Dudek N, Gofton W, Rekman J, McDougall A. Faculty and resident perspectives on using entrustment anchors for workplace-based assessment. *J Grad Med Educ*. 2019;11(3):287–294.

Supplemental – Interview Guides

Phase One – Faculty

Section 1

Workplace based assessments tools, such as In-training Evaluation Reports (ITERs), Daily Encounter Cards (DECs) and the mini-CEX (Mini Clinical Evaluation Exercise) are used by clinical supervisors to document their judgment of how a resident completed a particular clinical task. They typically have several items that are to be assessed using a common rating scale that has descriptive anchors. One common scale is as follows (interviewer to hand out sheet with Rating Scale A).

Rating scale A
1 – rarely meets expectations
2 – inconsistently meets expectations
3 – meets expectations
4 – sometimes exceeds expectations
5 – consistently exceeds expectations

What works well when using this scale? Why?

What does not work well when using this scale? Why?

Possible prompts: Do you find it easy to use? Does it allow you to record your impressions accurately? Why or why not?

Section 2

Another common rating scale for WBA tools is as follows (interviewer to hand out sheet with Rating Scale B).

Rating scale B
1 – poor
2 – fair
3 – good
4 – very good
5 – excellent

What works well when using this scale? Why?

What does not work well when using this scale? Why?

Possible prompts: Do you find it easy to use? Does it allow you to record your impressions accurately? Why or why not?

Section 3

A new rating scale has been developed to use on WBA tools (interviewer to hand out rating scale C).

Rating scale C
1 – requires complete guidance; "I had to do"
2 – able to perform but requires repeated direction; "I had to talk them through"
3 – some independence but intermittent prompting required; "I had to direct them from time to time"
4 – independent for most things but requires assistance for nuances; "I had to be there just in case"
5 – complete independence; "I did not need to be there"

What do you think of this scale?

How is it different from other rating scales you have used/seen for residency assessment?

Do you see potential advantages for this scale compared to rating scales A & B?

Do you see potential disadvantages for this scale compared to rating scales A & B?

Section 4
We would like to explore your thoughts on how some clinical performances would be rated on the various scales. (Interviewer to hand out example scenario A)

Example scenario A
PGY2 trainee working in a clinic is functioning beyond what one would expect with respect to their physical exam skills for some of the patients seen in this clinic. Their physical exam skills on their daily assessment form are rated as follows:

Scale A – 4 (sometimes exceeds expectations)
Scale B – 5 (excellent)
Scale C – 3 (some independence but intermittent prompting required)

What are your thoughts about the various ratings assigned?

If you had assigned the ratings would they have been the same? Why or why not?

Which rating scale would be easiest for you to use? Why?

What rating scale would be most useful for the program director? Why?

Which rating scale would be most useful for the resident? Why?

Do you think that the resident described in this scenario would accept the rating of 3 on scale C? Why or why not?

Dudek N, Gofton W, Rekman J, McDougall A. Faculty and resident perspectives on using entrustment anchors for workplace-based assessment. *J Grad Med Educ*. 2019;11(3):287–294.

Section 5

Let us consider another example scenario. (Interviewer to hand out example scenario B)

Example scenario B

PGY4 trainee working on a general surgery service is observed taking a consent for a laparoscopic cholecystectomy. As expected, he demonstrates proficiency with this task as a resident at his level should be able to obtain informed consent for a routine surgical procedure such as this one. He is thorough in his explanation of the procedure, clearly explains the risks and benefits in language that the patient understands and answers the patient's questions accurately. He does this in a time efficient manner and displays compassion for the patient who is clearly nervous, as they have never had an operation before. The clinical supervisor rates their performance of obtaining informed consent for this procedure as follows.

Scale A – 3 (meets expectations)
Scale B – 5 (excellent)
Scale C – 5 (complete independence)

What are your thoughts about the various ratings assigned?

If you had assigned the ratings would they have been the same? Why or why not?

Which rating scale would be easiest for you to use? Why?

What rating scale would be most useful for the program director? Why?

Which rating scale would be most useful for the resident? Why?

Do you think that the resident described in this scenario would accept the rating of 3 on scale A? Why or why not?

Section 6

Let us consider a final example scenario. (Interviewer to hand out example scenario C)

Example scenario C

PGY4 trainee in rheumatology (i.e. in the first part of their rheumatology fellowship). They are working in clinic and need to aspirate a swollen knee joint. The resident is aware of sterile technique and knows what tests the aspirate samples should be sent for. However, they required direction on patient positioning and landmarking, as well as size of needle to use. You would have expected that a resident at their level would know and be able to complete these technical portions of the procedure. You rate their performance on this procedure as follows:

Scale A – 2 (inconsistently meets expectations)
Scale B – 2 (fair)
Scale C – 2 (able to perform but requires repeated direction)

What are your thoughts about the various ratings assigned?

If you had assigned the ratings would they have been the same? Why or why not?

Which rating scale would be easiest for you to use? Why?

What rating scale would be most useful for the program director? Why?

