

Original article

## The Script Concordance Test in anesthesiology: Validation of a new tool for assessing clinical reasoning

Guillaume Ducos<sup>a</sup>, Corinne Lejus<sup>b</sup>, François Sztark<sup>c</sup>, Nathalie Nathan<sup>d</sup>, Olivier Fourcade<sup>a</sup>, Ivan Tack<sup>e</sup>, Karim Asehnoune<sup>b</sup>, Matthias Kurrek<sup>f</sup>, Bernard Charlin<sup>g</sup>, Vincent Minville<sup>a,\*</sup>

<sup>a</sup> Department of Anesthesiology and Intensive Care, EA 4564, Toulouse University Hospital, Toulouse, France

<sup>b</sup> Department of Anesthesiology and Intensive Care, Nantes University Hospital, Nantes, France

<sup>c</sup> Department of Anesthesiology and Intensive Care, Bordeaux University Hospital, Bordeaux, France

<sup>d</sup> Department of Anesthesiology and Intensive Care, Limoges University Hospital, Limoges, France

<sup>e</sup> Physiology Laboratory, Inserm U 858, Toulouse University Hospital - Rangueil, Toulouse, France

<sup>f</sup> Department of Anesthesia, University of Toronto, 150, College Street, Fitzgerald Building, Room 121, Toronto, Ontario, M5S 3E2, Canada

<sup>g</sup> Faculty of Medicine, University of Montreal, CP 6128, Succ. Centre-ville, Montréal, Québec, H3C 3J7, Canada

### ARTICLE INFO

Article history:  
Available online xxx

Keywords:  
Script Concordance Test  
Anesthesiology  
Resident

### ABSTRACT

**Objective:** To evaluate whether the Script Concordance Test (SCT) can discriminate between levels of experience among anesthesiology residents and attending physicians.

**Study type:** Multicenter (Toulouse, Nantes, Bordeaux and Limoges), prospective, observational study.

**Patients and methods:** A SCT made of 60 items was used to evaluate “junior residents” ( $n = 60$ ), “senior residents” ( $n = 47$ ) and expert anesthesiologists ( $n = 10$ ).

**Results:** There were no missing data in our study. Mean scores ( $\pm$ SD) were 69.9 ( $\pm$ 6.1), 73.1 ( $\pm$ 6.5) and 82.0 ( $\pm$ 3.5) out of a potential score of 100 for “junior residents”, “senior residents” and expert anesthesiologists, respectively. Results were statistically different between the 3 groups ( $P = 0.001$ ) using the Kruskal-Wallis test. The Cronbach's  $\alpha$  score was 0.63.

**Conclusions:** The SCT is a valid and useful tool for discriminating between anesthesia providers with varying levels of experience in anesthesiology. It may be a useful tool for documenting the progression of reasoning during anesthesia residency.

© 2015 Société française d'anesthésie et de réanimation (Sfar). Published by Elsevier Masson SAS. All rights reserved.

### 1. Introduction

Initial medical training aims at giving students the professional competence that will enable them to practice their specialty. Although a sound knowledge base is essential, it represents only the foundation upon which successful clinical reasoning is built. It is well documented that clinical expertise is acquired by learning techniques such as examination skills or the completion of technical procedures, which trigger diagnostic or treatment responses [1–3]. Successful clinical reasoning thus depends on the ability to organize various pieces of information from multiple data sets [1–3].

The Script Concordance Test (SCT) was developed over the last 20 years in order to evaluate clinical reasoning [1–5]. Script theory asserts that experienced practitioners have developed networks of

knowledge, specifically in the performance of routine care, called illness scripts [4,5]. Networks of these scripts are made of links between knowledge of the disease, clinical symptoms, possible complications and relevant treatment [6].

The Script Concordance Test has been validated as a tool for separating groups of participants with varying degrees of clinical expertise in gynecology [7], urology [8], internal medicine [9], radiotherapy [10], neurology [5], as well as in nursing education [11].

While Multiple Choice Questions (MCQ) and/or Open Short Answer Questions (OSAQ), commonly used in validation exams for anesthesiology courses, can evaluate knowledge that was acquired during residency training, these methods mainly estimate factual knowledge and do not assess networks of knowledge and clinical reasoning. To date, SCTs have never been used to assess anesthesiologists.

