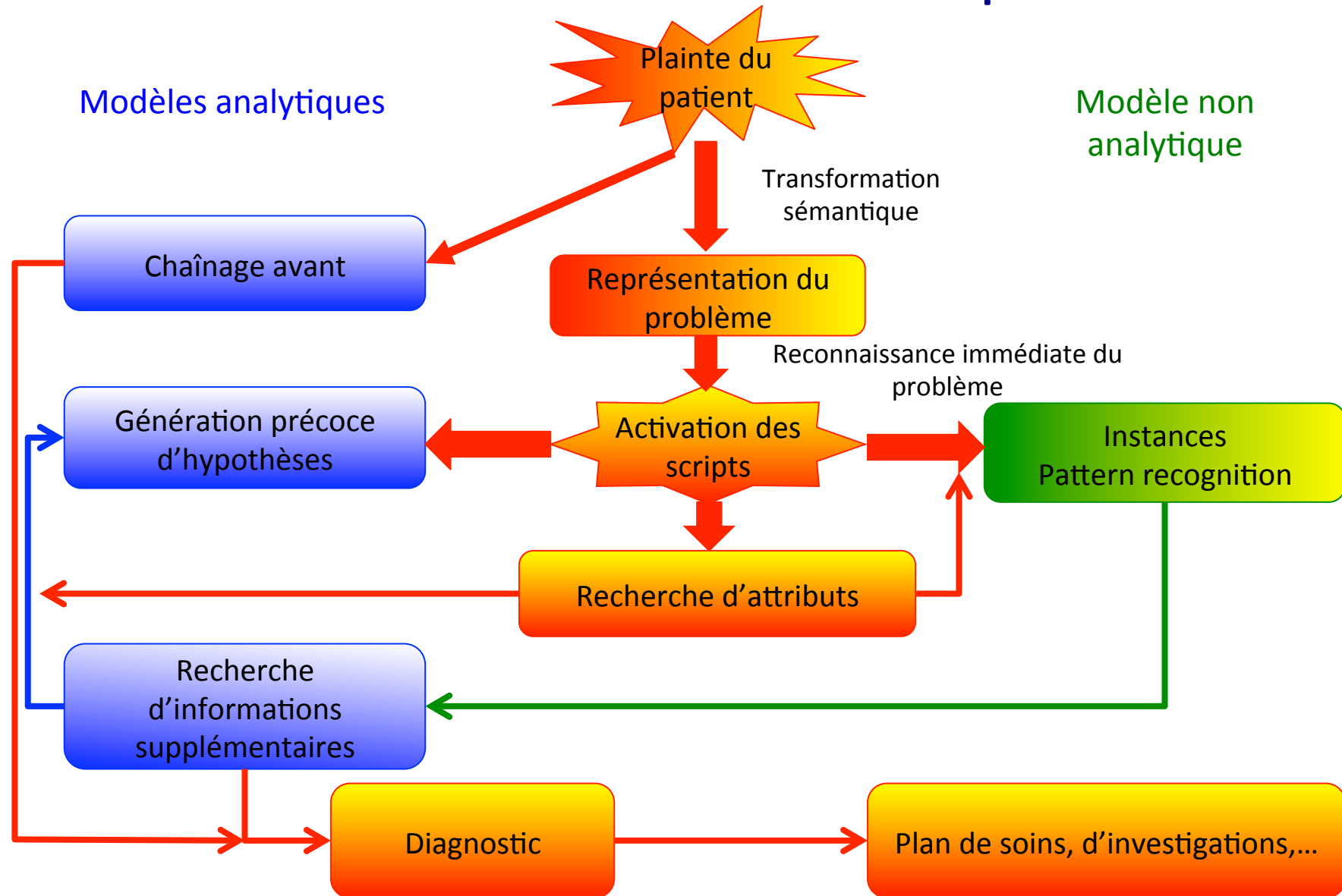
Théorie des scripts

Script illness : cognitive entity containing relatively little knowledge about pathophysiological causes of symptoms or complaints (because of encapsulation), but a wealth of clinically relevant information about the enabling conditions of diseases (= attributs) as a product of growing experience with how disease manifests in daily life



- ❖ Scripts d'autant plus riches que contacts avec les patients plus précoces et plus nombreux ;
- ❖ Scripts non transférables ;
- ❖ Apparaissent vers le milieu du cursus clinique.

Raisonnement clinique

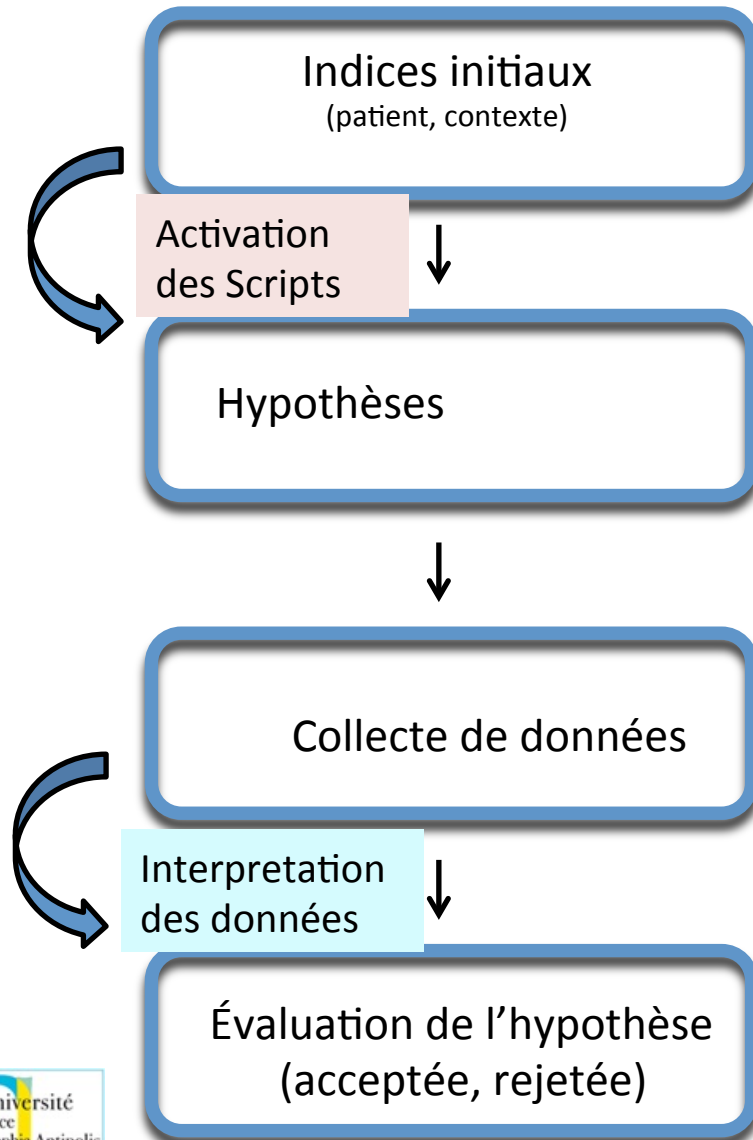


Test de Concordance des Scripts

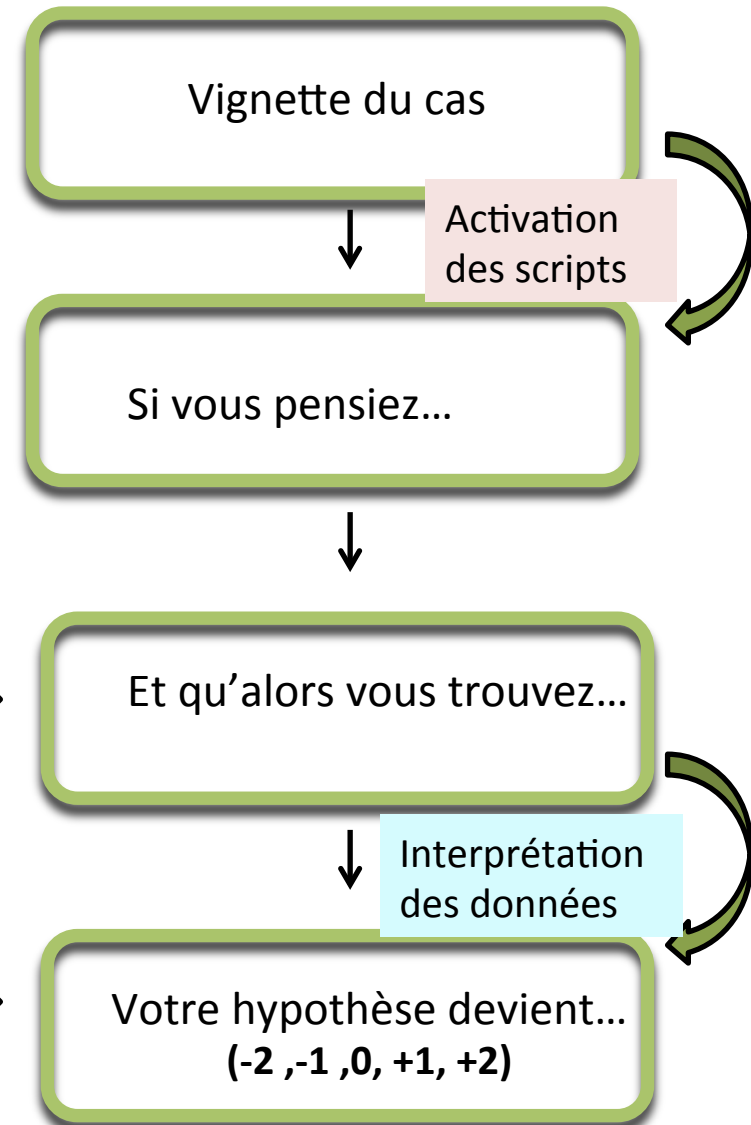
Vous visitez un parc d'attractions au Japon. Vous n'êtes pas encore très familiarisé(e) avec les idéogrammes. Vous cherchez un restaurant. Vous avez compris qu'il existait des restaurants « à thème »

Si vous pensiez...	..et qu'alors vous constatez	Votre hypothèse devient...				
Que ce restaurant sert une cuisine américaine	Qu'il y a une bouteille de Ketchup sur la table	-2	-1	0	+1	+2
Que ce restaurant sert une cuisine française	Qu'il y a une nappe à carreaux sur la table	-2	-1	0	+1	+2
Que ce restaurant sert une cuisine italienne	Qu'il y a une bouteille d'huile d'olive sur la table	-2	-1	0	+1	+2
-2 : Absolument pas probable -1 : peu probable	0 : ni plus, ni moins probable	+1 : Probable +2 : Très probable				

Raisonnement clinique : Étapes clés



Format TCS

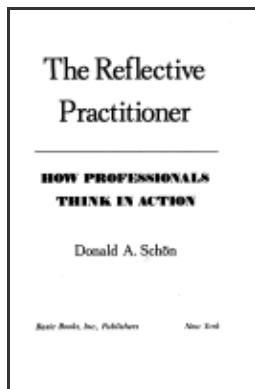


Scores combinés

- ❖ Gold standard : réponse du plus grand groupe d'experts (réponse modale) ;
- ❖ Autres réponses d'experts : différences d'interprétation, cliniquement pertinentes.

