

# LE TEST DE CONCORDANCE DE SCRIPT

## ***Un outil pour évaluer une dimension supérieure de la compétence clinique en chirurgie pédiatrique***

Joël LECHEVALLIER, Agnès LIARD-ZMUDA, Louis SIBERT

### **Présentation du problème**

L'objectif de cette présentation est de proposer l'utilisation du test de concordance de script pour évaluer les compétences des candidats à l'épreuve écrite de validation du DESC de chirurgie pédiatrique.

### **Revue de la littérature**

Le raisonnement clinique est un processus hypothético-déductif (*Elstein and al. 1978*) caractérisé par la génération précoce d'hypothèses (orientées par la collection d'information) et la prise de décision (par l'utilisation d'information permettant de confirmer ou de rejeter les hypothèses émises). De nombreux auteurs ont émis l'hypothèse qu'en médecine clinique, un praticien expérimenté diffère d'un collègue moins expérimenté par la mise en œuvre de réseaux élaborés de connaissance adaptés aux tâches qu'il effectue de façon régulière (*Feltovich and al. 1984, Bordage and al. 1984, Custers and al. 1996, Veitz and al. 1997*). Ces réseaux appelés *scripts* (*Feltovich and al. 1984, Schmidt and al. 1990, Charlin and al. 2000*) sont organisés pour atteindre des objectifs au sein des tâches concernées : diagnostic, stratégie d'investigation ou thérapeutique. Ces scripts se constituent dès que les étudiants sont confrontés avec leur première situation clinique et sont ensuite développés, enrichis et affinés au cours de la vie clinique (*Schmidt and al. 1990, Charlin and al. 2000*). Cette théorie met en évidence que ***dans une situation clinique donnée, le clinicien fait appel à la connaissance mémorisée en rapport avec une situation clinique de même nature antérieurement vécue.*** La réactivation des connaissances à partir de la mémoire est ensuite utilisée dans un processus déductif pour activer la recherche d'information qui permettra la confirmation ou au contraire l'infirmité des hypothèses évoquées (*Charlin and al. 1994*). ***L'élaboration de la connaissance*** (qui relève de la constitution de nombreux liens entre les différentes composantes des connaissances acquises et des expériences vécues) ***est la clef de la compétence clinique*** (*Bordage and al. 1994*). Charlin and al. (1998) a conçu un instrument, le test de concordance de script, qui mesure le degré d'élaboration d'un schéma de diagnostic.

Les principes du test de concordance de script (TCS) ont été déjà rapportés (*Charlin and al. 2000*). Cet instrument a été testé en radiologie, en gynécologie obstétrique et en urologie. Ces études ont mis en évidence que les seniors, les résidents, les étudiants ont des capacités de réponse différentes avec des niveaux de réponse différents avec l'expérience et des différences significatives entre chaque groupe (*Charlin and al. 1998*). Ces différentes publications rappellent que le TCS a le double avantage d'être relativement simple à construire et d'être un outil de mesure fiable. Ces résultats apportent aussi de nouvelles modalités d'évaluation de la connaissance clinique. Beaucoup de compétences diagnostiques et thérapeutiques sont sous-évaluées par les enseignants. En effet, il est habituel d'évaluer les résidents et les chefs de clinique en appréciant de façon subjective leur performance par un examen oral sur un cas clinique isolé. La plupart de ces méthodes d'évaluation mesure une compétence clinique factuelle plus que l'organisation des connaissances. De nombreuses études ont montré que les évaluations objectives et structurées par des patients standardisés permettaient de détecter des différences dans la performance clinique entre des résidents en urologie sans différence en fonction de leur année d'étude (*Sibert and al. 1997, 2000*). Les limites de ce système d'évaluation sont bien connues en particulier les difficultés logistiques (préparation des candidats, préparation de l'équipe enseignante, locaux, préparation des patients simulés...) (*Cohen and al. 1990, Grand'maison and al 1992*). ***En fait, il n'y a bien entendu pas un seul moyen d'évaluation qui permet d'évaluer tous les niveaux de la compétence.*** L'utilisation complémentaire de différents moyens d'évaluation de la compétence clinique devrait en théorie être encouragée. Dans cette présentation, nous proposons de développer le TCS en chirurgie pédiatrique pour une évaluation de niveau supérieur dans le cadre de la validation du DESC.

