The Script Concordance Test (SCT): Evaluating Clinical Reasoning Skills

Brett Windsor, PT, MPA, OCS, COMT, FAAOMPT
PhD Student (RMUoHP)
Assistant Professor of Health Sciences

College of Pharmacy & Health Sciences
Campbell University, Buies Creek, NC
Thursday January 24, 2013
Session Objectives

1. Define the term "Script Concordance Test."
2. Describe the history and development of the Script Concordance Test.
3. Briefly describe the cognitive theory of clinical expertise development as it relates to the development of clinical-reasoning skills.
4. Discuss the body of literature supporting the use of the Script Concordance testing in the education of health care professionals.
5. Recognize the role that the Script Concordance Test currently plays in the evaluation and prediction of clinical performance.
6. Describe the steps involved in compiling a "reference panel" appropriate to the scoring of a Script Concordance Test.
7. Demonstrate the construction of a Script Concordance Test case scenario and 3 associated test items.
8. Demonstrate the method used to accurately score a Script Concordance Test case scenario and 3 associated test items.
9. Discuss the implications of the Script Concordance Test and the opportunities for adoption into the education of pre-professional and post-professional physical therapy students.
My Session Goals...

1. Offer a complement to existing assessments
2. Explain the Script Concordance Test (SCT)
3. Provide some background on the SCT
4. Explain how you construct and score an SCT
5. Show you an example of an SCT...have you complete and score a short SCT test example
6. Discuss where I think SCT has a place in the physical therapy profession
7. Take some questions and feedback...ideas...
Clinical Reasoning

Not what is it...or how do we do it...or teach it...
A few questions to consider...

How do we recognize it?
How do we measure it?
How do we know if it gets better with experience?
Are the people that we call ‘experts’, really ‘experts’?
Clinical Reasoning

- Dory et al., 2012: CR lies at the heart of clinical practice
  - Requires knowledge and technical skill and intuition
  - Requires an ability to integrate knowledge and technical skill to make decisions about patient care
    - Knowledge and technical skill don’t guarantee good clinical decisions
- Eva, 2005: Core task of clinical teachers
  - Iterative Clinical reasoning process:
    - Hypothesis generation → Intuitive → Pattern recognition
    - Hypothesis testing → Analytical
  - Script Theory
    - “Sort through a cluster of features presented by a patient and accurately assign a diagnostic label”
    - Goal is to apply the most appropriate treatment strategy
Clinical Reasoning

• Cornerstone of medical practice...physical therapy practice...clinical practice

• Cornerstone of educational practice...
  – The Development of CR skills
  – The Measurement of CR skills
    1. Patient Management Problems
       – Short descriptions of clinical situations, followed by a series of questions
       – Intermediate effect → abandoned
       – Case specificity → led towards efforts to assess knowledge of students and practitioners...MCQ’s...
Measuring Clinical Reasoning

2. Longer form written examinations
   – Difficult to establish reliability, validity
   – Can’t reproduce decision-making structures
   – Institutions are beginning to drop them\textsuperscript{52}

3. Objective Structured Clinical Examination (OSCE)
   – Difficulty of standardization, scoring objectivity, practicability
   – In similar situations, professions show substantial variation in performance and thought patterns in any particular real or simulated case (Charlin et al., 2000)
Measuring Clinical Reasoning

4. Multiple Choice Question Test (MCQ)
   – Provide reliable, valid assessment of relevant knowledge
   – Practical, Accessible, Feasible
   – Good psychometric properties encouraged testers to extend the range of MCQ beyond knowledge
     • Rich descriptions of patient cases
     • Variations of the format → matching
   – Limitations:
     • Testing limited to cases that have only one answer – straightforward...
     • Can’t address the variability, uncertainty of clinical practice
     • Can’t address the variability in expert opinion, application of evidence to a case
     • Knowledge is not the same as clinical expertise...
The Multiple Choice Test: Fact

What is 2 + 2?