The reflective practitioner

The reflective practitioner: how professionals think in action



[Donald A. Schön](#)

[16 Avis](#)

Basic Books, 1983 - 374 pages

A leading M.I.T. social scientist and consultant examines five professions: engineering, architecture, management, psychotherapy, and town

[Plus »](#)

Une des différences majeures entre experts et juniors est l'aptitude des experts à prendre une décision en contexte d'incertitude, quel que soit le domaine.

Test de Concordance des Scripts

Vous visitez un parc d'attractions au Japon. Vous n'êtes pas encore très familiarisé(e) avec les idéogrammes. Vous cherchez un restaurant. Vous avez compris qu'il existait des restaurants « à thème »

Si vous pensiez...	..et qu'alors vous constatez	Votre hypothèse devient...				
Que ce restaurant sert une cuisine américaine	Qu'il y a une bouteille de Ketchup sur la table	-2	-1	0	+1	+2
Que ce restaurant sert une cuisine française	Qu'il y a une nappe à carreaux sur la table	-2	-1	0	+1	+2
Que ce restaurant sert une cuisine italienne	Qu'il y a une bouteille d'huile d'olive sur la table	-2	-1	0	+1	+2
-2 : Absolument pas probable -1 : peu probable	0 : ni plus, ni moins probable	+1 : Probable +2 : Très probable				

Test de Concordance des Scripts

Vous visitez un parc d'attractions au Japon. Vous n'êtes pas encore très familiarisé(e) avec les idéogrammes. Vous cherchez un restaurant. Vous avez compris qu'il existait des restaurants « à thème »

Vignette clinique courte

Si vous pensiez...	..et qu'alors vous constatez	Votre hypothèse devient...
Que ce restaurant sert une cuisine américaine	Qu'il y a une bouteille de Ketchup sur la table	-2 -1 0 +1 +2
Que ce restaurant sert une cuisine française	Qu'il y a une nappe à carreaux sur la table	-2 -1 0 +1 +2
Que ce restaurant sert une cuisine italienne	Qu'il y a une bouteille d'huile d'olive sur la table	-2 -1 0 +1 +2
-2 : Absolument pas probable -1 : peu probable	0 : ni plus, ni moins probable	+1 : Probable +2 : Très probable
Hypothèses pertinentes	Informations supplémentaires	Echelle de Likert

Ancrage

Lecture

Test de Concordance des Scripts

Incertitude

Vous visitez un parc d'attractions au Japon. Vous n'êtes pas encore très familiarisé(e) avec les idéogrammes. Vous cherchez un restaurant. Vous avez compris qu'il existait des restaurants « à thème »

Vignette clinique courte

Si vous pensiez...	..et qu'alors vous constatez	Votre hypothèse devient...
Que ce restaurant sert une cuisine américaine	Qu'il y a une bouteille de Ketchup sur la table	-2 -1 0 +1 +2
Que ce restaurant sert une cuisine française	Qu'il y a une nappe à carreaux sur la table	-2 -1 0 +1 +2
Que ce restaurant sert une cuisine italienne	Qu'il y a une bouteille d'huile d'olive sur la table	-2 -1 0 +1 +2
-2 : Absolument pas probable -1 : peu probable	0 : ni plus, ni moins probable	+1 : Probable +2 : Très probable
Hypothèses pertinentes	Informations supplémentaires	Echelle de Likert

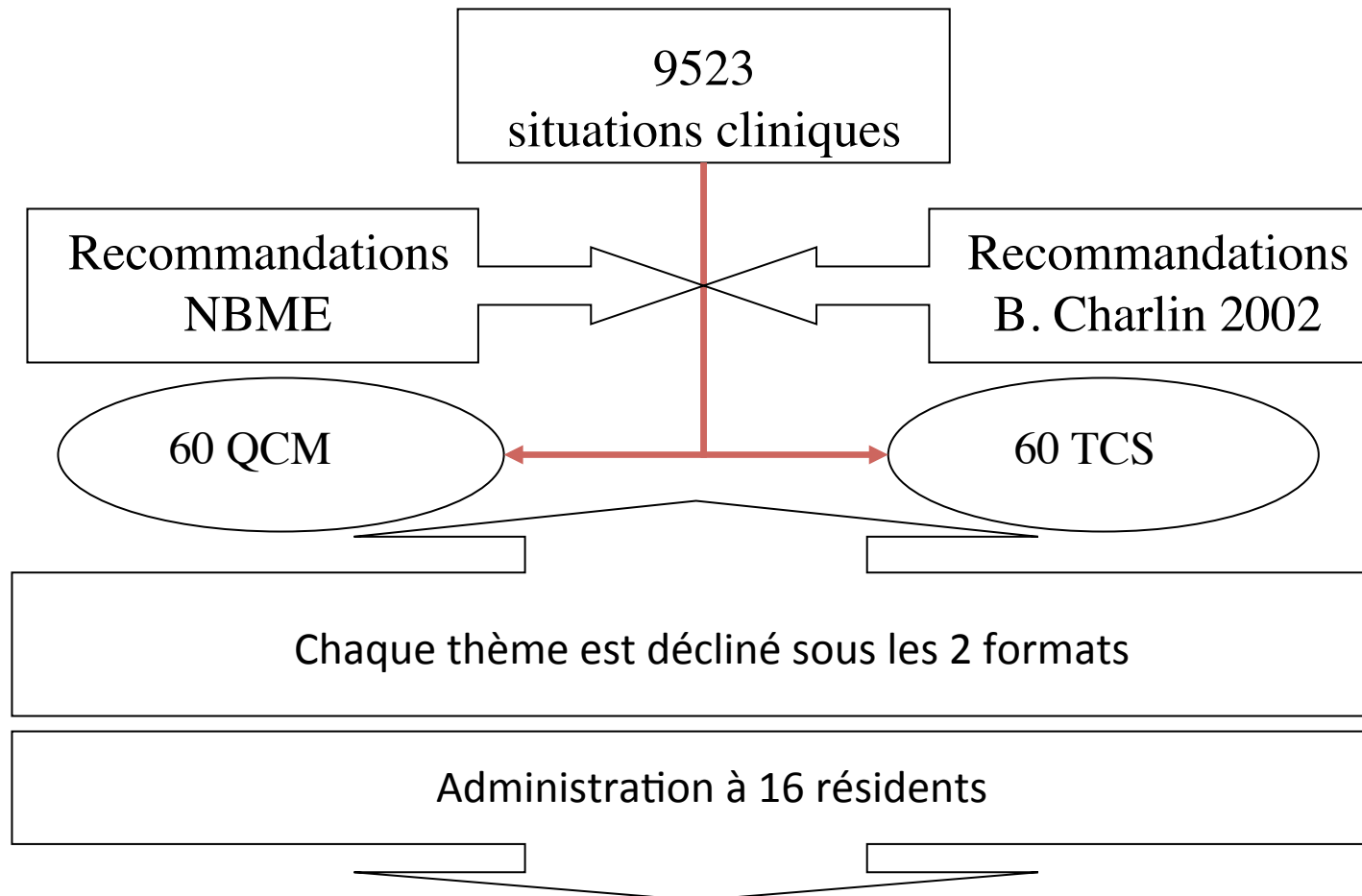
Ancrage

Lecture

Quelle est sa place parmi les autres outils d'évaluation ?

- ❖ Comparaison aux QCM ;
- ❖ Comparaison aux questions rédactionnelles format ECN.

TCS vs QCM à contexte riche



TCS vs QCM à contexte riche

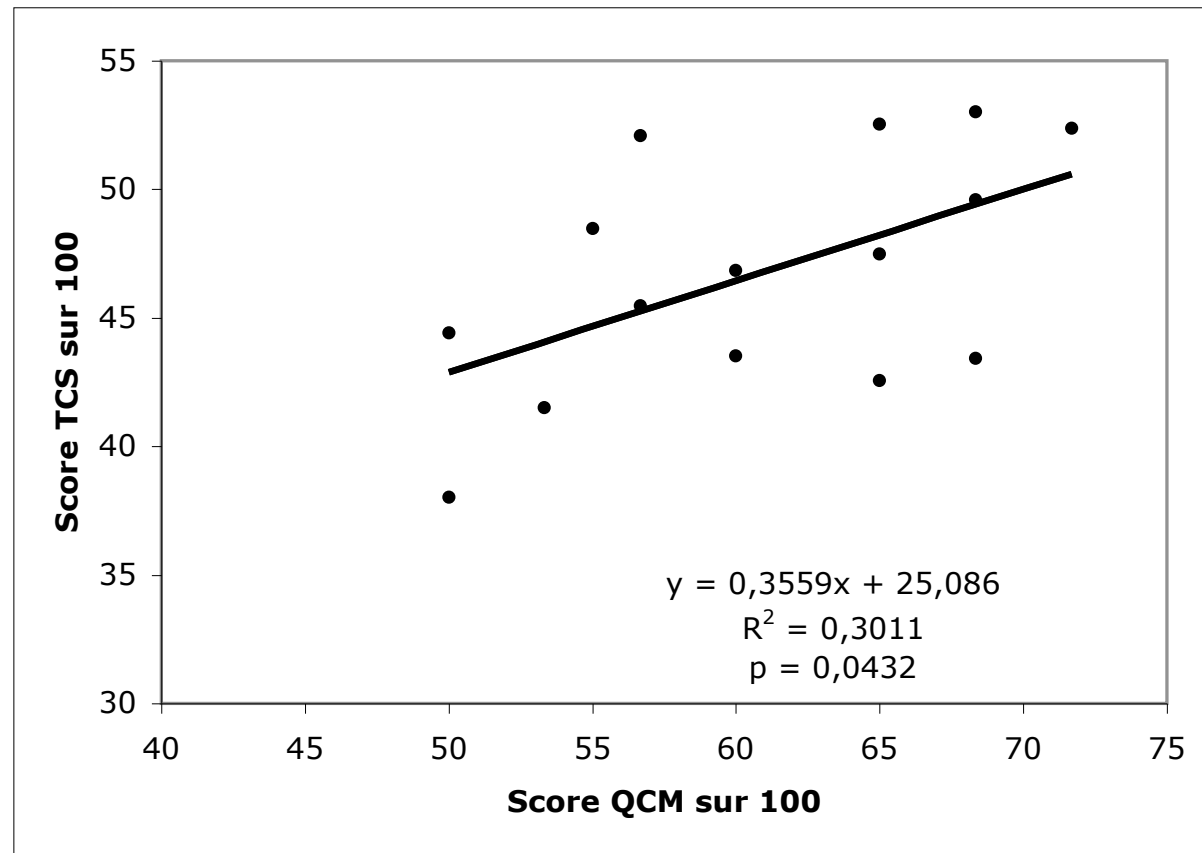
Une patiente de 17 ans consulte pour un « accident de préservatif ». Elle est au 12^{ème} jour de son cycle. Quelle est votre prescription de sortie ?