The purpose of our study was to evaluate the validity of the SCT as a tool for assessing clinical reasoning in anesthesiology and to discriminate between groups of anesthesia residents and practitioners with varying levels of clinical experience.

\* Corresponding author. Department of Anesthesiology and Intensive Care, Rangueil University Hospital, avenue Jean-Poulhès, 31059 Toulouse cedex 9, France. Tel.: +33 05 61 32 27 12; fax: +33 05 61 32 22 32.

E-mail address: [minville.v@chu-toulouse.fr](mailto:minville.v@chu-toulouse.fr) (V. Minville).

<http://dx.doi.org/10.1016/j.accpm.2014.11.001>

2352-5568/© 2015 Société française d'anesthésie et de réanimation (Sfar). Published by Elsevier Masson SAS. All rights reserved.

2. Material and methods

2.1. Test design

The anesthesiology Script Concordance Test included 20 cases, presented as small vignettes from common tasks performed in several

different areas of anesthesiology (for example: elective surgery, emergencies, pediatrics, obstetrics, regional anesthesia, preoperative consultations, postoperative analgesia, or management of perioperative coronary ischemia). There was no question about Intensive Care in the SCT. Each vignette (several lines long) described a specific clinical case or task, though sometimes incompletely (Fig. 1).

Vignette 8

You support a 28 years-old man in the emergency room. He presents a closed fracture of the right leg.

If you think about...	And you discover...	Hypothesis become...
Perform a regional anesthesia (femoral and sciatic nerves blocks)	Persistent paresthesia in the peroneal nerve territory	-2 Contraindicated -1 Less useful 0 Neither more or less usefull +1 More useful +2 Essential
Perform a regional anesthesia (spinal anesthesia)	A congenital lung dyspalsia with chronic respiratory failure	-2 Contraindicated -1 Less useful 0 Neither more or less usefull +1 More useful +2 Essential
Perform a general anesthesia	A Cormack score at 4	-2 Contraindicated -1 Less useful 0 Neither more or less usefull +1 More useful +2 Essential

Vignette 8		
You are caring for a 28 year old man in the emergency room. He presents a closed fracture of the right leg.		
If you think about...	And you discover...	The hypothesis becomes...
Performing regional anesthesia (femoral and sciatic nerve blocks)	Persistent paresthesia in the peroneal nerve territory	-2 Contraindicated -1 Less useful 0 Neither more or less useful +1 More useful +2 Essential
Performing regional anesthesia (spinal anesthesia)	A congenital lung dyspalsia with chronic respiratory failure	-2 Contraindicated -1 Less useful 0 Neither more or less useful
		+1 More useful +2 Essential
Performing general anesthesia	A Cormack score of 4	-2 Contraindicated -1 Less useful 0 Neither more or less useful +1 More useful +2 Essential

Fig. 1. Example of a clinical vignette with Script Concordance Test (SCT) items.

Each vignette then listed three items, organized in three parts distributed on one line:

- part one: a relevant hypothesis (diagnostic, therapeutic or investigative) was given by the test;
- part two: new elements (medical history, clinical signs, current treatment) not derived from the vignette, and interacting with the previously submitted hypothesis are given;
- part three: a response in the form of a Likert scale with five levels (from +2 to -2; e.g. “-2” corresponds to “contraindicated” and “+2” to “strongly indicated”) is associated with and noted for each item.

Data from individual items can be linked by a common case/task reported in a vignette, but do not apply to the following item; i.e. the questions were all independent of each other.

## 2.2. Development of vignettes

A total of 20 vignettes (with 60 items) were developed by expert consensus and used for the final SCT. The experts who developed the vignettes were different from the experts who participated in the panel.

## 2.3. Scoring of the SCT

Scoring was performed in accordance with current data in the literature as the basis of the modal response [8]. Panel experts took the SCT under the same conditions as candidates.

The answer that collected the greatest number of votes from experts was rated 1 (modal response), other answers chosen by one or more experts were rated as a fraction, while those that were not chosen were rated 0 (Table 1). Thus all questions were scored from 0 to 1.

