

**COLLEGE OF POLICING**  
**Independent Peer Review of**  
**Initial Recruit Online Assessment**



**October 2020**

# CONTENTS

<b>Section</b>	<b>Page No.</b>
<b>Overview</b>	3
<b>Executive Summary</b>	4
<b>Review of Exercises</b>	6
<b>Review of Candidate Communications</b>	14
<b>Marking Explanation</b>	16
<b>Review of Assessor Training &amp; Materials</b>	17
<b>Review of a Sample of Candidates</b>	21

## OVERVIEW

The College of Policing commissioned a Peer Review of the Initial Recruit Online Assessment process, with a specific focus on fair selection and inclusion.

This report outlines the key findings from a review of the assessor and candidate material and provided by the College. The material was reviewed for content, potential bias, language that may put candidates off, transparency and compared to general best practice guidelines in fair selection.

The report is structured into section headings which detail the key documentation/data that was reviewed, along with a traffic light system of:

- ✓ green (what works well),
- amber (minor improvements/tweaks could be made), and
- red (some important or larger changes need to be made).

We have not provided a separate list of recommendations as we have outlined them throughout the report. We have also made track changes and comments on the documents themselves, for ease, so that the College can see specifically where ideas and comments for improvements have been made.

## EXECUTIVE SUMMARY

Our independent review of the Initial Recruit Online Assessment concluded that overall, the College has produced a valid and objective set of selection exercises/material incorporating many aspects of best practice, and which provide clear guidance for both candidates, assessors, and quality assurers. Specific areas of best practice and recommendations for improvement have been identified throughout. It was evident that the exercises need only minor changes to some the language used, demonstrating that the CoP team have developed their learning about inclusive language in assessment.

- A review of the candidate exercise material revealed that overall, the exercises appear to be face valid and well written, in plain, simple English. We did not identify any source of bias or unfair discrimination that may negatively impact on different groups of candidates, however, we did make recommendations for some 'tweaks. In particular, it is recommended that the Scenario 1 answer options within the SJT are reviewed again as this may be having a small negative impact on some candidates.
- Minor changes for more inclusive language were suggested for the written exercise, particularly for replacing the term [REDACTED]. We did not believe this would affect outcomes for candidates but it could potentially make the exercise more inclusive.
- The online candidate assessment experience was positive – it was easy to navigate the system, with clear instructions and practice to get used to the recording aspect. There is provision throughout for candidates to get in touch with a technical support team if they are experiencing any technical or internet issues.
- We recommend a review of the automatic fail criteria, particularly in relation to [REDACTED], [REDACTED], [REDACTED]; and 'other'. From the data it was evident that the QACO overturned sometimes half of the automatic fail decisions, (which is a good double-check) so further examples and guidance may need to be provided for assessors.
- The candidate communications provide a good, detailed overview of what a candidate would be doing on the assessment, and provides answers to frequently asked questions. We recommend that a 'Candidate Guide' for the Competency/Values Based Interview' is written which outlines how to structure interview answers using the STAR format.
- In the future, the College may want to consider using a four-point rating scale for the exercises rather than a five-point scale. We would also recommend reversing the current scale as it is counter-intuitive.
- The assessor online training appears to be thorough and detailed, with knowledge checks and mock scoring exercises to undertake. We would recommend that assessors have to go back to re-read material if they get a knowledge check question incorrect.

- The quality assurance procedure appears to be excellent, and a 10% sample is in line with best practice. We were not able to comment on how easy it was for assessors/QA'ers to navigate through the LaunchPad system.
- Reasonable adjustments are made where requested and approved. The online assessment is compatible with screen reading software which is in line with best practice. Potentially, provision could be made for those candidates who do not have access to the internet/a computer (or even quiet space) to use a police force building in order to take the assessment.
- Competence in spelling, formatting and grammar does not form part of the assessment criteria and this is consistently mentioned to both the assessor and candidate.