- A. Lévonorgestrel, éthynyl estradiol (Adépal[®])
- B. Lévonorgestrel (Norlévo[®])
- C. Consultation au planning familial
- D. Méthotrexate (Méthotrexate[®])
- E. Dosage des β -HCG urinaires

Une patiente consulte pour un « accident de préservatif ». Vous discutez le traitement.

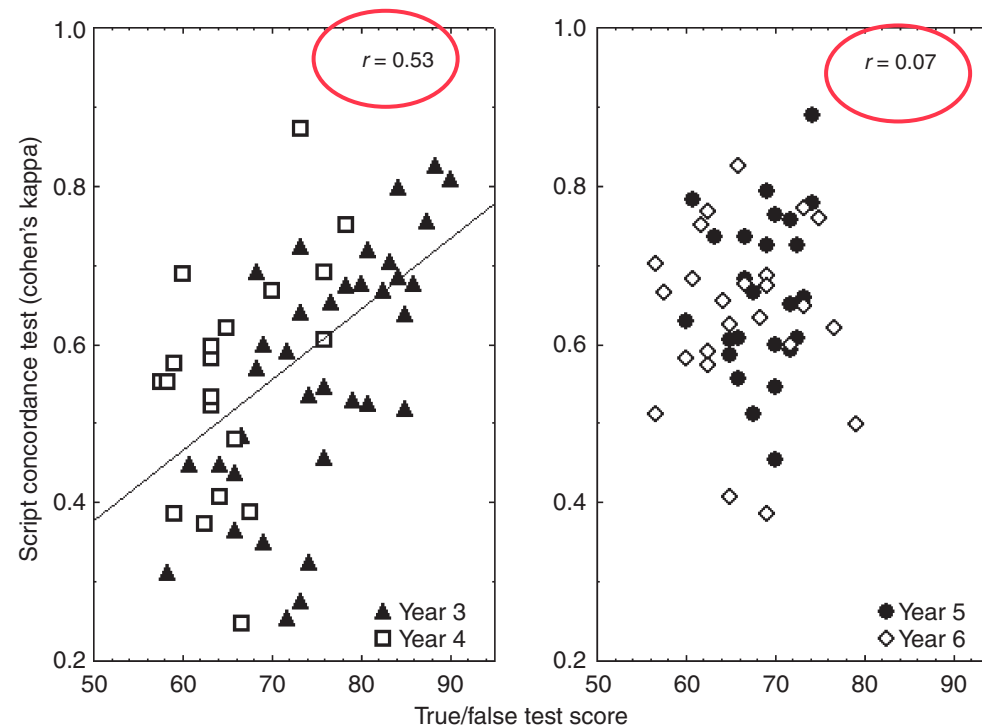
Si vous pensiez à	Et qu'alors vous trouvez	Cette nouvelle donnée rend le traitement proposé
Une prescription de lévonorgestrel (Norlévo [®])	Un « accident » survenu le 2 ^{ème} jour du cycle	-2 -1 0 +1 +2
Une prescription de lévonorgestrel (Norlévo [®])	Une patiente âgée de 15 ans	-2 -1 0 +1 +2
Une prescription de lévonorgestrel (Norlévo [®])	Des antécédents familiaux d'hypercholestérolémie	-2 -1 0 +1 +2
-2 : Totalelement contre indiqué -1 : Contre indiqué	0 : Ni plus, ni moins indiqué	+1 : Indiqué +2 : Formellement indiqué

TCS vs QCM à contexte riche

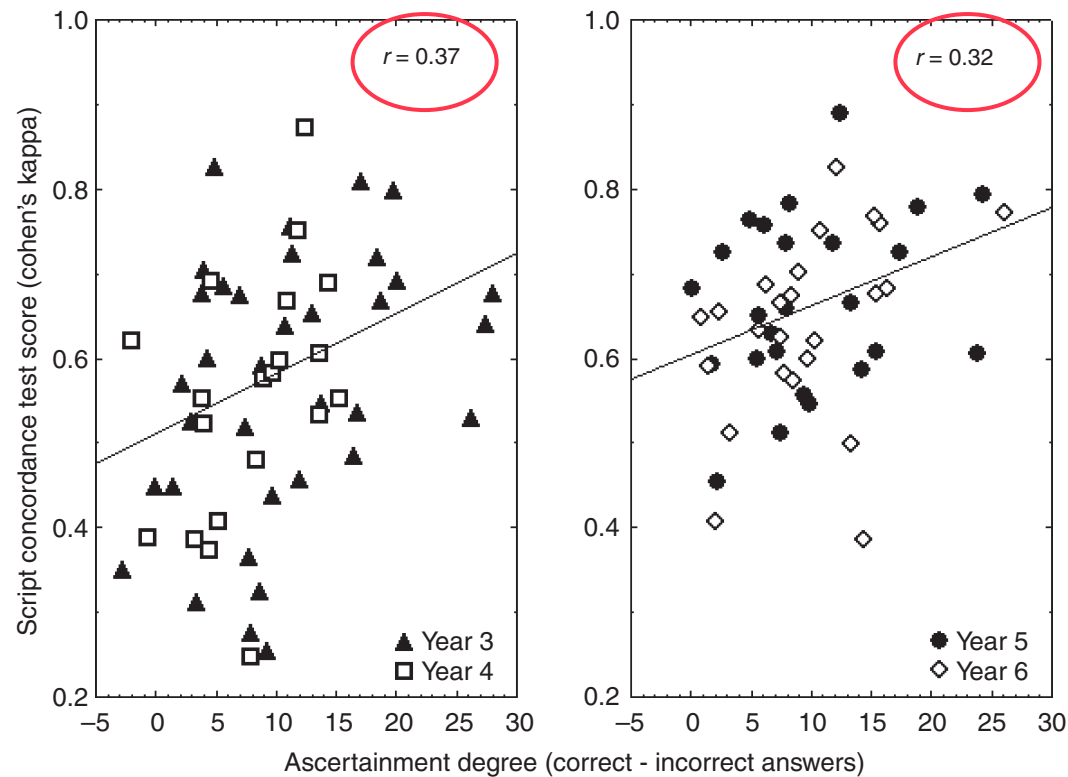


TCS vs QCM factuels

- ❖ 106 étudiants en médecine de 3^{ème} à 6^{ème} année ;
- ❖ Enseignement d'endocrinologie ;
- ❖ TCs vs QCM factuels (vrai / faux) :



TCS vs QCM factuels

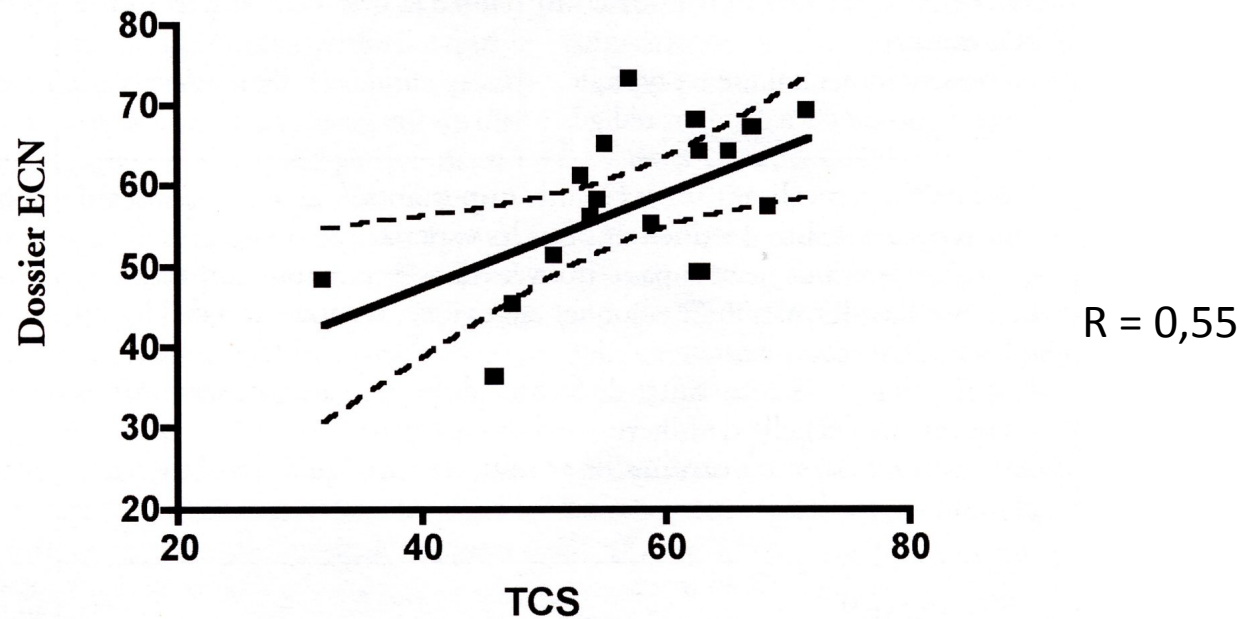


TCS vs dossier type ECN

- ❖ 18 étudiants de DCEM 2 à DCEM 4 en réanimation médicale ;
- ❖ Dossier type ECN ;
- ❖ TCS comportant 26 questions.

TCS vs dossier type ECN

Corrélation entre les scores obtenus par les étudiants au test de concordance de script (TCS) (abscisses) et au dossier de « type épreuves classantes nationales » (ECN) (ordonnées)



$R = 0,55$



L'examen

Aspects techniques pratiques

- ❖ Table de spécification (*blueprint*) ;
- ❖ Rédaction de scénarii et questions ;
- ❖ Ancrages ;
- ❖ Panel d'experts ;
- ❖ Contrôle qualité.