### **Quelle est la structure d'un test de concordance de script ?**

Le test consiste à mettre les candidats en présence d'une série de problèmes cliniques en leur demandant de faire un diagnostic, de choisir des investigations ou de prendre des décisions thérapeutiques en fonction des éléments ou des informations qui sont proposées. Ce test place donc les candidats dans une situation clinique authentique mais sous une forme écrite au cours de laquelle ils doivent interpréter les informations pour prendre leurs décisions.

### **Les étapes successives pour préparer un TCS**

#### **a) Spécification des objectifs éducationnels majeurs**

Il faut donc dans cette étape sélectionner les problèmes cliniques fondamentaux éducationnels envisagés dans le cadre du DESC de chirurgie pédiatrique par exemple :

- Situation de consultation :
  - a) augmentation douloureuse d'une bourse chez un petit garçon ;
  - b) marche les pieds en dedans chez un enfant de 5 ans
  - c) ...
- Situation d'urgence :
  - a) douleur abdominale aiguë chez un enfant de 4 ans
  - b) Ostéomyélite aiguë
  - c) ...

### **b) Ecrire des questionnaires sous forme de vignettes courtes**

Les situations cliniques sont présentées sous forme de petites vignettes, chacune suivie par une succession de questions en rapport avec celles-ci. Le format de la vignette est différent en fonction de l'objectif d'évaluation que l'on s'est fixé (compétence diagnostique, choix d'examens complémentaires ou option thérapeutique). Chaque question se compose de trois parties :

- la 1<sup>ère</sup> partie comporte une hypothèse diagnostique, une investigation envisagée ou une option thérapeutique.
- La 2<sup>ème</sup> partie apporte une nouvelle information (par exemple une information clinique, une imagerie, un test de laboratoire) qui pourrait avoir une influence sur l'hypothèse qui a été préalablement émise (voir exemple en appendice).
- La 3<sup>ème</sup> partie est une échelle de Likert en 5 points. Les vignettes successives sont complètement indépendantes les unes des autres et les options ou hypothèses peuvent changer d'une question à l'autre de telle sorte qu'il est habituellement relativement aisé de constituer un questionnaire pouvant comporter jusqu'à une centaine d'items.

### **Modalités de scorage**

*Le scorage est l'un des autres éléments très originaux du TCF.* On considère que la réponse donnée à chaque question pour un clinicien donné n'est jamais de type binaire de sorte que, pour une question donnée, on peut considérer que deux cliniciens expérimentés n'apporteront pas exactement la même réponse. Il est donc plus pertinent de demander à un panel de praticiens de niveau d'expérience différent de répondre à une question donnée (par exemple

on peut choisir un panel constitué de 4 PU-PH, 4 PH et 2 CCA en choisissant dans chaque catégorie 2 praticiens exerçant principalement en chirurgie viscérale et 2 autres exerçant principalement en chirurgie orthopédique).

Pour chaque item, chaque réponse se voit affecter une valeur correspondant à la proportion d'experts qui ont choisi la réponse donnée.

Par exemple, à une question donnée

- 10 % des experts ont répondu A,
- 30 % des experts ont répondu B,
- 60 % des experts ont répondu C,
- aucun n'a répondu D ou E.

Le crédit sera alors attribué au prorata des réponses données en phase avec le groupe correspondant du panel:

- l'étudiant qui répond A obtiendra 0,1 point (il a répondu comme 10% du panel),
- l'étudiant ayant répondu B obtiendra 0,3 points (il a répondu comme 30% du panel)
- l'étudiant ayant répondu C aura 0,6 points (il a répondu comme 60% du panel)
- l'étudiant ayant répondu D ou E aura 0 point.