A. 3
B. 5
C. 4
D. 7
The Multiple Choice Test

A. Pink
B. Off-Pink
C. Black
D. Pinkish
Facets of Clinical Practice

Inherently uncertain
Ill-defined
Non-linear
Unpredictable
Heterogeneous
Complex
Ambiguous

Legitimate disagreement amongst ‘experts’ as to how to best manage clinical scenarios
Assessment of *knowledge* ✓
Assessment of technical *skill* ✓
Assessment of straightforward clinical scenarios ✓

Assessment of clinical reasoning, decision-making in complex, ambiguous patient scenarios... ✅
A complementary technique?
Never touch a third rail! It could kill you.

A public service message from WNYCTA.
The Script Concordance Test (SCT)

- How closely an examinee can mimic the decision making process of a group of experts...
- Incorporates ambiguity...variability
- Written Test...online or paper
- 60-90 minutes
- Assesses process, not outcome
  - Does not dictate process,
- Responses of examinees are compared to a panel of experts
  - Allows for variability in response
  - The more ‘concordance’ with the expert panel, the higher the score
The Script Concordance Test

• A real-life, ill-defined clinical case is provided, in the form of a brief clinical vignette where the information provided is insufficient to reach a decision

• Each case followed by a number of test items
  1. A diagnostic hypothesis, investigation option, or treatment option relevant to the clinical situation
  2. New information is presented
  3. Examinees are asked to evaluate the impact of this new information on the likelihood that the proposed hypothesis is correct
1. A 69 year-old female who is obese and wheelchair bound is phone-referred to you by a familiar source. The patient presents that same day with episodic, intermittent pain centered over the right C2-3 area. The pain is of acute insidious onset; it began this morning upon waking. The pain varies between 6/10 & 8/10 on the Visual Analog Scale.

<table>
<thead>
<tr>
<th>If you were thinking of the following diagnosis…</th>
<th>And the following new information became available…</th>
<th>Then this hypothesis could become:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acute traumatic arthritis of the right C2-C3 zygapophyseal joint</td>
<td>Pain increased with active and passive left cervical spine side bending</td>
<td>2 -1 0 +1 +2</td>
</tr>
<tr>
<td>Instability of the right C2-3 segment into extension</td>
<td>Left active cervical spine rotation was limited to 30 degrees</td>
<td>2 -1 0 +1 +2</td>
</tr>
<tr>
<td>An acute vascular event</td>
<td>Diminished light touch sensation in the right index finger and thumb</td>
<td>2 -1 0 +1 +2</td>
</tr>
</tbody>
</table>

- 2 Significantly less likely
- 1 Slightly less likely
0 No change at all in the likelihood
+1 Slightly more likely
+2 Significantly more likely

A vignette such as this is a challenging and authentic clinical situation.

Credits on each item are derived from the answers given by a panel of reference.

A 5-point Likert Scale records the students’ answer.

Windsor-CIM 2013 All Rights Reserved.
SCT History

• Script Concordance Testing (SCT)
  – Developed in late 1990’s by Bernard Charlin
  – Measurement based on theory
    • Adaptation of cognitive psychology script theory
  – Measures process, not outcome
  – Judgments can be measured and compared to a reference panel of experienced practitioners
  – Scripts can be used to assess performance in the context of ill-defined problems in the context of uncertainty with no factual solutions
  – Easy to construct
  – Easy to score
  – Easy to administer
Test Construction

- The topic is decided
  - Know what you’re testing
    - Diagnosis, Evaluation, Intervention
- 2-3 writers compile real life clinical scenarios (20-30)
  - Must be some ambiguity
  - Ideally a patient you’ve actually seen (i.e. real)
- Write items for scenarios (3 per scenario)
- A reference panel of experts (10-15) takes test
- Filter the test for ‘reliable’ questions
- Reference panel responses used to develop answer key
- Examinees take the test
- Test scored
Test Scoring

- Aggregate Scoring Method\textsuperscript{28}
  - 10 experienced panel members take the test and their answers are collated
  - Credit for each answer is the number of panel members that have provided that answer, divided by the modal value for that item
Test Example
Instructions

The following examination represents a method called Script Concordance Testing (SCT). It is a testing method designed to assess the development of your clinical reasoning skills. It is important to understand that there are no "right" answers. The exam is designed to collect your assessment of each situation, and assess your decision-making processes.