Table de spécification

4 éléments :

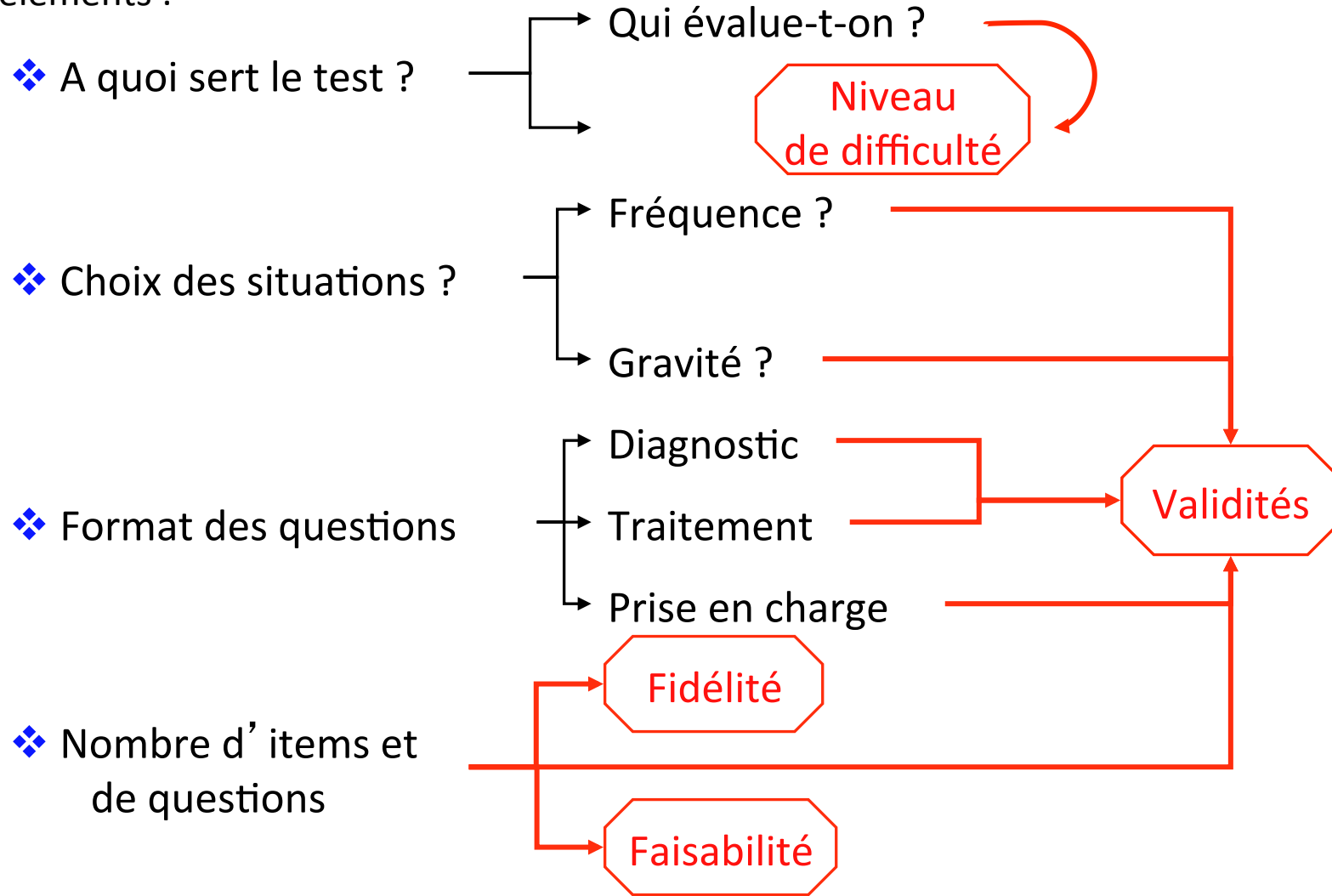


Table de spécification

	Cardio	Pneumo	Neuro	Gastro	Métabol.	Rhumato	Gériatr.	Gynéco	Psy	Trauma
Dg	X	X	X	X		X	X	X	X	X
Trt	X	X	X	X	X				X	
Prise en charge	X	X	X							

	Cardio	Pneumo	Neuro	Gastro	Métabol.	Rhumato	Gériatr.	Gynéco	Psy	Trauma
Dg	Douleur thoracique	Dyspnée aiguë	Céphalée	Douleur abdominale		Mono arthrite	AEG	GEU	Confusion	Trauma vertébral
Trt	Insuffisance cardiaque	Embolie pulmonaire	Crise comitiale	Hémorragie digestive	Deshydr.				Agitation	
Prise en charge	Synd. coronarien aigu	BPCO en poussée	AVC							

Table de spécification et construction de l'examen

En résumé :

❖ Architecture de l'examen : **4 impératifs** ;

❖ Sélection des situations ;

❖ Sélection des hypothèses pertinentes ;

❖ Sélection des informations supplémentaires pertinentes ;

→ Fixe le niveau d'incertitude.

❖ 1 – 2 concepteurs / rédacteurs
❖ Experts

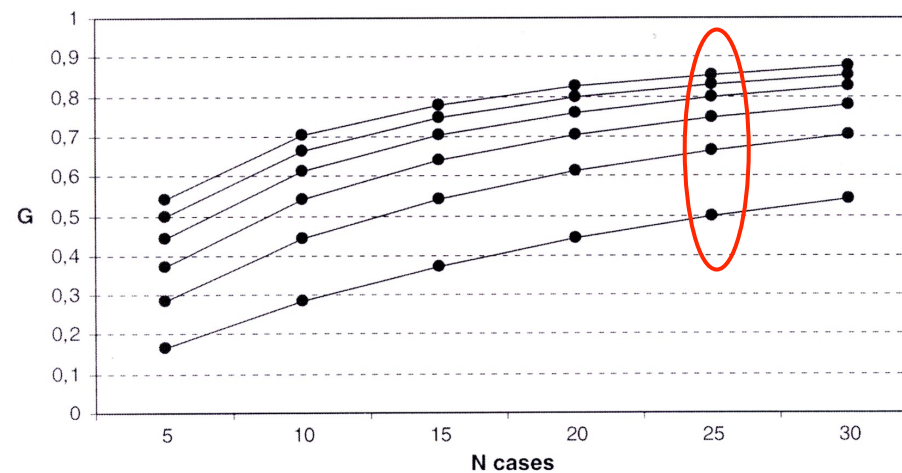
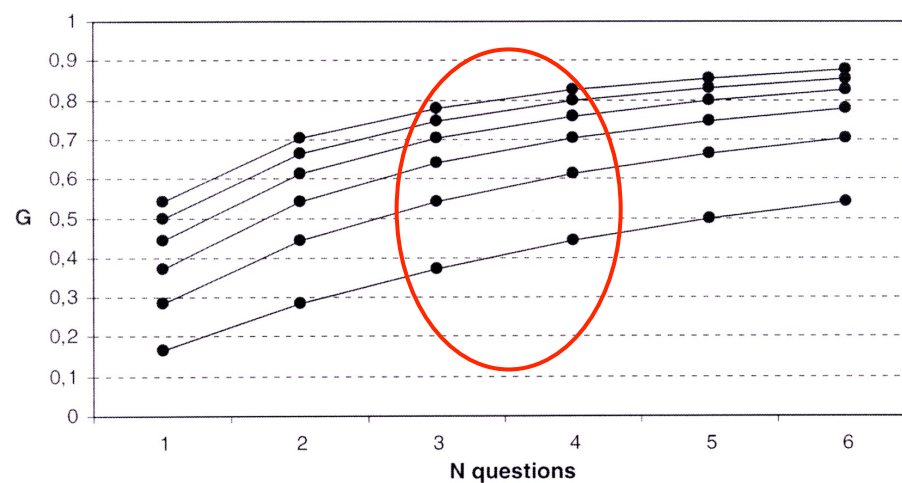
Niveau d'incertitude

- ❖ Pas de piège ;
- ❖ Pas de consensus, guidelines ;
- ❖ Plusieurs options possibles ;
- ❖ Données incomplètes ;
- ❖ Rapport bénéfice / risque, coût / efficacité.

Combien de cas et de questions ?

- ❖ Notion de spécificité de cas (*case specificity*) ;
- ❖ Nombreux cas avec 1 question ?
- ❖ 1 cas avec n questions ?
- ❖ Combien de cas /questions pour une fidélité acceptable ?

Combien de cas et de questions ?



Combien de cas et de questions ?

Table 3 D study of reliability by number of questions and cases

Number of questions per case	Number of cases	Total number of questions	Approximate testing time in minutes	G coefficient		
				Oncology	Pediatrics	Nursing
1	30	30	25	0.72	0.54	0.70
2	25	50	40	0.81	0.66	0.79
2	30	60	50	0.83	0.70	0.82
3	20	60	50	0.83	0.70	0.82
3	25	75	60	0.86	0.75	0.85
4	15	60	50	0.83	0.70	0.81
4	20	80	65	0.87	0.76	0.85
5	12	60	50	0.83	0.69	0.81
5	15	75	60	0.86	0.74	0.84

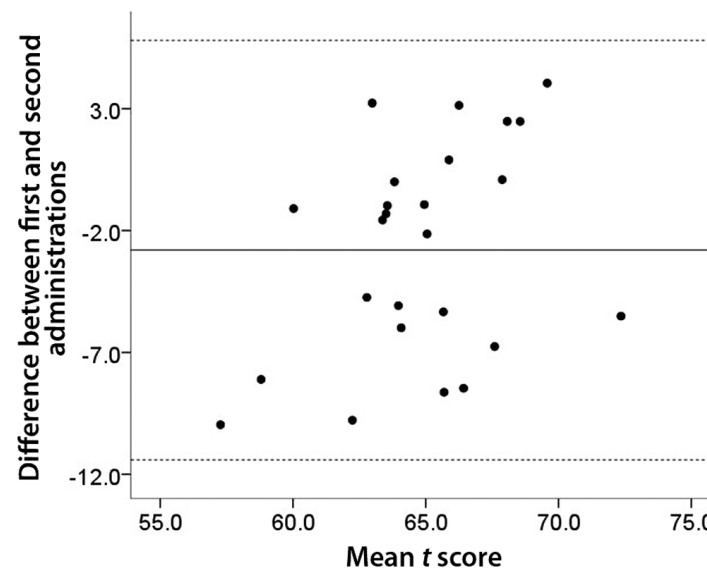
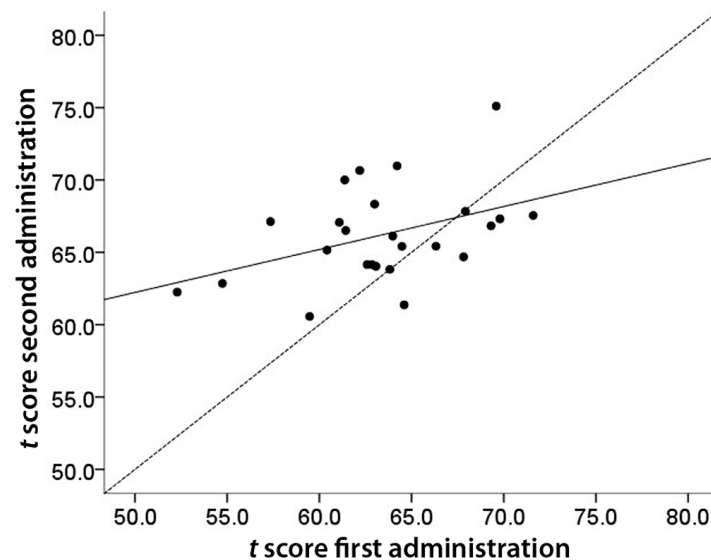
Conserver un format homogène +++++