We wondered if doing an online written assessment may negatively impact more BAME candidates whose first language is not English, compared to White candidates, or those with no keyboard/typing skills. However, ample time is given to complete the exercise so this should mitigate any potential effect. The data for the written exercise shows only a small difference between white and BAME candidate scores for this exercise.

- A review of a (very small) sample of candidates' interviews/written exercises/briefing exercises on the LaunchPad system found that the videos appeared to be of good quality, with clear audio and visuals. The guidance for assessors made it relatively easy to navigate the system in order to view videos/text and assign ratings.

## REVIEW OF EXERCISES

### Interview - Full instructions no screening

- ✓ This introduction to the interview was all clear and transparent. It included what CVFs were being assessed.
- ✓ The candidate is provided with tips before they start the interview, such as ensuring that they are in a quiet room, having a strong internet connection etc. This appears at the beginning of every exercise which serves as a good reminder to candidates to allow them to perform at their best.
- We recommend having a Candidate Guide to competency-based interviews (which includes how to structure answers) to help level the playing field. Some candidates will have had these sorts of interviews before, and others not so it would be useful positive action measure but available to all.

### Interview - Questions & Behavioural Checklist

- ✓ Overall, the interview questions are well written, with useful probing questions provided for candidates to structure their answer. They are in line with best practice for competency-based interviews. There was no sign of bias or unfair discrimination within these questions.
- For Integrity [REDACTED], it would be interesting to see if there are any significant differences between white and BAME candidates' scores for this question (if that data is available): "[REDACTED]"

There may be cultural differences in the interpretation of the word '[REDACTED]' in this question. The CoP could instead phrase the question as '[REDACTED]' [REDACTED].

### Interview - Explanatory Notes

- ✓ This document provides a list of all competences, behaviours and values being measured, and provides examples of what might be assessed as a 1, 3, or 5. It is very useful for the assessor for benchmarking purposes and for clarifying the difference between each rating. It is in line with best practice.
- ✓ The notes are clear, and written in plain English.
- The rating scale with 1 being the highest rating and 5 being the lowest is counter-intuitive (this is for all the exercises and QA documents). It is recommended that the rating scale is reversed to fit in line with best practice.
- P2 [REDACTED] reads "[REDACTED]" It is recommended that this is re-worded this to "[REDACTED]"



- ✓ The 'timer' and what it is used for is well explained.

## **Situational Judgement Test – Practice Scenario**

- ✓ This scenario is useful for candidates to get used to what they are about to undertake in the SJT and is in line with best practice.
- ✓ The practice scenario is not assessed.

## **SJT – Questions**

- ✓ There is good 'sign-posting' of the CVFs and what competency/value the scenarios relate to. This demonstrates transparency in the selection process.
- Scenario 1 lists options A-D as possible answers to choose from. When reviewing these, it was evident as an assessor why choosing option C would score most highly. However, it was less transparent why options A and D were scored as 2 points and 1 point respectively.

Some candidates from BAME cultures (and non BAME, including women) [REDACTED]. There may be a possibility that this impacts some candidates more than others so it would be worth 'deep diving' into this further and reviewing the average ratings for this scenario for men/women/white/BAME candidates (and any other group that you have data for) to see if there are any significant differences. It is not clear to me as an assessor why option A is any different to option D (in terms of the total marks that should be awarded).

A review of Cohen's D for Q1 in the SJT shows a small difference in the average ratings for White compared to BAME candidates. Whilst the difference is small, out of all of the 12 scenario questions, Q1 has the biggest observed difference between candidates. The answer options to this question are worth reviewing again by the CoP team.

- ✓ The remaining scenarios and answer options appeared to be transparent, clear, well written and easy to understand. There was no evidence of bias or unfair discrimination that we could observe.
- Scenario 10 makes reference to an '[REDACTED]' – in this context it appears fine but with modern inclusive language, we would recommend rewording this to '[REDACTED]'. This would be more inclusive language to use and in line with the latest terminology.