Enfin, le crédit accordé pour chaque réponse est pondéré de sorte que la note maximale soit 1 et la note minimale soit 0. Dans l'exemple choisi,

- la réponse C rapportera 1 point (0,6 / 0,6),
- la réponse B rapportera 0,5 points (0,3 / 0,6),
- la réponse A rapportera 0,16 points (0,1 / 0,6),
- les réponses C et D rapporteront 0 point.

## References

BORDAGE, G. *Elaborated knowledge: A key to successful diagnostic thinking. Academic medicine*, 1994 69, pp. 883-885.

CHARLIN, B. *Le schéma comme structure de connaissances sous-jacente aux hypothèses dans l'investigation clinique médicale. Mémoire pour l'obtention du diplôme de maîtrise ès arts en sciences de l'éducation, Université de Sherbrooke 1994.*

CHARLIN, B., BRAILOVSKY, C.A., BRAZEAU-LAMONTAGNE, L., SAMSON, L., LEDUC, C. (1998a) *Script questionnaires: their use for assessment of diagnostic knowledge in radiology Medical Teacher*, 1998 20, pp.567-571.

CHARLIN, B., BRAILOVSKY, C. A., LEDUC, C., BLOUIN D. (1998b) *The Diagnostic Script Questionnaire: A new tool to assess a specific dimension of clinical competence Advances in Health Sciences Education*, 1998, 3, pp. 51-58.

CHARLIN, B., TARDIF, J., BOSHUIZEN, H. (2000a) *Scripts and medical diagnostic knowledge: theory and applications for clinical reasoning instruction and research. Academic Medicine*, 2000, 75, pp.182-190.

CHARLIN, B., ROY, L., BRAILOVSKY, C. A., & VAN DER VLEUTEN, C. P. M. (2000). *The Script*

Concordance Test: A tool to assess the reflective clinician. *Teaching and Learning in Medical Education*, 12, 189-195.

COHEN, R., RIEZNICK, R. K., TAYLOR, B. R., PROVAN, J., ROTHMAN, A. (1990) Reliability and validity of the objective structured clinical examination in assessing surgical residents *American journal of Surgery*, 160, pp. 302-305.

CUSTERS, J. F. M., REGHER, G., NORMAN, G. R. Mental representations of medical diagnostic knowledge: a review *Academic medicine*, 1996 71, pp. S55-S61.

ELSTEIN, A. S., SHULMAN, L. S., SPRAFKA, S. A. (1978) *Medical problem solving: an analysis of clinical reasoning* (Cambridge, MA, Harvard University Press)

FELTOVICH, P. J. (1983) Expertise: reorganizing and refining knowledge for use, *Professions Education Research Notes*, 4, pp. 5-9.

FELTOVICH, P. J., BARROWS, H. S. (1984) Issues of generality in medical problem solving. In H. G. Schmidt and M. L. De Volder (Eds.), *Tutorials in problem-based learning: A new direction in teaching the health professions*, Assen, Holland: Van Gorcum.

GRAND'MAISON, P., LESCOP, J., RAINSBERRY, P., BRAILOVSKY C. A. (1992) Large-scale use of an objective structured clinical examination for licensing family physicians *Canadian Medical Association Journal*, 146, pp. 1735-1740.

Mc NICOLL, A., BRAILOVSKY, C. A., BERTRAND, R., CARDINET, J. *EtudGen*, programme pour l'analyse de la généralisabilité pour Macintosh. CESSUL 1992, 1996, 1999. In D. BAIN et G. PINI. *Pour évaluer vos évaluations: la généralisabilité, mode d'emploi*. Centre de recherche psychopédagogiques, Genève, 1996.