A perfect score of 100 would mean that a given candidate selected the modal response for all items (i.e. the answer provided by a majority of experts), though this has never been found in the literature [8]. The results are expressed as scores returned on 100 points.

## 2.4. Experts

Previous studies have shown that the ideal number of experts was between 10 and 20 [9]. For the present study, the selected experts [10] were experienced academic anesthesiologists from the same institutions as the residents. The expert panel answered the questionnaire during the academic year 2010–2011 under the same conditions as the residents.

## 2.5. Participants

Participants were volunteer anesthesiology residents from four French medical schools (Toulouse, Limoges, Bordeaux and Nantes).

**Table 1**  
Scoring of the Script Concordance Test (SCT) via modal responses.

	-2	-1	0	+1	+2
Expert panel responses			3	5	2
Rating			3/10	5/10	2/10
Rating adapted to the modal response			3/5	5/5	2/5
Final rating			0.6	1	0.4

In this example, +1 is the majority response given by the panel of experts (5/10). Therefore, +1 is the modal response for this item, which corresponds to a rating of 1 point. The 2 other responses mentioned by the experts correspond to ratings of a fraction of a point, reflecting the number of experts who agree. The responses that were never selected by the experts are rated as “0” points.

**Table 2**

Distribution of anesthesiology residents from the four French medical schools (Toulouse, Limoges, Bordeaux and Nantes).

	1st year	2nd year	3rd year	4th year	5th year	Total
Bordeaux	12	6	2	4	2	26
Limoges	4	2	3	2	0	11
Toulouse	17	6	8	7	6	44
Nantes	7	6	6	3	4	26
Total	40	20	19	16	12	107
	Junior residents n = 60		Senior residents n = 47			

(Table 2). They took the SCT during their residency under standardized conditions (one-hour test under surveillance, which corresponds to 1 minute per item) during the 2010–2011 academic year. We separated the participants by degree of clinical experience: the group “junior residents” consisted of first and second year residents (n = 60) and the group “senior residents” of third, fourth and fifth year residents (n = 47), which were compared with the expert panel. We used all participants (residents and experts) available at the time of the study, without taking into account power calculations (convenience sample).

## 2.6. Statistical analysis

The results for each participant were calculated (score per item and total score) with software available from the University of Montreal website (<http://www.sctmed.ca>). The results for experts were calculated using a key excluding their own response to the scoring grid. The results of the SCT remained anonymous.

Data were presented as mean  $\pm$  SD and interquartile ranges (IQR). Differences between groups were evaluated with the Kruskal-Wallis test and the comparison between groups was analyzed using the Mann-Whitney test with Bonferroni’s corrections.

A *P* value < 0.05 was considered statistically significant. Internal consistency was estimated using Cronbach’s coefficient  $\alpha$ . Statistical analysis was performed using Statview software (SAS Institute Inc., Cary, North Carolina).

## 3. Results

The distribution of results for each group is presented in Fig. 2 Part A for junior residents, in Part B for senior residents and in Part C for the panel of experts.

The average score for the “junior resident” group was  $69.6 \pm 6.1$  (range 56.7, 81.1) and  $72.2 \pm 6.5$  (range 47.8, 81.8) for the “senior resident” group. The average score for the “expert” group was  $82.9 \pm 7.1$  (range 77.4, 89.1). There was a statistically significant difference between results for the “junior” and “senior” residents ( $P = 0.0076$ ), as well as between the latter two groups and the “experts” group ( $P < 0.0001$  and  $P = 0.0003$ , respectively). The Script Concordance Test was thus able to discriminate between the different levels of experience in the tested anesthesiologist group. Cronbach’s coefficient  $\alpha$  for all groups was 0.63.

## 4. Discussion

Assessment of clinical reasoning is important, even though we do not know much about its pattern of acquisition. One theory on how clinicians develop their decision-making skills is based on a script theory that comes from cognitive psychology [1]. The SCT is based on that theory and aims at assessing script development. It represents an attractive and innovative tool for evaluating clinical reasoning during physician training.