## **Situational Judgement Test – online candidate experience**

- ✓ On the welcome screen there is a picture of a female black officer. There is diversity in the images that are used which demonstrates inclusivity.
- ✓ The online experience of undertaking the SJT was that it was easy to navigate and all the questions appeared to be clear and jargon free. There was nothing within this exercise that we could observe that may account for the disproportionality in outcomes.

- ✓ A video test was undertaken as well as a voice test, which is in line with best practice.
- ✓ It was thought that there was sufficient time given to candidates to read and respond to the scenarios. This was with the view that some candidates may have English as an additional language and therefore may potentially need more time to translate and respond to scenarios, compared to others. However, the CoP has factored in sufficient time for this.
- On the negative side, having such a long time to respond to scenarios may mean that a candidate can get help from others. As this test is not in front of the camera, or being recorded, there does not appear to be any way of mitigating this risk.
- The instructions appeared on the screen however they were not read out verbally. It may be useful for those whose first language isn't English if the instructions could also be read out (as they are for the interview exercise).

As this is the first exercise in the process for some candidates, it could also be seen as a more welcoming first experience.

- There appeared to be no explanation to the candidate of what the 'blur my screen' function was for. This needs to be made transparent.





## **Briefing Exercise - Candidate Instructions**

- ✓ The candidate brief is clear, jargon free and written in plain English. Instructions appear clear and easy to read, and nicely spaced out for the candidate so as not to appear overbearing.
- ✓ There was nothing within this brief or exercise overall that was considered to favour white/BAME/men/women/older/younger candidates over others. We could not find any evidence that would suggest that the briefing exercise could lead to an unfair disproportionality in outcomes for different groups of candidates.
- ✓ The scenario questions are both read out and appear on the screen which is very helpful for candidates.
- ✓ A link to the Competency & Values Framework is provided for all candidates (throughout all of the exercises) which demonstrates transparency in the assessment process.
- ✓ The brief makes it clear to candidates that the online assessment process will not be measuring 'technical skills'. We were initially concerned that those with a lack of IT skills may find the online assessments difficult, but when looking at the assessment process holistically, there are many references made to both candidates and assessors that IT ability, formatting, spelling etc will not be part of any decision-making process.
- ✓ There is provision for a candidate to get in touch with a technical support team if they are experiencing any technical or internet issues. The instructions clearly say "please do not finish the exercises if you are experiencing technical issues."
- The brief stipulates that a strong internet connection is required. We wondered what adjustments were made for those who did not have access to such facilities or those who did not have access to the internet at all (perhaps for reasons to do with social mobility)? Is there another way that candidates can complete the online assessment, for example, by going into a local police force and doing the assessments there? Perhaps this provision could be made in the future if it was Covid compliant.
- It was not clear if the candidate was moved on to answering the questions after the preparation time had ended. We understand that the preparation time will not be assessed but it was unclear as to how long a candidate could spend preparing, and if there was an end point after which the exercise questions would automatically commence.

## **Briefing Exercise - Explanatory Notes**

- ✓ The explanatory notes for assessors were very helpful. They were written in clear, jargon free English and there appeared to be sufficient differentiation between the examples (and ratings) provided.

## Briefing Exercise - Questions

- ✓ A review of the questions asked during the briefing exercise indicated no evidence of any potential bias to any particular group of candidates.
- ✓ The questions appeared clear, easy to understand and jargon free.

## Briefing Exercise - Behavioural Indicators

- ✓ The behavioural indicators and examples provided appeared to be clear, justifiable, transparent and jargon free. There was nothing to suggest that they would disproportionately impact one group of candidates any more than another.