SCHMIDT, H. G., NORMAN, G. R., BOSHUIZEN, H. P. A. (1990) A cognitive perspective on medical expertise: Theory and implications *Academic Medicine*, 65, pp. 611-621.

SIBERT, L., GRAND'MAISON, P., CHARLIN, B., GRISE, P. (1997). Evaluation de la compétence clinique en urologie: approche innovatrice basée sur l'observation de la performance. *Progrès en urologie*, 7, pp. 581-589.

SIBERT, L., GRAND'MAISON, P., DOUCET, J., WEBER, J., & GRISE, P. (2000) Initial experience of an Objective Structured Clinical Examination in evaluating urology residents, *European Urology*, 37, pp. 621-627.

SIBERT L., CHARLIN B., CORCOS J., GAGNON R., LECHEVALLIER J., P.GRISE P.. Assessment of clinical reasoning competence in urology with the Script Concordance test: An exploratory study across two sites from different countries. *European Urology* 2002;41:227-233.

ZEITZ, C. M. (1997) Some concrete advantages of abstraction: how experts' representations facilitate reasoning. In P. J. FELTOVICH, K. KM. FORD, R. R. HOFFMAN (Eds.) *Expertise in Context: Human and machine*. Memlo Park, CA: AAAI / MIT press

## Annexe 1: Canevas de questions et de grilles de correction.

*Le format dépend du type de compétence évaluée (Faire un diagnostic, Demander une exploration complémentaire, Choisir un traitement).*

### Evaluation d'une aptitude à proposer un diagnostic

Si vous pensez à...	Et que vous observez...	Votre hypothèse devient...
<i>Une hypothèse diagnostique</i>	<i>Une information clinique, un cliché d'imagerie, un résultat d'examen complémentaire</i>	-2 -1 0 +1 +2

- 2 Mon hypothèse est pratiquement éliminée
- 1 Mon hypothèse devient peu probable
- 0 Cette information n'a aucun effet sur mon hypothèse diagnostique
- +1 L'hypothèse devient plus probable
- +2 Mon hypothèse devient pratiquement certaine

### Pour l'évaluation d'une aptitude à choisir une investigation complémentaire

Si vous pensiez demander...	Et qu'alors vous trouvez	L'effet sur la nécessité de demander ce test est le suivant
<i>Option d'examen para clinique</i>	<i>Information clinique ou paraclinique</i>	-2 -1 0 +1 +2

Avec la grille de réponse suivante :

- 2 absolument contre-indiqué
- 1 peu utile ou plutôt néfaste
- 0 non pertinent ou inutile dans cette situation
- +1 utile et souhaitable
- +2 indispensable

### Pour l'évaluation d'une compétence à proposer un choix thérapeutique

Si vous pensiez faire	Et qu'alors vous trouvez	L'effet sur la pertinence de réaliser ce geste est le suivant
<i>Une option thérapeutique</i>	<i>Nouvelle information clinique ou résultat d'un examen complémentaire</i>	-2 -1 0 +1 +2

Avec la grille de réponse suivante :

- 2 absolument contre-indiqué
- 1 peu utile ou plutôt néfaste
- 0 non pertinent dans cette situation
- +1 utile et souhaitable
- +2 indispensable

## Annexe 2: Exemples de TCS pour la pathologie rachidienne

*Vous êtes amené(e) à examiner une jeune fille de 13 ans dont la maman a noté une déformation de la taille pendant les vacances, alors qu'elle était en maillot de bain.*