Which rating scale would be most useful for the resident? Why?

Do you think that the resident described in this scenario would accept the rating of 2 on scale A, on scale B, on scale C? Why or why not?

Section 7
Do you have any additional thoughts that you would like to share on the various types of rating scales used for WBA?

Phase One - Residents

Section 1
Workplace based assessments tools, such as In-training Evaluation Reports (ITERs), Daily Encounter Cards (DECs) and the mini-CEX (Mini Clinical Evaluation Exercise) are used by clinical supervisors to document their judgment of how a resident completed a particular clinical task. They typically have several items that are to be assessed using a common rating scale that has descriptive anchors. One common scale is as follows (interviewer to hand out sheet with Rating Scale A).

Rating scale A
1 – rarely meets expectations
2 – inconsistently meets expectations
3 – meets expectations
4 – sometimes exceeds expectations
5 – consistently exceeds expectations

What works well when you receive an evaluation using this scale? Why?

What does not work well? Why?

Section 2
Another common rating scale for WBA tools is as follows (interviewer to hand out sheet with Rating Scale B).

Rating scale B
1 – poor

2 – fair
3 – good
4 – very good
5 – excellent

What works well when receiving an evaluation using this scale? Why?

What does not work well? Why?

Section 3
A new rating scale has been developed to use on WBA tools (interviewer to hand out rating scale C).

Rating scale C
1 – requires complete guidance; "I had to do"
2 – able to perform but requires repeated direction; "I had to talk them through"
3 – some independence but intermittent prompting required; "I had to direct them from time to time"
4 – independent for most things but requires assistance for nuances; "I had to be there just in case"
5 – complete independence; "I did not need to be there"

What do you think of this scale?

Do you see potential advantages for this scale compared to rating scales A & B?

Do you see potential disadvantages for this scale compared to rating scales A & B?

Section 4
We would like to explore your thoughts on how some clinical performances would be rated on the various scales. (Interviewer to hand out example scenario A)

Example scenario A
PGY2 trainee working in a clinic is functioning beyond what one would expect with respect to their physical exam skills for some of the patients seen in this clinic. Their physical exam skills on their daily assessment form are rated as follows:

Scale A – 4 (sometimes exceeds expectations)
Scale B – 5 (excellent)
Scale C – 3 (some independence but intermittent prompting required)

What are your thoughts about the various ratings assigned?

If you had received these ratings how would you feel? For A? For B? For C?

Which rating scale would be most useful for your future learning? Why?

Do you think that the resident described in this scenario would accept the rating of 3 on scale C? Why or why not?

Section 5

Let us consider another example scenario. (Interviewer to hand out example scenario B)

Example scenario B

PGY4 trainee working on a general surgery service is observed taking a consent for a laparoscopic cholecystectomy. As expected, he demonstrates proficiency with this task as a resident at his level should be able to obtain informed consent for a routine surgical procedure such as this one. He is thorough in his explanation of the procedure, clearly explains the risks and benefits in language that the patient understands and answers the patient's questions accurately. He does this in a time efficient manner and displays compassion for the patient who is clearly nervous, as they have never had an operation before. The clinical supervisor rates their performance of obtaining informed consent for this procedure as follows.

Scale A – 3 (meets expectations)
Scale B – 5 (excellent)
Scale C – 5 (complete independence)

What are your thoughts about the various ratings assigned?

If you had received these ratings how would you feel? For A? For B? For C?

Which rating scale would be most useful for your future learning? Why?

Do you think that the resident described in this scenario would accept the rating of 3 on scale A? Why or why not?

Section 6

Let us consider a final example scenario. (Interviewer to hand out example scenario C)

Example scenario C

PGY4 trainee in rheumatology (i.e. in the first part of their rheumatology fellowship). They are working in clinic and need to aspirate a swollen knee joint. The resident is aware of sterile technique and knows what tests the aspirate samples should be sent for. However, they required direction on patient positioning and landmarking, as well as size of needle to use. You would have expected that a resident at their level would know and be able to complete these technical portions of the procedure. You rate their performance on this procedure as follows:

Scale A – 2 (inconsistently meets expectations)
Scale B – 2 (fair)
Scale C – 2 (able to perform but requires repeated direction)

What are your thoughts about the various ratings assigned?

If you had received these ratings how would you feel? For A? For B? For C?

Which rating scale would be most useful for your future learning? Why?

Do you think that the resident described in this scenario would accept the rating of 2 on scale A, on scale B, on scale C? Why or why not?

Section 7
Do you have any additional thoughts that you would like to share on the various types of rating scales used for WBA?

Phase Two – Faculty

Section 1
Workplace based assessments tools, such as In-training Evaluation Reports (ITERs), Daily Encounter Cards (DECs) and the mini-CEX (Mini Clinical Evaluation Exercise) are used by clinical supervisors to document their judgment of how a resident completed a particular clinical task. They typically have several items that are to be assessed using a common rating scale that has descriptive anchors. One common scale is as follows:

Rating scale A (Interviewer to give hard copy of scale to participant)
1 – rarely meets expectations
2 – inconsistently meets expectations
3 – meets expectations
4 – sometimes exceeds expectations
5 – consistently exceeds expectations

Do you have experience using this type of rating scale? If no, move to section 2.