## Briefing Exercise - Online Candidate Experience

- ✓ The briefing exercise appears to have good face validity and resembles some of the work that a PC may encounter in their role. This is likely to be a good and realistic preview of the work that a PC may undertake in the future.
- At the start of the briefing exercise, the lady providing tips refers to 'the interview'. This was quite confusing as a candidate as it made me go back to my email folder and try the link again 4 times to check I was doing the correct exercise. This could unsettle candidates before they undertake their assessment, particularly those who are not as familiar with IT.
- There is a 'warm up' question that asks: " [REDACTED] [REDACTED]". This question sounded a little strange when candidates may [REDACTED] [REDACTED]. Perhaps a more appropriate question might be "[REDACTED]" or "[REDACTED]?"
- It was unclear what the 'blur my video' function was for. It was later explained well by the CoP in person but it is not explained to the candidate when the video/audio check is done at the start of the assessment. It may confuse candidates, and for transparency it is recommended that some reference is made to what it can be used for. We could not see any reference to it in the Q&A documents for candidates.

## REVIEW OF CANDIDATE COMMUNICATIONS

- ✓ This is a clear, understandable document outlining what the candidate will do and the system requirements they will need. It is jargon free and written in plain English.

### Candidate Communications - SJT (no stages)

- ✓ As above – this is a clear, understandable document outlining what the candidate will do and the system requirements they will need. It is jargon free and written in plain English.

### Candidate Journey (no stages)

- ✓ This is a really useful diagram which shows what the candidate needs to do from beginning to the end of the online assessment process, and which exercises to take in which order. It also provides an overview of the approximate time each exercise will take to complete.

### Online Assessment Process Candidate Declaration v1.1

- [REDACTED]  
[REDACTED] Does this also apply to [REDACTED]  
[REDACTED]  
it needs to be made explicit in this document. It is not as transparent for the candidate as it is for the assessor in the assessor training.

### Candidate Journey - Online Assessment Process Candidate Guide

- ✓ This is a very useful document outlining the entire assessment process to the candidate. It clearly outlines the candidate journey, the exercises they will undertake, how to prepare, and frequently asked questions. It is really helpful and transparent.
- ✓ P22 outlines reasonable adjustments that can be made. A discussion with the CoP also confirmed that the online assessment process is screen reader compatible (speech to text or text to speech). Extra time is provided for those candidates who are eligible, and use of a reading rule and coloured background/overlays can be accommodated. Other requests from candidates are reviewed and assessed based on a case by case basis and specialist recommendations. This is in line with best practice in assessment.
- ✓ P 10 – the CBI refers to being able to use examples from both your work and personal life. This is great practice as sometimes candidates have a lot of relevant experience from personal/outside of work examples such as community, voluntary or sporting initiatives.

- For the CBI it might be useful to outline that candidates should try to think of one example and describe it in detail for each question.
- P19 – How to plan for your attendance – it would be useful if the CoP could provide some guidance here about how to structure competency-based interviews (for all candidates) to ensure that candidates are on a level playing field as much as possible. They could be provided with information on STAR (situation, task, action, result) and how answers can be structured. These positive action measures can often help those from less affluent backgrounds, and also any BAME candidates who have never undertaken a competency-based interview before.
- ✓ There appeared to be a good, comprehensive list of Q and A's.

### **Online assessment process generic feedback report**

- This document looks comprehensive; however, it is very generic so we did not think that it would be hugely helpful for a candidate's development.
- A more bespoke document highlighting overall strengths and a couple of development points in relation to the competency and values being assessed would be more helpful. However, we understand that this would be an enormous and potentially costly task to undertake.

### **Candidate FAQ's**

- ✓ There is a detailed list of candidate questions with answers which is really helpful for those candidates with queries or questions. It will give some reassurance to candidates and is transparent, in line with best practice.