### Questionnaire diagnostique

Si vous pensez à...	Et que vous trouvez...	L'effet sur votre hypothèse diagnostique sera
<i>Une scoliose thoracique droite de 30°</i>	<i>Pas de gibbosité</i>	-2 -1 0 +1 +2
<i>Une scoliose thoracique droite de 30°</i>	<i>Une gibbosité thoracique droite de 3 cm</i>	-2 -1 0 +1 +2
<i>Une scoliose lombaire</i>	<i>Un rachis douloureux</i>	-2 -1 0 +1 +2
<i>Une scoliose thoracolombaire</i>	<i>Un déséquilibre latéral droit de 3 cm</i>	-2 -1 0 +1 +2
<i>Une scoliose</i>	<i>La mère a été opérée à 15 ans d'une arthrodèse vertébrale</i>	-2 -1 0 +1 +2
<i>Une scoliose lombaire</i>	<i>Une inégalité de longueur des membres inférieurs de 1,5 cm</i>	-2 -1 0 +1 +2
<i>Une tumeur maligne du rachis lombaire</i>	<i>Pas de raideur</i>	-2 -1 0 +1 +2
<i>Une tumeur maligne du rachis lombaire</i>	<i>Pas d'amaigrissement au cours des 6 derniers mois</i>	-2 -1 0 +1 +2
<i>Un spondylolisthésis</i>	<i>Une importante raideur du rachis lombaire</i>	-2 -1 0 +1 +2
<i>Un spondylolisthésis</i>	<i>Une rétraction bilatérale des ischiojambiers</i>	-2 -1 0 +1 +2

- 2 l'hypothèse est pratiquement éliminée
- 1 l'hypothèse devient moins probable
- 0 l'information n'a aucun effet sur l'hypothèse
- +1 l'hypothèse devient plus probable
- +2 il ne peut s'agir pratiquement que de cette hypothèse

## Questionnaire d'investigation complémentaire

Si vous pensez demander...	Et que vous observez...	Cette investigation sera...
<i>Une radiographie du rachis en entier de face en charge</i>	<i>Une gibbosité thoracique droite</i>	-2 -1 0 +1 +2
<i>Une radiographie du rachis en entier de face en charge</i>	<i>Un rachis équilibré de face et de profil sans gibbosité</i>	-2 -1 0 +1 +2
<i>Une échographie rénale</i>	<i>Une touffe pileuse sur la ligne médiane en L2</i>	-2 -1 0 +1 +2
<i>Une échographie rénale</i>	<i>Un rachis cervical court</i>	-2 -1 0 +1 +2
<i>Une échographie rénale</i>	<i>Pas d'antécédent d'infection urinaire</i>	-2 -1 0 +1 +2
<i>Une IRM médullaire</i>	<i>Pas de raideur rachidienne</i>	-2 -1 0 +1 +2
<i>Une IRM médullaire</i>	<i>Abolition des réflexes cutanés abdominaux</i>	-2 -1 0 +1 +2

- 2 absolument contre-indiqué  
 -1 peu utile ou plutôt néfaste  
 0 non pertinent ou inutile dans cette situation  
 +1 utile et souhaitable  
 +2 indispensable

## Pour l'évaluation d'une compétence thérapeutique

Si vous pensez proposer...	Et que vous observez...	L'indication thérapeutique devient...
<i>Une arthrodèse vertébrale</i>	<i>Une courbure thoracique de 25°</i>	-2 -1 0 +1 +2
<i>Une arthrodèse vertébrale</i>	<i>Une courbure thoracique de 60°</i>	-2 -1 0 +1 +2
<i>Un traitement orthopédique</i>	<i>Une courbure thoracique de 25°</i>	-2 -1 0 +1 +2
<i>3 séances de rééducation du rachis par semaine</i>	<i>Une courbure thoracique de 15°</i>	-2 -1 0 +1 +2
<i>Une arthrodèse vertébrale postérieure isolée</i>	<i>Age osseux de 10 ans</i>	-2 -1 0 +1 +2
<i>Une arthrodèse vertébrale postérieure isolée</i>	<i>Une courbure thoracique de 60° et une cyphose thoracique de 30° entre T11 et L1</i>	-2 -1 0 +1 +2

Avec la grille de réponse suivante :

- 2 absolument contre-indiquée  
 -1 peu utile ou plutôt néfaste  
 0 non pertinente dans cette situation  
 +1 utile et souhaitable  
 +2 indispensable