In this examination, there are ten (10) individual case scenarios, each one representing a real clinical situation. Each scenario provides you with some basic information about the case. You are then asked to indicate what decisions you would make given the addition of some new information. Each scenario asks you to make three decisions.

It is important that you only consider the information that is presented. DO NOT make any other assumptions about the case.

You will make a decision by choosing a number on a Likert Scale, -2, -1, 0, 1, or 2. Each number represents a choice that is explained in the scale directly below the scenario. Please choose the number that most accurately represents the decision that you would make in the given situation.

You have sixty (60) minutes in which to complete the examination.

Please let me know if you have any questions.
1. A 69 year-old female who is obese and wheelchair bound is phone-referred to you by a familiar source. The patient presents that same day with episodic, intermittent pain centered over the right C2-3 area. The pain is of acute insidious onset; it began this morning upon waking. The pain varies between 6/10 & 8/10 on the Visual Analog Scale.

<table>
<thead>
<tr>
<th>If you were thinking of the following diagnosis...</th>
<th>And the following new information became available...</th>
<th>Then this hypothesis would become...</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acute traumatic arthritis of the right C2-C3 zygapophyseal joint</td>
<td>Pain increased with active and passive left cervical spine side bending</td>
<td>-2 -1 0 +1 +2</td>
</tr>
<tr>
<td>Instability of the right C2-3 segment into extension</td>
<td>Left active cervical spine rotation was limited to 30 degrees</td>
<td>-2 -1 0 +1 +2</td>
</tr>
<tr>
<td>An acute vascular event</td>
<td>Diminished light touch sensation in the right index finger and thumb</td>
<td>-2 -1 0 +1 +2</td>
</tr>
</tbody>
</table>

-2 Significantly less likely
-1 Slightly less likely
0 No change at all in the likelihood
+1 Slightly more likely
+2 Significantly more likely
2. A 35 year-old male presents to you with paresthesia in the left medial forearm, extending from the medial epicondyle region to the distal portion of the 5th digit. The symptoms began suddenly 3 weeks ago after falling asleep in an armchair.

<table>
<thead>
<tr>
<th>If you were thinking of the following diagnosis...</th>
<th>And the following new information became available...</th>
<th>Then this hypothesis would become:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Radiculopathy of the C8 nerve root</td>
<td>Fatigable weakness of the left wrist flexors</td>
<td>-2  -1  0  +1  +2</td>
</tr>
<tr>
<td>Peripheral neuropathy of the ulnar nerve</td>
<td>Fatigable weakness of the left ulnar deviators</td>
<td>-2  -1  0  +1  +2</td>
</tr>
<tr>
<td>An apical lung tumor</td>
<td>Fatigable weakness of the left elbow flexors</td>
<td>-2  -1  0  +1  +2</td>
</tr>
</tbody>
</table>

-2  Significantly less likely
-1  Slightly less likely
0   No change at all in the likelihood
+1  Slightly more likely
+2  Significantly more likely
3. Your patient is a 55 year-old female who was involved in an MVA one year ago. She presents with painful active cervical spine rotation associated with dizziness. She has been seeing a chiropractor for the last six months, which has not helped her symptoms at all.