If yes:
What works well when using this scale? Why?

What does not work well when using this scale? Why?

Do you find it easy to use? Why or why not.

Does it allow you to record your impressions accurately? Why or why not?

Section 2
Another common rating scale for WBA tools is as follows

Rating scale B: (Interviewer to give hard copy of scale to participant)

1 – poor
2 – fair
3 – good
4 – very good
5 – excellent

Do you have experience using this type of rating scale? If no, move to section 3.

If yes:
What works well when using this scale? Why?

What does not work well when using this scale? Why?

Do you find it easy to use? Why or why not.

Does it allow you to record your impressions accurately? Why or why not?

Section 3
A new rating scale has been developed to use on WBA tools as follows:

Rating scale C (Interviewer to give hard copy of scale to participant)

1 – requires complete guidance; "I had to do"
2 – able to perform but requires repeated direction; "I had to talk them through"
3 – some independence but intermittent prompting required; "I had to direct them from time to time"
4 – independent for most things but requires assistance for nuances; "I had to be there just in case"
5 – complete independence; "I did not need to be there"

Are you familiar with a tool that uses this rating scale (or very similar)? If yes, please tell me about your experiences using it.

What do you think of this scale?

Do you find that there are advantages to using this scale compared to rating scales A & B?

Do you find that there are disadvantages for this scale compared to rating scales A & B?

Do you find that you use more ratings on this scale compared to scales A & B? If yes, why do you think that is? If no, why do you think that is?

Which rating scale is easiest for you to use? Why?

What rating scale do you think is most useful for the program director? Why?

Which rating scale do you think is most useful for the resident? Why?

In your experience, do resident accept low ratings on scale C? Can you describe experiences where they did accept them? Where they did not accept them?

Section 4
In this section we explore the thoughts of the participants regarding the results from the phase one analysis. The results will contain the theoretical explanations that the phase one participants provided regarding why the various rating scales do or do not work well for resident assessment. The interviewer will present each result and ask the participant if they agree or disagree and why?

Section 5
Do you have any additional thoughts that you would like to share on the various types of rating scales used for WBA?

Phase Two – Residents

Section 1
Workplace based assessments tools, such as In-training Evaluation Reports (ITERs), Daily Encounter Cards (DECs) and the mini-CEX (Mini Clinical Evaluation Exercise) are used by clinical supervisors to document their judgment of how a resident completed a particular clinical task. They typically have several items that are to be assessed using a common rating scale that has descriptive anchors. One common scale is as follows:

Rating scale A (Interviewer to give hard copy of scale to participant)

1 – rarely meets expectations
2 – inconsistently meets expectations
3 – meets expectations
4 – sometimes exceeds expectations
5 – consistently exceeds expectations

Do you have experience being evaluated using this type of rating scale? If no, move to section 2.

If yes:
What works well when you receive an evaluation using this scale? Why?

What does not work well? Why?

Section 2
Another common rating scale for WBA tools is as follows

Rating scale B: (Interviewer to give hard copy of scale to participant)

1 – poor
2 – fair
3 – good
4 – very good
5 – excellent

Do you have experience being evaluated using this type of rating scale? If no, move to section 3.

If yes:
What works well when receiving an evaluation using this scale? Why?

What does not work well? Why?

Section 3
A new rating scale has been developed to use on WBA tools as follows:

Rating scale C: (Interviewer to give hard copy of scale to participant)

1 – requires complete guidance; "I had to do"
2 – able to perform but requires repeated direction; "I had to talk them through"
3 – some independence but intermittent prompting required; "I had to direct them from time to time"
4 – independent for most things but requires assistance for nuances; "I had to be there just in case"
5 – complete independence; "I did not need to be there"

Are you familiar with a tool that uses this rating scale (or very similar)? If yes, please tell me about your experiences being evaluated using it.

What do you think of this scale?

Do you find that there are advantages to being evaluated using this scale compared to rating scales A & B?

Do you find that there are disadvantages to being evaluated using this scale compared to rating scales A & B?

Do you find that you receive more variation in the ratings on your evaluations compared to scales A & B? If yes, why do you think that is? If no, why do you think that is?

Which rating is most useful for you? Why?

In your experience, have you sometimes received lower ratings on this scale compared to either scale A & B? If yes, how do you feel about that? Why?

Section 4

Dudek N, Gofton W, Rekman J, McDougall A. Faculty and resident perspectives on using entrustment anchors for workplace-based assessment. *J Grad Med Educ*. 2019;11(3):287–294.

In this section we explore the thoughts of the participants regarding the results from the phase one analysis. The results will contain the theoretical explanations that the phase one participants provided regarding why the various rating scales do or do not work well for resident assessment. The interviewer will present each result and ask the participant if they agree or disagree and why?

Section 5

Do you have any additional thoughts that you would like to share on the various types of rating scales used for WBA?