## MARKING EXPLANATION

- ✓ A Quality Assurance Team (QA) monitors the performance of the assessors by selecting a random sample of candidates and co-assessing the candidates' performance. This ensures standardisation, fairness and consistency.
- ✓ Some detailed consideration has been given to the cut off scores for the exercises, and various options were presented to a Board, where the most favourable outcome was chosen. Whilst we have not seen the data behind this or the different options, we are satisfied that due care and consideration has been given to this process.
- The cut off score for the **written and briefing exercises** in this document was a little confusing for the assessor and may require further clarity/explanation. Each of these exercises will have a total maximum of 20 but they are combined so that the cut offs are out of 40 (see page 5). It may be possible that a candidate scores really well on one exercise but not on the other and still gets through as they meet the 21/40 cut of threshold. It would be useful to understand the rationale for this.

## REVIEW OF ASSESSOR TRAINING & MATERIALS

### Assessor Online Training

- ✓ The assessor online training appears to be very thorough. It takes the assessor through the online assessment process, the role and responsibilities of an online assessor, assessor best practice, assessment principles and assessment criteria. Please note that we did not see or undertake the online assessor training, but read the written version of it. We were unsure how long the assessor training would take to complete.
- ✓ There are 'knowledge check' questions throughout the online training to check for understanding. This is in line with best practice in assessment.
- It would be useful if an assessor got a question wrong, that the online module takes them back to the relevant section within the assessor training to look at it again, before letting them proceed on to the next question.
- ✓ There is a benchmarking session which is used to assess a mock candidate against the criteria which is good practice to ensure consistency and calibration amongst assessors. Benchmarking training is also provided which is in line with best practice.
- P26 provides an example of indirect discrimination which could potentially be viewed as direct discrimination rather than indirect. An example of indirect discrimination could be: where we have stipulated a requirement which negatively impacts a group of people with a particular protected characteristic e.g. having a height requirement, or a minimum number of years' experience (unless of course the requirement can be justified as important for the role).
- ✓ P30-31 provides an explanation of reliability, validity and standardisation which is crucial for assessors to understand.
- ✓ A good and thorough explanation of unconscious bias has been provided as part of the learning. There are examples for expectancy effect, recency effect and contrast effect.
- ✓ There is a document that outlines different types of bias which is detailed and serves as a helpful reminder for assessors.
- ✓ An explanation of competency and values frameworks is provided in detail.
- ✓ Assessors are asked to assess a mock candidate. The CBI launchpad assessor user guides provide detailed and a step by step guide to entering ratings onto the system. We did not get the opportunity to try the system but it appeared be written in a logical and clear manner.
- ✓ All instances of automatic fail will be moderated, through a QA process, before being granted. This is an example of best practice and a good double check for objectivity in the process.
- ✓ Pages 22-24 outline common biases when giving ratings – such as central tendency, leniency and stringency effect, and similar to me (affinity) bias.

- ✓ Mock interview rating practice is undertaken, which we would expect to see in any best practice assessor training.
- ✓ There is a good explanation of what a 1 rating might look like. Clear explanations for benchmark ratings are given to the assessor once they have entered the ratings based on the mock interview video. These appeared to be bias free.
- P77 reads .... “ [REDACTED] ”. We would recommend changing the word “ [REDACTED] ” to ‘ [REDACTED] ’.
- ✓ There is a section called Considerations which outlines aspects such as the environment, technology, time management and preparation. This is very useful for the assessor.
- ✓ Assessors are asked not to have multiple candidates’ assessments (for the written exercise) open at the same time but instead to review one and assess it fully before then opening up another candidates’ exercise. This will help prevent the ‘contrast effect’ bias.
- ✓ Assessors are asked to read the whole written script (for the Written Exercise) once in full before they go back through it and begin to identify the evidence and confirm their ratings. This is an example of best practice in assessing.
- ✓ There is capacity for assessors to make comments on Launchpad whilst classifying and evaluating evidence.
- It was not clear why a candidate may be asked to retake the interview but we assume for reasons to do with IT and video quality.
- ✓ The guidance recommends that assessors should allocate the same amount of time to mark each candidate. This is in line with best practice standards.