![Diagram showing left and right sides of patient’s body with notes on cervical spine rotation]

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>New Information Available</th>
<th>Likelihood</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atlantoaxial Instability</td>
<td>The patient has longstanding rheumatoid arthritis</td>
<td>-2 -1 0 +1 +2</td>
</tr>
<tr>
<td>A tear of the tunica intima of one of the vertebral arteries</td>
<td>Patient reports sharp pain at ten (10) degrees of active right rotation</td>
<td>-2 -1 0 +1 +2</td>
</tr>
<tr>
<td>A cervical myelopathy</td>
<td>Patient experienced regular upper cervical spine pain post treatment</td>
<td>-2 -1 0 +1 +2</td>
</tr>
</tbody>
</table>

-2 Significantly less likely
-1 Slightly less likely
0 No change at all in likelihood
+1 Slightly more likely
+2 Significantly more likely
4. A 35 year-old truck driver has been referred to you following a fall onto an outstretched hand six weeks ago. He presents to you for initial evaluation with pain in the left lateral humeral epicondyle region.

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>New Information</th>
<th>Hypothesis Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lateral epicondylalgia</td>
<td>The patient had pain in the lateral elbow prior to the recent fall</td>
<td>-2 -1 0 +1 +2</td>
</tr>
<tr>
<td>A mechanical restriction in the ulnohumeral joint</td>
<td>The patient describes medial forearm paresthesia</td>
<td>-2 -1 0 +1 +2</td>
</tr>
<tr>
<td>A fracture of the anatomical neck of the scaphoid bone</td>
<td>There is no pain reproduced with radial deviation + overpressure</td>
<td>-2 -1 0 +1 +2</td>
</tr>
</tbody>
</table>

-2: Significantly less likely  
-1: Slightly less likely  
0: No change at all in the likelihood  
+1: Slightly more likely  
+2: Significantly more likely
5. Your patient is a 43 year-old male who presents with continuous episodic pain in the mid left cervical spine. The pain frequently refers to the left superior and posterior shoulder region and the mid-vertebral border of the scapula as a vague, non-specific ache.

![Diagram of body with red arrows indicating pain areas]

<table>
<thead>
<tr>
<th>If you were thinking of the following diagnosis...</th>
<th>And the following new information become available...</th>
<th>Then this hypothesis would become:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disc pathology at the C5-6 level</td>
<td>The patient is experiencing intermittent paresthesia in the left UE</td>
<td>-2 -1 0 +1 +2</td>
</tr>
<tr>
<td>Cervical myelopathy</td>
<td>The patient complains of paresthesia in the left index finger</td>
<td>-2 -1 0 +1 +2</td>
</tr>
<tr>
<td>A radiculopathy of the C6 nerve root</td>
<td>The patient has experienced occasional shooting, electric pain into the left arm</td>
<td>-2 -1 0 +1 +2</td>
</tr>
</tbody>
</table>

-2 Significantly less likely
-1 Slightly less likely
0 No change at all in the likelihood
+1 Slightly more likely
+2 Significantly more likely
### Scoring the Test

<table>
<thead>
<tr>
<th>Reference Panel of 10 members (Score for individual items = number of members choosing option/10)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Scenario 1</strong></td>
</tr>
<tr>
<td>Item B</td>
</tr>
<tr>
<td>Item C</td>
</tr>
<tr>
<td><strong>Scenario 2</strong></td>
</tr>
<tr>
<td>Item B</td>
</tr>
<tr>
<td>Item C</td>
</tr>
<tr>
<td><strong>Scenario 3</strong></td>
</tr>
<tr>
<td>Item B</td>
</tr>
<tr>
<td>Item C</td>
</tr>
<tr>
<td><strong>Scenario 4</strong></td>
</tr>
<tr>
<td>Item B</td>
</tr>
<tr>
<td>Item C</td>
</tr>
<tr>
<td><strong>Scenario 5</strong></td>
</tr>
<tr>
<td>Item B</td>
</tr>
<tr>
<td>Item C</td>
</tr>
</tbody>
</table>