Un homme de 56 ans est adressé par les pompiers pour malaise sur la voie publique. Il a été découvert inanimé par des passants. Il reprend peu à peu conscience. En dehors d'une désorientation temporo spatiale et d'une agitation, son examen neurologique est sans particularité. Les constantes vitales sont normales. Vous planifiez sa prise en charge.

<i>Si vous pensiez prescrire :</i>	<i>Et qu'alors vous trouvez :</i>	<i>L'effet sur la nécessité de demander ce test est le suivant :</i>
Un dosage de lactates veineux	Une intervention des pompiers environ 2 heures avant votre examen	-2 -1 0 +1 +2
Un EEG	Une morsure du bord latéral de la langue	-2 -1 0 +1 +2
Une injection de 1 mg de clonazépam (Rivotril®)	Une morsure du bord latéral de la langue	-2 -1 0 +1 +2
-1 : Inutile -2 : Tout à fait inutile	0 : Ni plus, ni moins utile	+1 : Plus utile +2 : Très utile

Fidélité du test

- ❖ Cohérence interne du test ;
- ❖ Mesurée par le coefficient α de Cronbach ($\geq 0,80$) ;
- ❖ Habituellement atteint avec 25 cas, 75 questions, et 60 minutes de composition.



Rédaction de scénarii et questions

Scénario : informations nécessaires à la description du problème :

- ❖ Âge et sexe ;
- ❖ Motif de consultation ou d' admission ;
- ❖ Contexte : urgences, cabinet de consultation, hospitalisation,...
- ❖ Données cliniques,...

1 à 2
rédacteurs

Rédaction de scénarii et questions

Sélection des hypothèses *pertinentes* :

- ❖ Pertinentes dans le contexte (vignette clinique) ;
- ❖ Pas de piège.

1 à 2
rédacteurs

Rédaction de scénarii et questions

Informations supplémentaires *pertinentes* :

- ❖ Donnée d'examen clinique ;
- ❖ Résultat biologique ;
- ❖ Imagerie statique ;
- ❖ Vidéo,...

1 à 2
rédacteurs

Rédaction de scénarii et questions

Rédaction des questions et informations supplémentaires :

- ❖ Balayer l'ensemble des points de l'échelle de Likert ;



Limitation des choix
au hasard

- ❖ Niveau de variabilité équilibrée ;



Discrimination

- ❖ Chaque question est indépendante ;

- ❖ Pas d'hypothèse ou d'information commune d'une question à l'autre ;

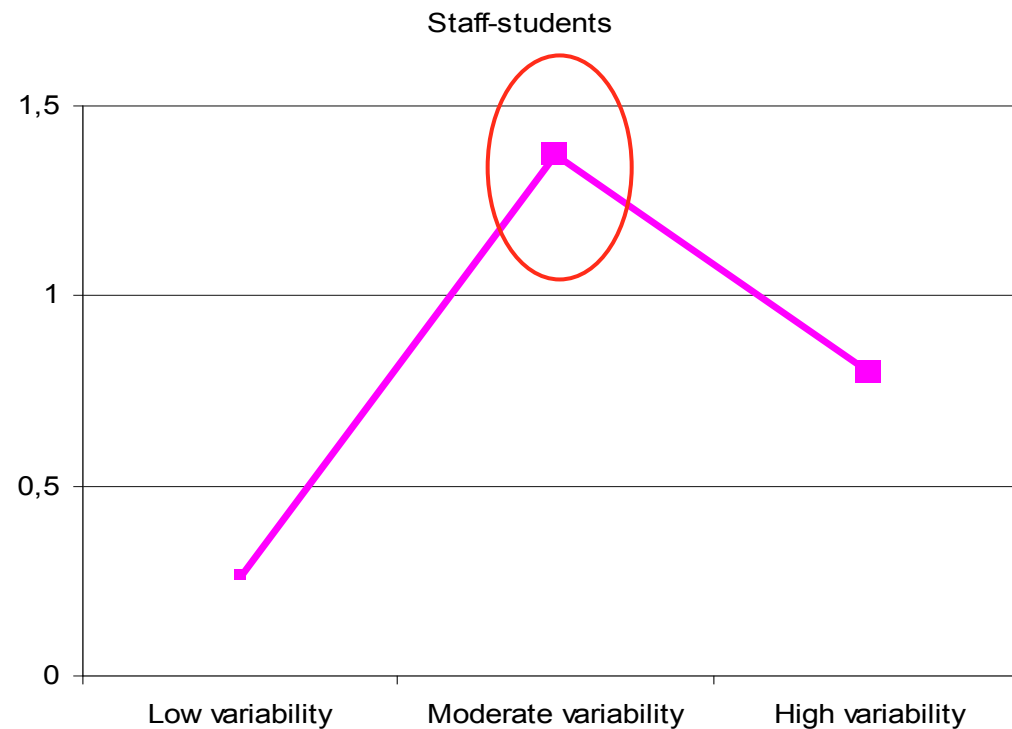


Limitation du
cueing effect

- ❖ Format homogène des questions ;

- ❖ Intérêt de questions à consensus ?

Impact de la variabilité des questions sur leur pouvoir discriminant



Qu'est-ce qu'une question *variablement* correcte ?

❖ Question correcte :

Réponses attendues (et observées) réparties autour de la valeur modale (1 à 2 gradations de l'échelle de Likert) ;

❖ Questions discordantes :

- ✓ Valeurs bimodales ;
- ✓ Réparties sur toute l'échelle ;
- ✓ Centrées aux 2 extrémités.

Échelle de Likert - Ancrages

Recommandations :

- ❖ Préciser la signification des différents points pour éviter toute ambiguïté, **en particulier le 0** ;
- ❖ Garder la même présentation.

-1 : Moins probable -2 : Très peu probable	0 : Ni plus, ni moins probable	+1 : Plus probable +2 : Très probable
---	-----------------------------------	--

Échelle de Likert - Ancrages

Échelles :

❖ Ancrage pour investigation (utilité) :

Si vous pensiez prescrire...	Et qu'alors vous trouvez	Cette nouvelle information rend l'examen
Option d'investigation	Nouvelle information	-2 -1 0 +1 +2
-1 : Inutile -2 : Complètement inutile	0 : Ni plus, ni moins utile	+1 : Utile +2 : Très utile

Impact négatif

Pas de prescription

❖ Ancrage pour investigation (risque bénéfice) :

Si vous pensiez prescrire...	Et qu'alors vous trouvez	Cette nouvelle information rend l'examen
Option d'investigation	Nouvelle information	-2 -1 0 +1 +2
-1 : Contre indiqué -2 : Totalement contre indiqué	0 : Ni plus, ni moins indiqué	+1 : Indiqué +2 : Fortement indiqué

Prescription

Impact positif

Échelle de Likert - Ancrages

Échelles :

❖ Ancrage pour traitement (utilité) :

Impact négatif



Pas de prescription

❖ Ancrage pour traitement (risque bénéfice) :

Pas d'impact

Prescription



Impact positif

Si vous pensiez prescrire...	Et qu'alors vous trouvez	Cette nouvelle information rend le traitement
Option de traitement	Nouvelle information	-2 -1 0 +1 +2
-1 : Inutile -2 : Complètement inutile	0 : Ni plus, ni moins utile	+1 : Utile +2 : Très utile

Si vous pensiez prescrire...	Et qu'alors vous trouvez	Cette nouvelle information rend le traitement
Option de traitement	Nouvelle information	-2 -1 0 +1 +2
-1 : Contre indiqué -2 : Complètement contre indiqué	0 : Ni plus, ni moins indiqué	+1 : Indiqué +2 : Très indiqué

Modifications des échelles de Likert

❖ Format unique :

Si vous pensiez prescrire...	Et qu'alors vous trouvez	L'impact de cette nouvelle information sur l'hypothèse est :
Option d'investigation	Nouvelle information	-2 -1 0 +1 +2
-1 : Négatif -2 : Totalelement négatif	0 : Ni plus, ni moins positif	+1 : Positif +2 : Très positif

- ❖ Garantir le balayage d l'échelle de Likert :
 - ✓ Remplacer les valeurs extrêmes (-2, +2) par des intitulés moins catégoriques :
 - ✓ - 2 : éliminée ou presque éliminée ;
 - ✓ +2 : certaine ou presque certaine.

Ce qui change d'un format à l'autre...