In our study of 107 residents in Anesthesiology/Intensive care as well as 10 experts, we found that the SCT was able to

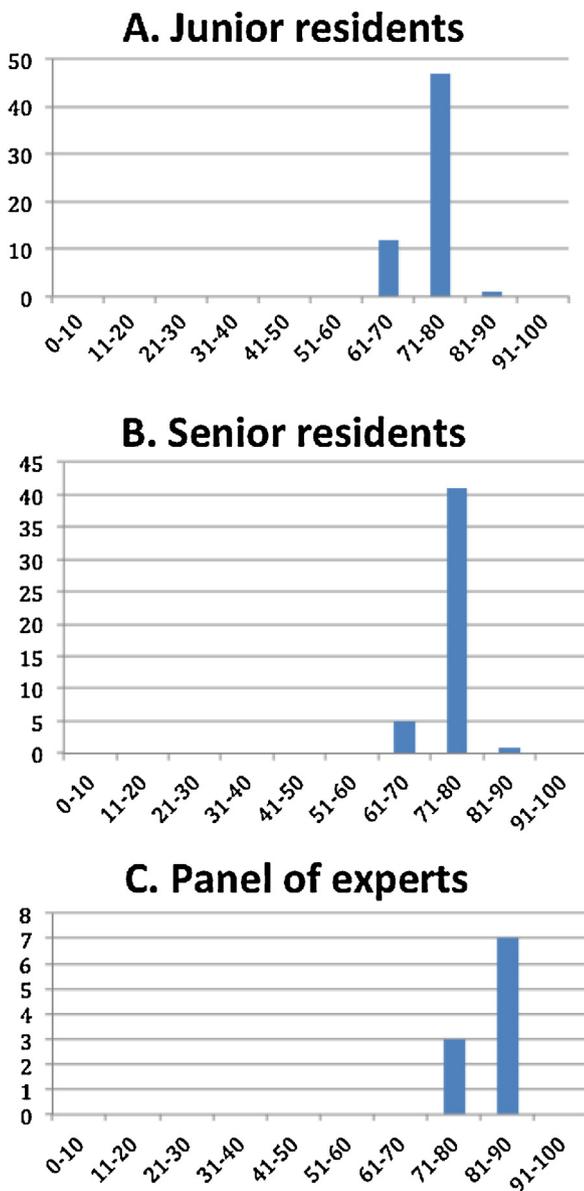


Fig. 2. a, b, c: Script Concordance Test (SCT) score distributions for each group.

discriminate between groups of different clinical experience, thus validating this instrument for anesthesiology as was previously done for radiotherapy [10], general surgery [12], otolaryngology [13], urology [8] and gynecological surgery [7]. The groups identified (“junior resident”, “senior resident” and “expert”) are all different from each other and reflect the accumulation of experience. The results of the SCT for “senior residents” are located between experts and “junior residents”, but there were significant differences between groups (cf. results section), as previously described in other specialties [5].

Unlike most recent studies published on the subject [13], which used a computer-based examination, our participants took the SCT using paper and pencil. This method has several advantages: no missing data, time control of the test performance (1 hour is allotted, 1 minute/item), and supervision of participants to avoid outside support. Moreover, the SCT is highly reproducible (standardized conditions) and feasible. It can easily be inserted into a lecture program during the course of initial and continuing training.

To our knowledge, this is the first study testing the validity of the SCT in the field of anesthesiology. We demonstrated that the SCT could differentiate between groups based on different levels of

clinical experience in anesthesiology. It will be interesting to investigate whether it can evaluate the evolution of clinical reasoning by testing the same participants longitudinally. Finally, the SCT could be used to identify participants with low scores who may benefit from remedial educational measures such as theoretical training, practical workshops and exercises using clinical simulation.

Our study has some limitations. The small number of experts ( $n = 10$ ) was sufficient given the present state of knowledge [14]. Anesthesiology is a broad specialty and the questions covered areas that some individuals on the expert panel may not have practiced for some time. However, the results are consistent with previous data about SCTs in another specializations [5,7–11]. Finally, the SCT is an innovative yet new method of evaluation and both the participants and experts may not have been familiar with the test, even though prior information about the test and its modalities was given. In our study, all groups had the same practice and experience with the test.