**TOTAL** | **15**

**X 10** | **150**
Evidence...Specifics

- Test construction
  - Fournier et al., 2008 for how to construct
  - Evidence for internal consistency strong\(^{42}\)
    - Consider case versus item though...overestimation
  - Optimal number of items or cases
    - 36 cases (1-2 items each)\(^{7}\)
    - 130 items\(^{53}\)
    - 48-102 items\(^{29}\)
    - Adding items rather than cases is more effective in increasing reliability
    - Ceiling effect after 3-4 items per case
    - 25 cases, 3 items, results in reliability coefficients of 0.75 – 0.86
    - Eliminate individual items with poor item-total correlations
    - IDEAL: 100 items, approx. 90 minute test.
Evidence...Specifics

- Panel Composition
  - One study has examined ideal number of experts on a panel\textsuperscript{21}
    - 38 experts used – reliability = 0.76
    - Samples of at least 10 panelists provided satisfactory estimates of internal consistency
    - Little gain when panels exceed 20
    - 10-20 members considered ideal
    - Larger panels lead to higher mean test scores
Evidence...Specifics

• Examination of Panel’s responses (Dory et al)
  — Overall disagreement of panel members should not be considered measurement error
  — Should be considered reflective of the ill-defined nature of the clinical case...KEY...
  — Items the elicit a unanimous answer are better suited to knowledge based MCQ’s
  — High variability raises questions:
    • Wording of question
    • Concerns on expertise of panel members
Evidence...Specifics

• Scoring
  – 5 point Likert scales optimal (originally 7 point scale)
    • Two studies have compared 5- and 3- point scales \(^5^3\)
    • Traditional scoring method weights alternative responses against
      the modal answer – confirmed reliable and discriminatory \(^5^3\)
  – Discriminative ability
    • Aim for panelist scores mean of 80; standard deviation (SD) of 5 to
      help examinees and educators interpret test scores.
    • Effect sizes between students and experts have been in the range
      of 0–4.74, typically around 2 \(^5^4\)
    • Effect sizes between residents and experts have been in the range
      of 0–3.2, typically around 1 \(^5^4\)
    • Students would be expected to achieve a mean score of around 70
      and residents a mean score of around 75.
Validity

• Lubarsky et. al. 42
  – Published research generally supports the use of SCT to assess the interpretation of clinical data under conditions of uncertainty42
    • High internal consistency reliability supports the internal structure of SCT scores
    • SCT scores correlate poorly with assessments of pure factual knowledge
    • validity of SCT scores is weakly supported by evidence pertaining to examinee response processes and educational consequences

• Kreiter39
Limitations...

- The tests are not easy to write
  - Learning curve
  - Writing in ambiguity is harder than not
- Up to 30% of the items in a test get tossed
- Large learning curve for students
- Difficulties assembling a panel
  - Who is an expert?
  - Numbers required...
- No literature base in physical therapy
- Very strongly entrenched assessment methods
Discussion:
Application to Physical Therapy
Adoption of the SCT...

- Emergency Medicine
- Pediatric Emergency medicine
- Dermatology
- Neurosurgery
- Seating mobility (PT)
- Nursing students
- Veterinary students
- Otorhinolaryngology
- Radiation Oncology
- Neurology
- Internal Medicine
- gynecologic surgery
- geriatric urinary incontinence
- Thrombophilia testing
- Urology
Validation of the Seating and Mobility Script Concordance Test