Vous visitez un parc d'attractions au Japon. Vous n'êtes pas encore très familiarisé(e) avec les idéogrammes. Vous cherchez un restaurant. Vous avez compris qu'il existait des restaurants « à thème »

Intitulé des 3 colonnes

Si vous pensiez...	..et qu'alors vous constatez	Votre hypothèse devient...				
Que ce restaurant sert une cuisine américaine	Qu'il y a une bouteille de Ketchup sur la table	-2	-1	0	+1	+2
Que ce restaurant sert une cuisine française	Qu'il y a une nappe à carreaux sur la table	-2	-1	0	+1	+2
Que ce restaurant sert une cuisine italienne	Qu'il y a une bouteille d'huile d'olive sur la table	-2	-1	0	+1	+2
-1 : Improbable -2 : Très improbable	0 : Ni plus, ni moins probable	+1 : Probable +2 : Très probable				

Signification des ancrages

Panels d' experts

Rôles :

2 groupes
différents

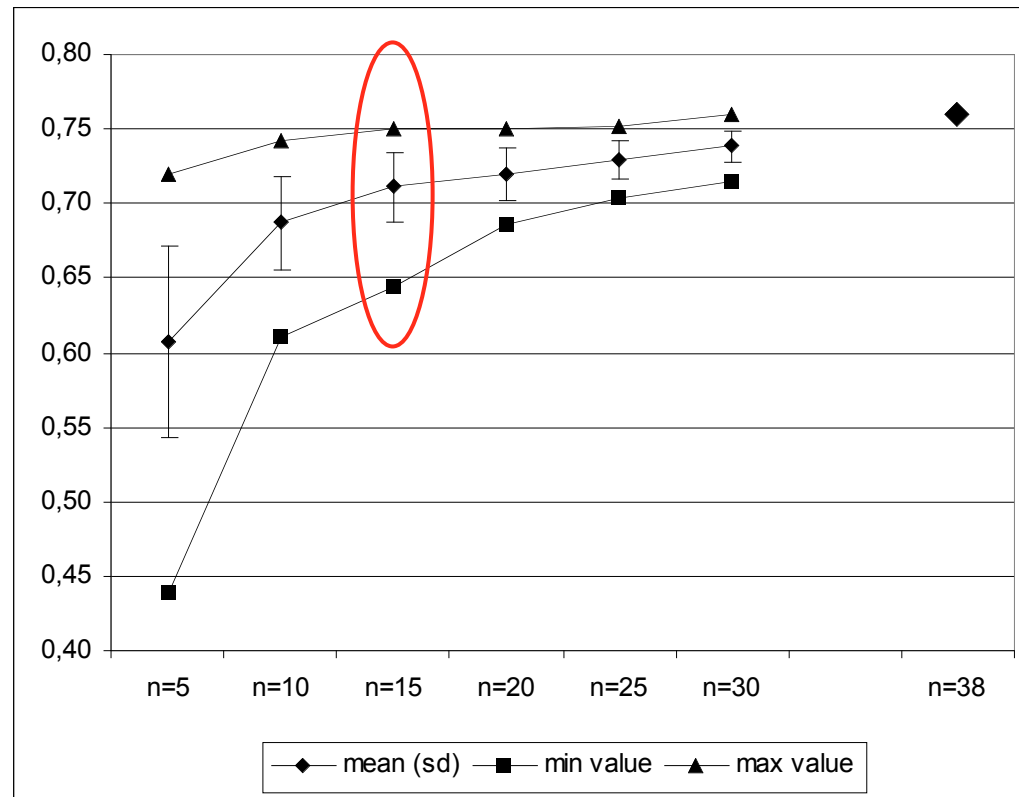
- ❖ Préparation de l' examen ;
- ❖ Validation de scénarii et questions ;
- ❖ Établissement des scores.

- ❖ Focaliser sur les points importants ;
- ❖ Vérification du niveau de difficulté (M3) ;
- ❖ Ancrages ;
- ❖ Balayer l'ensemble des points de l'échelle de Likert ;
- ❖ Indépendance des questions ;
- ❖ Niveau de variabilité équilibré.

Panel d'experts

- ❖ Experts dans le domaine clinique exploré :
Multiples critères de choix : diplôme, nombre d'années d'exercice, membre de collège,...
- ❖ 10 à 20 minimum.

Fidélité moyenne en fonction de la taille du panel d'experts



Composition du panel d'experts

Mais :

- ❖ Augmentation des scores avec le nombre d'experts ;
- ❖ Augmentation des scores si enseignants ;
- ❖ Augmentation des scores si même culture ;
- ❖ Mais concordance parfaite des scores obtenus ($R = 0,98$).

Impact de la composition du panel d'experts sur les scores en cas de TCS multi disciplinaire ?

- ❖ Prise de décision per opératoire par des résidents de niveau d'expertise croissant (R1 à R5) ;
- ❖ Panel de jeunes (diplôme ≤ 5 ans) chirurgiens *généralistes* :

Table 1
Examination scores (62-item test) by resident level

Resident level	n	Mean score, %	SD, %
R1	8	52.5	9.96
R2	6	62.5	5.12
R3	9	68.3	9.19
R4	6	75.7	9.61
R5	7	68.0	6.44

Impact de la composition du panel d'experts sur les scores en cas de TCS multi disciplinaire ?

- ❖ Duggan (2012) : panel multi spécialistes : 158 questions : Cronbach α : 0,62 (5^{ème} année) à 0,78 (6^{ème} année) ;
- ❖ Humbert (2011) : panel interniste : 75 questions : Cronbach α : 0,73 ;
- ❖ Gagnon (2011) : Pas d'impact de l'optimisation du panel sur le Cronbach α .

Impact de la composition du panel d'experts sur les scores en cas de TCS multi disciplinaire ?

1 groupe d'experts par champ (discipline) :

	Experts / champ	Items / questions	Score α
Duggan	10 – 16	51 / 158	0,62 et 0,78
Ramaekers	28 en tout	30 / 120	0,79 et 0,80
Petrucci	6	53 / 153	0,67 → 0,81

D'après

Duggan P. *BMC Med Educ* 2012

Ramaekers S. *Assessment Evaluation Higher Education* 2012

Petrucci A. *Am J Surg* 2012

Séminaire TCS
Nancy, 07 janvier 2016

Impact de la composition du panel d'experts sur les scores en cas de TCS multi disciplinaire ?

1 groupe d'experts par champ (discipline) :

	Experts	Etudiants	Items / Questions	Score α
Nouh	22	202	53 / 153	0,81 → 0,89
Petrucci	6 x 5	202 + 25	53 / 153	0,61 → 0,81

Établissement des scores

Correction : Établissement des scores :

❖ Choix de 10 experts :

		-2	-1	0	+1	+2
Unit 1	Question 1	7	1	2	0	0

❖ Transformation modale du choix des 10 experts :

		-2	-1	0	+1	+2
Unit 1	Question 1	1	0.14	0.28	0	0

❖ Choix des 8 résidents :

		R 1	R 2	R 3	R 4	R 5	R 6	R 7	R 8
Unit 1	Question 1	- 1	- 1	- 1	2	1	0	- 2	0

❖ Calcul des scores des 8 résidents :

		R1	R2	R3	R4	R5	R6	R7	R8
Unit 1	Question 1	0.14	0.14	0.14	0	0	0.28	1	0.28

Plus de signification aux scores ?

- ❖ Etablissement des scores individuels des membres du panel d'experts ;
- ❖ Transformation du score des étudiants en *t-score* :

$$t\text{-score}_{\text{étudiants}} = [(SCT_{\text{étudiants}} - \text{mean } SCT_{\text{experts}}) / SD \text{ } SCT_{\text{experts}}] \times 5 + 80$$

80 : score max habituellement rencontré chez les experts

5 : SD habituellement rencontrée chez les experts

Au total : qui fait quoi ?

Planificateur(s)	❖ Table de spécification ❖ Situations cliniques ❖ Hypothèses ❖ Informations supplémentaires ❖ Établissement des scores des étudiants ❖ Décision de passage
Experts 1	❖ Situations cliniques ❖ Hypothèses ❖ Informations supplémentaires
Experts 2	❖ Établissement des scores des étudiants

Les erreurs à ne pas faire...

Erreur	Conséquence	Prévention	Correction
Trop spécialisé	Inadapté	❖ Programme CNESER 2013 ❖ Table de spécification	Relecture
Focaliser sur une partie de l'échelle de Likert	Choix au hasard	❖ Choix de situations/hypothèses/informations pertinent ❖ Répartir les items dans l'examen	Relecture
Niveau d'incertitude inadapté	QCM ou bruit de fond	❖ Choix de situations/hypothèses/informations pertinent	Relecture
Questions non indépendantes	<i>Cueing effect</i>	❖ Choix de situations/hypothèses/informations pertinent	Relecture
Informations communes d'une question à l'autre	<i>Cueing effect</i>	❖ Choix de situations/hypothèses/informations pertinent	Relecture
Format hétérogène des questions	<i>Cueing effect</i>	❖ Choix de situations/hypothèses/informations pertinent	Relecture

Contrôle qualité *a priori*

Vignette	❖ Décrit une situation problématique, même pour les experts ❖ Décrit une situation clinique adaptée au niveau de compétence des étudiants ❖ Question et contexte sont explicites ❖ Présentation clinique typique ❖ Vignette correctement rédigée	☐ Oui ☐ Non ☐ Oui ☐ Non ☐ Oui ☐ Non ☐ Oui ☐ Non ☐ Oui ☐ Non
Questions	❖ Questions focalisant sur les points importants ❖ Hypothèses pertinentes ❖ 2 options identiques non rencontrées dans 2 questions consécutives ❖ Information supplémentaire pertinente / hypothèse ❖ Ancrage explicite, en particulier la valeur neutre (0) ❖ Questions construites pour balayer les 5 points de l'échelle de Lickert ❖ Questions construites pour obtenir une variabilité moyenne	☐ Oui ☐ Non ☐ Oui ☐ Non ☐ Oui ☐ Non ☐ Oui ☐ Non ☐ Oui ☐ Non ☐ Oui ☐ Non ☐ Oui ☐ Non
Panel d'experts	❖ Entre 10 et 20 ❖ Professionnels dont la présence dans le panel est licite compte tenu du niveau de compétence des étudiants ❖ Les professionnels passent le test dans les mêmes conditions que les étudiants	☐ Oui ☐ Non ☐ Oui ☐ Non ☐ Oui ☐ Non

Contrôle qualité *a posteriori*

Analyse d'items :

- ❖ Indices de difficulté ;
- ❖ Indice de discrimination ;
- ❖ Fidélité globale du test (score de Cronbach) ;
- ❖ Impact de chaque item sur la fidélité du test ;
- ❖ Réponse « déviante ».

Prévoir plus
de cas pour
optimiser le
test (+ 20%)

www.cme.umontreal.ca/tcs

Que faire des experts « déviants » ?