## 5. Conclusion

In summary, the SCT appears to measure a dimension of reasoning and knowledge that is different from those evaluated by traditional assessment tools. It explores the interpretation of data in a clinical context and in relation to clinical experience. This study provides evidence in favor of the SCT as a reliable and valid tool for evaluating clinical reasoning skills among anesthesiology residents. A low score on a SCT may identify residents who could benefit from additional educational interventions.

## Disclosure of interest

The authors declare that they have no conflicts of interest concerning this article.

## Acknowledgements

We thank Professors Samii and Geeraerts as well as Drs. Bayoumeu, Penonnard, Olivier and Schnell for their expert advice.

**Summary statement:** the Script Concordance Test is a useful tool for documenting the progression of reasoning during residency training. This work should be attributed to the Department of Anesthesiology, Toulouse University Hospital - Rangueil, Toulouse, France. Support was provided solely from institutional and departmental sources.

**Authors' contributions:** C.L., N.N., O.F. and F.S. are local coordinators for pedagogy in anesthesiology and helped organize resident participation. K.A., M.K., and C.B. helped draft the manuscript. O.F. and I.T. participated in study design and manuscript drafting. G.D. and V.M. originated the study, participated in its design and coordination, helped draft the manuscript and perform the statistical analyses. All authors read and approved the final manuscript.

## References

- [1] Collard A, Gelaes S, Vanbelle S, Bredart S, Defraigne JO, Boniver J, et al. Reasoning versus knowledge retention and ascertainment throughout a problem-based learning curriculum. *Med Educ* 2009;43:854–65.
- [2] Charlin B, Tardif J, Boshuizen HP. Scripts and medical diagnostic knowledge: theory and applications for clinical reasoning instruction and research. *Acad Med* 2000;75:182–90.
- [3] Charlin B, Roy L, Brailovsky C, Goulet F, van der Vleuten C. The Script Concordance test: a tool to assess the reflective clinician. *Teach Learn Med* 2000;12:189–95.
- [4] Lubarsky S, Charlin B, Cook DA, Chalk C, van der Vleuten CPM. Script concordance testing: a review of published validity evidence. *Med Educ* 2011;45:329–38.

- [5] Lubarsky S, Chalk C, Kazitani D, Gagnon R, Charlin B. The Script Concordance Test: a new tool assessing clinical judgement in neurology. *Can J Neurol Sci* 2009;36:326–31.
- [6] Dory V, Gagnon R, Vanpee D, Charlin B. How to construct and implement script concordance tests: insights from a systematic review. *Med Educ* 2012;46:552–63.
- [7] Park AJ, Barber MD, Bent AE, Dooley YT, Dancz C, Sutkin G, et al. Assessment of intraoperative judgment during gynecologic surgery using the Script Concordance Test. *Am J Obstet Gynecol* 2010;203. 240.e1–6.
- [8] Sibert L, Darmoni SJ, Dahamna B, Hellot MF, Weber J, Charlin B. On line clinical reasoning assessment with Script Concordance test in urology: results of a French pilot study. *BMC Med Educ* 2006;6:45.
- [9] Marie I, Sibert L, Roussel F, Hellot MF, Lechevallier J, Weber J. [The script concordance test: a new evaluation method of both clinical reasoning and skills in internal medicine]. *Rev Med Interne* 2005;26:501–7.
- [10] Lambert C, Gagnon R, Nguyen D, Charlin B. The script concordance test in radiation oncology: validation study of a new tool to assess clinical reasoning. *Radiat Oncol* 2009;4:7.
- [11] Deschênes MF, Charlin B, Gagnon R, Goudreau J. Use of a script concordance test to assess development of clinical reasoning in nursing students. *J Nurs Educ* 2011;50:381–7.
- [12] Meterissian SH. A novel method of assessing clinical reasoning in surgical residents. *Surg Innov* 2006;13:115–9.
- [13] Kania RE, Verillaud B, Tran H, Gagnon R, Kazitani D, Huy. et al. Online script concordance test for clinical reasoning assessment in otorhinolaryngology: the association between performance and clinical experience. *Arch Otolaryngol Head Neck Surg* 2011;137:751–5.
- [14] Gagnon R, Charlin B, Coletti M, Sauvé E, van der Vleuten C. Assessment in the context of uncertainty: how many members are needed on the panel of reference of a script concordance test? *Med Educ* 2005;39:284–91.