ABSTRACT The purpose of this study was to develop the seating system for the Seating and Mobility Script Concordance Test (SMSC). The SMSC was validated against instrumental and clinical performance measurement, and measured the reliability of the SMSC. The SMSC purpose is to provide a method for assessing knowledge of seating and mobility prescription. A sample of 106 therapists and 15 spinal cord injury patients contributed to the development of the seating system. Validity evidence was obtained using 15 seating and mobility experts, 10 orthotists, and other therapists with varying levels of seating and mobility expertise. Proxy measures of clinical expertise were used for statistical validity evidence since no criterion measures existed. The SMSC was found to differ between seating and mobility experts and orthotists' assessment scores (p = 0.08). The proxy measure of clinical expertise, seating and mobility hours/week, was found to predict SMSC assessment scores (p = 0.002). The statistical evidence of the SMSC may include evidence of reduced test performance but satisfactory convergent and discriminant evidence of construct validity. Although the SMSC may be a promising approach for assessing seating and mobility expertise, limitations exist in the current context. Future application of the SMSC should only be used after further development of the tool occurs.

INTRODUCTION The appropriateness of a consumer's seating and mobility (SM) system varies considerably depending on the consumer, proficiency, and expertise of the professional assessing the use (Ehrman & Lange, 1999). As present, there is a scarcity of skilled and knowledgeable therapists to evaluate and recommend SM devices (Ehrman & Lange, 1996; Kuhl & Filseth, 1987; Roseman, 2007). With ever-changing and emerging technologies available in the area of SM, it is necessary that clinicians continually update their knowledge, skills, and clinical competence in order to provide quality care. The need for a tool such as the SMSC to assess clinician's seating and mobility expertise in the clinic setting is evident. The SMSC represents a tool to evaluate seating and mobility expertise through evaluation of seating and mobility education activities. Yet, the effectiveness of professional training is self
Other uses of the SCT...

- Predictor of clinical reasoning performance end residency\(^3\)
- Identifying poor performers\(^26\)
- Assessment of pre-clinical medical students\(^29\)
- Intraoperative decision making skills\(^46\)
- Identify behaviors of highly professional resident physicians\(^28\)
- Online testing\(^30\)
Application to PT practice...

• The SCT can be adapted to fit any clinical specialty within physical therapy
• Complements current assessment methods
  – MCQ
  – Technical/Practical testing
• Two specific possibilities
  1. Residency-Fellowship
     • Assessment of Clinical Reasoning
     • Predictor of performance
     • Clinical Outcome Measure
  2. Specialty Certification
     • Provides a truer measure of clinical expertise
Example...Deschennes et al. 2011

  - 30 students took the test
  - Panel of fifteen experts
  - A statistically significant difference was found between the scores of the experts and the novices
  - Score’s reliability was high
Application to PT practice...

1. Residency & Fellowship
   – Predictor of Success?
   – Measure of outcomes
     • Pre-Test & Post Test
   – Niche Testing
     • Safety Testing (i.e. VBI Screening)
     • Identification of deficiencies
   – Testing of Residency-Fellowship Instructors
     • Are their reasoning abilities at the level of Expert
   – Comparing the outcomes of different programs?
Application to PT practice...

2. Specialist Certification
   – Current testing is all MCQ
   – Practical component would be useful, but it’s logistically impossible
   – SCT can be done on-line...
   – Could be a complementary component and supplement the current formats
Application to PT practice...

3. Continuing Education
   - The need to assess the effectiveness of CE is increasing
   - SCT is easy:
     • To create
     • To administer
     • To score....
     • Can be adapted to any clinical area of expertise
   - Could provide great feedback to instructors...
Application to PT practice...

4. Teaching
   – Fantastic tool
   – No ‘absolute’ right and wrong in the classroom
     • Opens up rich discussions with no ‘threat’
     • Decrease Likert Scale to 3-point from 5-point
   – Can be used as an examination-assessment method
   – Great for identifying areas for development of educational focus.
For now...

• The next step...
  – Publish a concept paper of some kind
  – Dissertation
    • Use of the SCT to determine level of concordance between the novice and expert manual therapist in the case of screening for vertebral artery insufficiency in patients with neck pain
  – Hopefully a greater interest and spread of the technique in the PT community
  – Collaboration
Questions are guaranteed in life; Answers aren't.
References

References

References

References