Table 2 Comparison of optimization methods with panel of 45

	Alpha value	Correlation with original data	Effect size		
			Students-residents	Students-panel	Residents-panel
No exclusion	0.86		2.08	2.57	0.48
Removal of members	0.86	0.99	2.05	2.44	0.45
Distance-from-mode					
2 Anchors	0.86	0.99	2.06	2.53	0.47
1 Anchor	0.86	0.99	2.08	2.52	0.46
Judgment-by-experts	0.86	0.99	1.99	2.47	0.43

Table 3 Comparison of optimization methods on panels of 15 (20 samples of 15 out of 45)

	Mean alpha value (SD)	Mean correlation with original data (SD)	Mean effect size (SD)		
			Students-residents	Students-panel	Residents-panel
No exclusion	0.86 (0.01)	0.99 (0.01)	1.97(0.12)	3.44 (0.70)	1.18 (0.33)
Removal of members	0.86 (0.01)	0.98 (0.01)	2.08 (0.11)	3.07 (0.74)	0.88 (0.32)
Distance-from-mode					
2 Anchors	0.86 (0.01)	0.98 (0.01)	1.93 (0.11)	3.16 (0.75)	1.05 (0.37)
1 Anchor	0.85 (0.01)	0.98 (0.01)	2.01(0.12)	2.32 (0.17)	0.37 (0.07)
Judgment-by-experts	0.85 (0.01)	0.98 (0.01)	2.03 (0.12)	3.73 (0.53)	1.24 (0.26)

- ❖ Si recherche de réponses « acceptables » (validité apparente : « élimination » de réponses plutôt que d'experts ;
- ❖ *Distance from mode 1* ou 2 la plus pratique.

Une patiente de 86 ans est adressée aux urgences par sa maison de retraite pour « altération de l'état général ». Elle « décline » depuis 48 heures. Elle a comme principaux antécédents une démence type Alzheimer, une HTA, et un DNID. La température auriculaire est à 38,6°C. La saturation transcutanée en O₂ est à 92 p. cent (les ambulanciers l'ont transportée sous O₂). Elle est somnolente, sans signe de localisation. La pression artérielle est à 110-70 mm Hg, et la fréquence cardiaque est à 110/min. Vous recherchez l'origine de cette hyperthermie.

<i>Si vous pensiez au diagnostic suivant :</i>	<i>Et qu'alors vous trouvez :</i>	<i>L'hypothèse devient :</i>
Une pyélonéphrite aiguë	BU : leucocytes : ++, sang : ++, protéines : ++, nitrites : 0, glucose : +++, corps cétoniques : ++	-2 -1 0 +1 +2
Une pneumonie bactérienne	Un traitement par donépézil (Aricept [®]), halopéridol (Haldol [®]), méprobamate (Equanil [®]), indapamide (Fludex [®]), et glipizide (Glibénèse [®])	-2 -1 0 +1 +2
Une déshydratation globale	Pas de foyer infectieux cliniquement décelable	-2 -1 0 +1 +2
-1 : Improbable -2 : Très improbable	0 : Ni plus, ni moins probable	+1 : Probable +2 : Très probable

10 experts :

		-2	-1	0	+1	+2
Item 2	Question 1	6	3			1
	Question 2			4	5	1
	Question 3		2	5	3	

Une patiente de 86 ans est adressée aux urgences par sa maison de retraite pour « altération de l'état général ». Elle « décline » depuis 48 heures. Elle a comme principaux antécédents une démence type Alzheimer, une HTA, et un DNID. La température auriculaire est à 38,6°C. La saturation transcutanée en O₂ est à 92 p. cent (les ambulanciers l'ont transportée sous O₂). Elle est somnolente, sans signe de localisation. La pression artérielle est à 110-70 mm Hg, et la fréquence cardiaque est à 110/min. Vous recherchez l'origine de cette hyperthermie.

<i>Si vous pensiez au diagnostic suivant :</i>	<i>Et qu'alors vous trouvez :</i>	<i>L'hypothèse devient :</i>
Une pyélonéphrite aiguë	BU : leucocytes : ++, sang : ++, protéines : ++, nitrites : 0, glucose : +++, corps cétoniques : ++	-2 -1 0 +1 +2
Une pneumonie bactérienne	Un traitement par donépézil (Aricept [®]), halopéridol (Haldol [®]), méprobamate (Equanil [®]), indapamide (Fludex [®]), et glipizide (Glibénèse [®])	-2 -1 0 +1 +2
Une déshydratation globale	Pas de foyer infectieux cliniquement décelable	-2 -1 0 +1 +2
-1 : Improbable -2 : Très improbable	0 : Ni plus, ni moins probable	+1 : Probable +2 : Très probable

10 experts :

		-2	-1	0	+1	+2
Item 2	Question 1	6	3			1
	Question 2			4	5	1
	Question 3	2	1	4	2	1

1^{ère} étape

Sélection d'une situation significative, d'hypothèses pertinentes et d'informations supplémentaires relevantes :

- ❖ Situation significative : situation fréquente, exemplaire, éducative ou potentiellement sévère (sélection selon le système PUIGE par ex) ;
- ❖ Hypothèses pertinentes : hypothèses que générerait un clinicien expérimenté dans le contexte défini par la vignette clinique ;
- ❖ Informations relevantes : informations supplémentaires que le clinicien expérimenté aurait cherché dans le contexte défini par la vignette clinique, compte tenu des hypothèses générées ;

1^{ère} étape (2)

❖ Situation clinique : douleur thoracique aux urgences (item 228) ;

❖ 3 hypothèses :

- ✓ Syndrome coronarien aigu (item 334) ;
- ✓ Dissection aortique (item 228) ;
- ✓ Embolie pulmonaire (item 224) ;

❖ 3 informations supplémentaires :

- ✓ Réalisation d'un électro cardiogramme ;
- ✓ Réalisation d'un cliché thoracique ;
- ✓ Dosage des D-dimères ;

2^{ème} étape

Rédaction de la vignette clinique :

- ❖ 4 à 5 rubriques indispensables :
 - ✓ Age ;
 - ✓ Sexe,
 - ✓ Motif de recours aux soins ;
 - ✓ Lieu de la consultation ;
 - ✓ ± Données cliniques et / ou biologiques ± d'imagerie ;
- ❖ Hypothèses et informations supplémentaires compte-tenu du contexte défini par la vignette ;

2^{ème} étape (2)

Vous voyez aux urgences un homme de 65 ans qui consulte pour une douleur thoracique évoluant depuis une heure environ. Il a comme principaux antécédents une hypertension artérielle, un tabagisme non sevré et un cancer broncho-pulmonaire, actuellement en rémission.

La douleur est rétrosternale, irradiant vers le haut et l'arrière. Il n'a plus mal lors de la réalisation de l'ECG.

3^{ème} étape

Rédaction des hypothèses et des informations supplémentaires :

Si vous pensiez à...	Et que vous apprenez que	Votre hypothèse devient...				
Un syndrome coronarien aigu	L'ECG inscrit des troubles de repolarisation en antérieur	-2	-1	0	+1	+2
Une dissection aortique	Le cliché thoracique montre un élargissement du médiastin	-2	-1	0	+1	+2
Une embolie pulmonaire	Les D-dimères sont à 1500 µg/l	-2	-1	0	+1	+2
-1 : Improbable -2 : Très improbable	0 : Ni plus, ni moins probable	+1 : Probable +2 : Très probable				

4^{ème} étape

Gestion de l'incertitude (si nécessaire) :

Vous voyez aux urgences un homme de 65 ans qui consulte pour une douleur thoracique évoluant depuis une heure environ. Il a comme principaux antécédents une hypertension artérielle, un tabagisme non sevré et un cancer broncho-pulmonaire, actuellement en rémission. **Il est porteur d'un pacemaker.**

Le patient parle mal le français. La douleur est montrée au niveau de la poitrine et de la nuque.

Si vous pensiez à...	Et que vous apprenez que	Votre hypothèse devient...				
Un syndrome coronarien aigu	Le tracé ECG est électro entraîné	-2	-1	0	+1	+2
Une dissection aortique	Le cliché thoracique en decubitus montre un élargissement du médiastin	-2	-1	0	+1	+2
Une embolie pulmonaire	Les D-dimères sont à 1500 µg/l	-2	-1	0	+1	+2
-1 : Improbable -2 : Très improbable	0 : Ni plus, ni moins probable	+1 : Probable +2 : Très probable				

5^{ème} étape

❖ 3 questions homogènes ?

Oui : 3 questions diagnostiques ;
3 résultats d'investigations ;

❖ 3 questions indépendantes ?

Oui :

❖ Différents points de l'échelle de Likert balayés ?

Si vous pensiez à...	Et que vous apprenez que	Votre hypothèse devient...				
Un syndrome coronarien aigu	Le tracé ECG est électro entraîné	-2	-1	0	+1	+2
Une dissection aortique	Le cliché thoracique en decubitus montre un élargissement du médiastin	-2	-1	0	+1	+2
Une embolie pulmonaire	Les D-dimères sont à 1500 µg/l	-2	-1	0	+1	+2
-1 : Improbable -2 : Très improbable	0 : Ni plus, ni moins probable	+1 : Probable +2 : Très probable				

C'est prêt !

Vous voyez aux urgences un homme de 65 ans qui consulte pour une douleur thoracique évoluant depuis une heure environ. Il a comme principaux antécédents une hypertension artérielle, un tabagisme non sevré et un cancer broncho-pulmonaire, actuellement en rémission. Il est porteur d'un pacemaker.

Le patient parle mal le français. La douleur est montrée au niveau de la poitrine et de la nuque.

Si vous pensiez à...	Et que vous apprenez que	Votre hypothèse devient...				
Un syndrome coronarien aigu	Le tracé ECG est électro entraîné	-2	-1	0	+1	+2
Une dissection aortique	Le cliché thoracique en decubitus montre un élargissement du médiastin	-2	-1	0	+1	+2
Une embolie pulmonaire	Les D-dimères sont à 1500 µg/l	-2	-1	0	+1	+2
-1 : Improbable -2 : Très improbable	0 : Ni plus, ni moins probable	+1 : Probable +2 : Très probable				

C'est prêt !

Rédigé par :	
Vérifié par :	
Dernière vérificationle :	
Items du programme (CNESER 2013) :	228, 334, 224
Spécialité(s) du panel de référence :	Médecine d'urgence, médecine générale, pneumologie, cardiologie

Instructions pour les étudiants

- ❖ Lisez la vignette clinique ;
- ❖ Le test se lit à l'horizontale : chaque ligne se compose d'une hypothèse *et* d'une information supplémentaire cohérente avec cette hypothèse *et* d'une échelle de Likert ;
- ❖ Vous n'avez pas à juger les hypothèses ou les informations supplémentaires correspondantes. Elles ont été validées par des experts ;
- ❖ Vous devez juger l'impact de l'information supplémentaire sur l'hypothèse, que vous mesurez avec l'échelle de Likert ;
- ❖ Les différentes lignes sont indépendantes les unes des autres : il n'y a pas d'information commune d'une ligne à l'autre ;
- ❖ Vous ne devez pas essayer de déduire votre réponse à une ligne de votre réponse à la ligne précédente ;
- ❖ Les valeurs de l'échelle de Likert sont des *catégories* et non des *scores* ;
- ❖ Vous devez choisir **une** valeur de l'échelle de Likert par ligne (question) ;
- ❖ Les valeurs -2 et -1 correspondent à un impact négatif de l'information supplémentaire sur l'hypothèse initiale ; les valeurs +1 et +2 correspondent à un impact positif de l'information supplémentaire sur l'hypothèse initiale ;
- ❖ La valeur 0 correspond à l'absence d'impact (et non pas à un impact négatif) ;
- ❖ Vous devez donc choisir une seule de ces valeurs en fonction de l'impact de l'information supplémentaire sur l'hypothèse initiale que vous estimez ;
- ❖ Le test est construit de manière à ce que les différentes valeurs soient retenues ;
- ❖ Il n'y a pas de logique à toujours choisir certaines valeurs (-1 à +1 habituellement) ;
- ❖ Il n'y a pas de logique à choisir des valeurs au hasard.

