Supplemental Digital Appendix 1

Using SRL Microanalysis to Identify and Remediate Specific Test-taking Deficiencies

Prior to utilizing this protocol, faculty should consider whether secondary issues such as life distractions, medical illness, substance abuse, learning disabilities, or personality disorders may be affecting the underperforming medical learner.[1] Addressing these secondary issues is outside the scope of this protocol.

The first test-taking deficiency is *lack of script recognition*, which is a problem utilizing the task strategy of disease scripts, and probably involves both the forethought (strategic planning) and performance phases (task strategies) of SRL. Historical features consistent with this problem include a below-average fund of knowledge across multiple assessments and difficulty completing tests within the allotted time. During the SRL microanalytic protocol, a learner who lacks script specificity offers no or factually incorrect interpretations of clinical data while reading the clinical stem and cannot answer Items 1-4 on the QRF. These learners often re-read the question or parts of the question multiple times, which may explain the struggle with time management. Because they cannot arrive at the correct diagnosis based on the clinical stem, these learners may use the answers to get an idea about what the diagnosis may be. For example, if the clinical stem presents a case of community-acquired pneumonia, the learner may not recognize the disease script but may conclude that it must be an infection of some kind when he sees the different antibiotics in the multiple choice answers. This approach is problematic; while it may help on tests, it is not how one practices clinical medicine, as patients do not present with a list of 5 possible treatment options. The solution for this learner is to study diseases in the context of clinical presentation, engage test questions in terms of differential diagnosis from the first 1-2 sentences, cover up answers until he has selected and justified the diagnosis, and utilize

items 1-4 on the QRF while doing practice questions to ensure that he follows these steps. See

https://youtu.be/PiEsywpnDCg for an example of a learner who lacks script recognition. (Note: The learners depicted in these videos are volunteer actors reading scripts developed from the authors' experience with actual struggling learners.)

The second type of test-taking deficiency is *lack of script specificity*, which also represents a problem using the task strategy of disease scripts for clinical diagnosis. This is a learner who can recognizing a general disease category (such as depression) or but not the specific clinical subtypes or severity of disease (such as depression in the elderly, depression with manic features, or postpartum depression) and the differences in diagnostic strategy, therapy, or prognosis that accompany these different disease subtypes. Historical features consistent with this deficiency include the learner's report that he can usually narrow down the answer choices to two, both of which may be appropriate choices for management of the general disease category, and then has to guess because he doesn't have disease scripts that are well-developed or specific enough to select the correct answer for the clinical scenario at hand. This learner will struggle to give complete and accurate answers to Items 2-4 on the QRF. The solution for this learner is to enrich and sharpen his disease scripts beyond just general categories to include the many different clinical subtypes and levels of severity of disease and to link these different clinical subtypes with the appropriate diagnostic, therapeutic, and prognostic considerations. When doing practice questions, the learner should focus on Items 2-4 on the QRF and seek every opportunity to increase his clinical exposure to enrich his disease scripts. See

https://youtu.be/nGzeO3pzXPg for a video example of a learner who lacks script specificity.

The third test-taking deficiency is *premature closure*, a metacognitive monitoring problem. This learner makes an early decision on the diagnosis and ignores or downplays

information inconsistent with the diagnosis. These learners may list multiple clinical findings on

QRF Item #4 that are inconsistent with their chosen diagnosis but refuse to change the diagnosis

– or this learner may fail to acknowledge the inconsistencies at all. This learner may report

finishing tests quickly but achieving low scores. Premature closure is evident during the SRL

protocol by the learner's answer to Item 4. The solution for this learner is to approach the

question in a slower, more deliberate manner, use the QRF when doing practice questions, and

focus on Item 4 as a checkpoint beyond which he may not pass until he has ascertained that his

selected diagnosis explains all markedly abnormal findings in the clinical stem. See

https://www.youtube.com/watch?v=lR_zC5QoFLY for an example of a learner with premature

closure.

The fourth test-taking deficiency is *underconfidence*, which is a metacognitive

monitoring and self-evaluation problem. Learners with this deficiency may have a history of

several examination failures and have consequently learned to distrust their clinical reasoning,

even when it may be correct. During the SRL protocol, these learners characteristically underrate

their confidence on Items 5-6 and 8, even when they have identified the correct diagnosis and

predicted the correct answer. An underconfident learner may predict the correct answer and

systematically talk himself out of it after seeing the other answer choices. The solution for these

learners is to improve their calibration accuracy, which is a measure of how well the learner's

confidence matches their abilities.[2] A simple formula can be used to calculate calibration

accuracy:

$$\text{Calibration accuracy} = 4 - |C - A|$$

Here, C is the confidence score from the Likert scale, with 1 for "not at all confident" and 5 for

"extremely confident," and A=5 for correct answers and A=1 for incorrect answers. The score

ranges from zero (total inaccuracy) to four (total accuracy). The solution for this learner is to use the QRF when doing practice questions, collect calibration accuracy data on himself, and re-calibrate his confidence with his abilities. Simple exercises such as this have been shown to improve calibration accuracy in other medical education settings.[3] See

https://youtu.be/v_maDtOXE8Y for a video example of the underconfident learner.

The fifth test-taking deficiency is inappropriate causal attributions, which is a deficiency in the self-reflection phase of SRL. Historical features that may suggest this deficiency include the learner's report of test failure despite completing hundreds and hundreds of practice questions. This deficiency is evident from superficial or inappropriate answers to QRF Item 14, where learners cannot describe why the right answer is correct and the wrong answers incorrect. A learner with this deficiency may justify correctly selecting levofloxacin as the appropriate treatment for community acquired pneumonia requiring hospitalization with a statement like, "That's what we always do," rather than identifying the specific pathogens that should be covered and affirming that levofloxacin has that spectrum of activity. When answering the question incorrectly, the learner may attribute his error to a poorly worded test question rather than considering where his own knowledge gaps lie. The solution for this learner is to use the QRF and focus on item 14. When he answers a question correctly, he must ensure that he is right for the right reasons and that he knows why each of the wrong answers is wrong. This means slowing down the pace of practice questions considerably in order to allow for in-depth reflection and consideration of knowledge gaps revealed by both correct and incorrect answers. See https://youtu.be/WP0i3bOHCfA for an example of a learner who makes inappropriate causal attributions.

The last test-taking deficiency is inappropriate adaptive inferences, which also is a deficiency in the self-reflection phase of SRL. This learner may be able to identify knowledge gaps after answering a practice question, but is unable to articulate an effective learning plan (QRF Item 15). A common response to Item 15 that indicates inappropriate adaptive inferences is, "I need to read more about [disease X]." In the learner's defense, he may be merely repeating what he may have been told repeatedly throughout his medical education. All physicians should read more, but this alone is an inefficient approach to improving test performance. Each practice question should reveal specific knowledge gaps for the learner, identified in Item 15. The learner's task is to then develop a specific, durable, realistic study plan to address that deficiency. The teacher should ensure that the study plan is targeted at the specific knowledge gap (a learner who identifies the clinical presentation of the disease but selects the wrong therapy should not be spending time reading about the clinical presentation of disease) and that the study plan involves meaningful interaction with the content, the ability to organize the content in a way that makes sense to the learner, and elaboration upon the content to facilitate memory encoding. Study plans should allow for easy review of specific points that were missed (e.g., single notebook, flashcards, mobile or web-based note-taking applications). The learner should use the QRF when doing practice questions and utilize Item 15 and teacher feedback to develop targeted plans to address identified deficiencies. See https://youtu.be/5-QAzP_sgJY for a video example of a learner who makes inappropriate adaptive inferences.

In our experience implementing this protocol, most learners display one predominant test-taking deficiency and we counsel them on how to implement the appropriate SRL-based remediation strategy as outlined above. However, in the case where a learner displays features of more than one test-taking deficiency, we recommend that these students focus on correcting one

deficiency at a time with the appropriate SRL-based remediation strategy. In these cases, longitudinal follow up and reassessment with the protocol would be appropriate to ensure improvement in one test-taking deficiency and identifying the next target deficiency.

References

1.   Lucey CR, Boote RM. Working with Problem Residnets: A Systematic Approach. In: Holmboe ES, Hawkins RE, eds. *Practical Guide to the Evaluation of Clinical Competence*. 1st ed. ed. Philadelphia, PA: Mosby, Inc.; 2008:205.

2.   Chen PP. Exploring the accuracy and predictability of the self-efficacy beliefs of seventh-grade mathematics students. *Learning and Individual Differences.* 2002;14(1):77-90.

3.   Leggett H, Sandars J, Burns P. Helping students to improve their academic performance: a pilot study of a workbook with self-monitoring exercises. *Med Teach.* 2012;34(9):751-753.