Qu'attendent les étudiants ?

Bonjour Pr,

Beaucoup d'étudiants se sont posés la question des groupes de sous-colles ces dernières semaines.

Plusieurs personnes se sont demandés comment cela fonctionne et si les sous-colles d'adapteront au format iECN.

Nous avons donc reçu un mail du Pr E. concernant les sous-colles (concernant uniquement l'inscription et non le déroulement des sous-colles)

Les étudiants ont donc plusieurs questions a propos de ses sous colles, je vous les transmet :

1) Qu'est-ce que les sous-colles, comment ca se déroule, des changements dans leur déroulement est-il prévu en vue de la réforme iECN ?

2) Sachant qu'une partie des étudiants ne pourra pas intégrer de groupe de sous-colle, les étudiants se demandent si les "sujets" de sous-colle pourront être accessible sur la base de données local de sides. (la préoccupation majeur des étudiants étant la peur de ne pas trouver assez de support d'entraînement sur la nouvelle forme d'Ecn)

Je pense personnellement que la sous-colle a été créé non pas pour augmenter la quantité de dossier accessible mais pour fournir des explications sur les points compliqués de ces dossiers. (l'ancien ecn comprend des millions de dossiers accessibles). **J'ai entendu beaucoup d'étudiants qui souhaite s'inscrire à des sous-colles pour avoir des QCM et des TCS et non pas pour avoir des explications.**

J'espère que le mail n'est pas trop confus.

Merci d'avance, (et encore merci pour l'énergie que vous mettez dans cette réforme.)

X.

Rentabilité d'annales ?

Une patiente de 86 ans est adressée aux urgences par sa maison de retraite pour « altération de l'état général ». Elle « décline » depuis 48 heures. Elle a comme principaux antécédents une démence type Alzheimer, une HTA, et un DNID. La température auriculaire est à 38,6°C. La saturation transcutanée en O₂ est à 92 p. cent (les ambulanciers l'ont transportée sous O₂). Elle est somnolente, sans signe de localisation. La pression artérielle est à 110-70 mm Hg, et la fréquence cardiaque est à 110/min. Vous recherchez l'origine de cette hyperthermie.

<i>Si vous pensiez au diagnostic suivant :</i>	<i>Et qu'alors vous trouvez :</i>	<i>L'hypothèse devient :</i>
Une pyélonéphrite aiguë	BU : leucocytes : ++, sang : ++, protéines : ++, nitrites : 0, glucose : +++, corps cétoniques : ++	-2 -1 0 +1 +2
-1 : Improbable -2 : Très improbable	0 : Ni plus, ni moins probable	+1 : Plus probable +2 : Beaucoup plus probable

	-2	-1	0	+1	+2
Panel 1			8	2	
Scores Panel 1	0	0	1	0,25	0
Panel 2				6	4
Scores Panel 2	0	0	0	1	0,67
Panel 3			7	1	2
Scores Panel 3	0	0	1	0,14	0,28

Utilisation des TCS pour l'apprentissage ?

Vignette 1 :

Vous voyez aux urgences un homme de 78 ans pour dyspnée aiguë. Il a des antécédents de cardiopathie ischémique, d'hypertension artérielle, et une broncho-pneumopathie chronique obstructive. Il est assis sur son brancard. L'auscultation retrouve des râles bronchiques diffus. La température auriculaire est à 37,8°C.

Si vous pensiez à...	Et qu'alors vous apprenez que...	Cette nouvelle information rend le diagnostic...				
a- Une poussée d'insuffisance cardiaque congestive	L'électrocardiogramme inscrit une arythmie complète par fibrillation auriculaire	-2	-1	0	+1	+2
b- Une exacerbation de broncho-pneumopathie chronique obstructive	L'expectoration est plus abondante que d'habitude	-2	-1	0	+1	+2
c- Une embolie pulmonaire	On vient de lui découvrir un cancer de la prostate	-2	-1	0	+1	+2
-1 : Improbable		0 : Ni plus, ni moins probable			+1 : Probable	
-2 : Très improbable					+2 : Très probable	

Questions du programme :	355, 232, 230, 205, 224, 307
Propositions de réponses :	a : +1, +2 b : +1 c : +1

Commentaires : Les trois situations retenues correspondent à trois étiologies de dyspnée aiguë avec orthopnée. Les trois situations peuvent entraîner une hyperthermie, habituellement modérée en ce qui concerne l'embolie pulmonaire et l'insuffisance cardiaque congestive. La fibrillation auriculaire est un facteur significativement prédictif d'insuffisance cardiaque congestive chez un patient dyspnéique aux urgences¹. L'augmentation du volume de l'expectoration est l'un des trois critères classiques (avec l'aggravation de la dyspnée et des expectorations purulentes) d'exacerbation de BPCO d'origine infectieuse². Le cancer est un des facteurs de risque de la maladie thrombo-embolique veineuse. Il apparaît dans les scores de Wells³ ou de Genève⁴.

Pour aller plus loin :

1- Wang CS, Fitzgerald JM, Schulzer M, Mak E, Ayus NT. Does this dyspneic patient in the emergency department have congestive heart failure ? *JAMA* 2005;294:1944-1956

2- Anthonisen NR, Manfreda AJ, Warren CP, Hershfield ES, Harding GK, Nelson NA. Antibiotic therapy in exacerbations of chronic obstructive pulmonary disease. *Ann Intern Med* 1987;106:196-204

3- Wells PS, Anderson DR, Rodger M et al. Derivation of a simple clinical model to categorize patients probability of pulmonary embolism : increasing the model utility with the SimpliRED D-dimer. *Thromb Haemost* 2000;83:416-420

4- Le Gal G, Righini M, Roy PM, Sanchez O, Aujesky D ; Bounameaux H, Perrier A. Prediction of pulmonary embolism in the emergency department : the revised Geneva score. *Ann Intern Med* 2006;144:165-172

Le raisonnement médical : comment l'enseigner en stage ?

Tableau 3 : Principes d'enseignement au lit du malade issus des théories de psychologie cognitive sur l'organisation des connaissances	
OBJECTIF	QUELQUES MOYENS pour le précepteur
Organiser les connaissances	Utiliser des diagnostics prototypiques. Développer les étapes du raisonnement en faisant des liens explicites vers la physiopathologie. Créer des liens avec des situations cliniques antérieures.
Activer les connaissances	Favoriser une bonne représentation du problème clinique à l'aide de transformations sémantiques ou schémas physiopathologiques. Faire résumer, synthétiser fréquemment les données cliniques.
Collecter des informations cliniques pertinentes et discriminantes	Utiliser des hypothèses diagnostiques pour cibler la collection des données cliniques. Reconnaître les éléments discriminant diverses hypothèses diagnostiques. Ne pas valoriser l'exhaustivité pour elle-même. Faire justifier les données recueillies.
Intégrer les processus de raisonnement	Utiliser un contexte de raisonnement hypothético-déductif. Permettre une approche non analytique (<i>Pattern recognition</i>). Favoriser une approche mixte entre reconnaissance spontanée d'un tableau clinique (approche non-analytique) suivie d'une approche hypothético-déductive (approche analytique), ou <i>vice versa</i> . Utiliser des cas cliniques nouveaux pour les apprenants. Exposer les étudiants à plusieurs présentations (formes) d'un même problème ou diagnostic afin de faciliter le transfert des connaissances à des situations nouvelles.

- ❖ Visites
« didactiques »
- ❖ Interrogation des étudiants
- ❖ Identification du problème
- ❖ Résumé
- ❖ Hypothèses argumentées
- ❖ Raisonnement à voix haute
- ❖ Rappels
(physiopath, pharmacologie, anatomie,...)
- ❖ ARC

Merci de votre attention !

BMC Medical Informatics and Decision Making



Correspondence

Open Access

Script Concordance Tests: Guidelines for Construction

Jean Paul Fournier¹, Anne Demeester² and Bernard Charlin^{*3}

Address: ¹Département de Pédagogie Médicale, Faculté de Médecine de Nice Sophia Antipolis, France; ²École Régionale de Sages-femmes, Faculté de Médecine de Moncton, France and ³PAASS, direction de la recherche, Faculté de Médecine, Université de Montréal, CP 6128, Succursale centre-ville, Montréal, Québec, H3C 3J7, Canada

Email: Jean Paul Fournier - fournier.jp@chu-nice.fr; Anne Demeester - Anne.Demeester@mail-ap-bm.fr; Bernard Charlin^{*} - bernard.charlin@umontreal.ca

^{*} Corresponding author

Published: 6 May 2008

BMC Medical Informatics and Decision Making 2008, 8:18 doi:10.1186/1472-6947-8-18

This article is available from: <http://www.biomedcentral.com/1472-6947/8/18>

Received: 28 November 2007

Accepted: 6 May 2008

How to construct and implement script concordance tests: insights from a systematic review

Valérie Dory,^{1,2} Robert Gagnon,³ Dominique Vanpee^{2,4} & Bernard Charlin³

Med Educ 2012

Le test de concordance de script en 20 questions

Ped Med 2013

Twenty questions on script concordance tests

Didier GIET^{1,3}, Valérie MASSART¹, Robert GAGNON² et Bernard CHARLIN²

2013, 1–10, Early Online



AMEE GUIDE

Script concordance testing: From theory to practice: AMEE Guide No. 75

STUART LUBARSKY¹, VALÉRIE DORY², PAUL DUGGAN³, ROBERT GAGNON⁴ & BERNARD CHARLIN⁴

¹McGill University, Canada, ²Université catholique de Louvain, Belgium, ³University of Adelaide, Australia,

⁴University of Montreal, Canada



Séminaire TCS
Nancy, 07 janvier 2016