## **QACO User Guide**

- ✓ Quality assurance of the online assessment process takes place.
- ✓ QACOs receive training in the online assessment process and have completed at least two of the online modules. They will have performed the QACO role before at least twice in the SEARCH Day One Assessment Centres. They will also have performed the assessor role for any exercise in the Online Assessment Process for PC’s for a minimum of two days.
- ✓ QACOs are selected on the basis of both experience and other essential/desirable skills including empathy, interpersonal skills, being supportive and an understanding of diversity and inclusion.
- ✓ Each assessor is QA’d twice in an allocation.
- ✓ 10% of candidates will be QA’d which is a good sample size considering the scale of the operation and how many candidates are applying. A 10% sample is in line with best practice.

- ✓ Automatic Fails are quality assured and logged. The QA team must decide whether the evidence submitted by the assessor is sufficient to meet the criteria outlined in the Penalising Unacceptable Language & Behaviour Procedure regardless of their performance in the other exercises.
- ✓ Constructive feedback is provided to assessors and further support and training is provided to those who score a 4/5. If assessors continue to score 4/5's by a QACO, they may be moved off the exercise and asked to assess a different exercise instead.

## Review of Automatic Fail Criteria

- ✓ Assessors are permitted to award a candidate with an automatic fail (A-F) but these are always quality assured by a QACO. The A-F's need to be logged so that they are always QA'd.
- ✓ A review of the automatic fail criteria was undertaken. For 1 ( [REDACTED] ), the examples provided seemed perfectly appropriate and bias free.
- For 2 ( [REDACTED] ), we perceived that 2a and 2b [REDACTED] needs to be put into context. For example, some candidates may be nervous and [REDACTED] They may simply [REDACTED] We recommend that this is revised or the examples deleted (perhaps there could be an example [REDACTED] )
- For 3 ( [REDACTED] ) – we recommend including [REDACTED] as part of 3g.
- For 4 [REDACTED] – [REDACTED] . Some degree of common sense needs to be applied to these automatic fails to take into account a candidate who [REDACTED] is constant or repeated, this may warrant a fail for '[REDACTED]'. Category 4c also needs to be put into context and a sense check applied. [REDACTED] and this should be taken into consideration.
- Category 5 is too broad. It could lead to bias or unfair discrimination, however, there is quality assurance in place. We would question if this category is really needed at all for the online assessment process.
- A review of the A-F data for each exercise shows that almost half of the A-F decisions are overturned. This would suggest that more information needs to be provided to assessors with regards to what does and does not constitute an A-F, perhaps through the example of some scripted 'video' or 'written' evidence.
- There could be an unconscious bias occurring (such as the horns effect) if/when an assessor marks a candidate with an A-F. We recommend drawing this to the

assessors' attention during the assessor training, as something to be **consciously mindful** of.

- It is worth the CoP breaking down the number of A-F's per exercise further to see if there are any significant differences between males/females being awarded them, or any differences between older/younger candidates, or between white/BAME candidates. **If the sample sizes allow it**, this would be a useful piece of work and may help identify if there is more that the CoP can do to brief candidates about A-F's, or ask forces to sift candidates fully prior to coming to the online assessment.
- As automatic fails are being used, it seems only fair and transparent to inform candidates about them (without necessarily giving them all the behavioural indicators). There is insufficient information for candidates about this (it does not need to be as detailed as it is in the training for assessors).

## REVIEW OF A SAMPLE OF CANDIDATES

A review of a sample of nine candidates was undertaken (three interviews, three written exercises and three briefing exercises), with the aim of seeing how easy it was for assessors to undertake the task of assessing using the online system, how clearly candidates came across on the videos, how nervous/relaxed they appeared, and how easy it was to navigate the Launchpad system. We were to some extent also able to cross reference our (approximate) ratings given to a candidate with those that the assessor gave. However, it should be noted that we have not undertaken the relevant assessor training so do not have the level of expertise that the assessors would have for a given exercise, and we have only sampled a few candidates from a very large pool. Therefore, we cannot draw any reliable conclusions about ratings awarded.

- ✓ Launchpad was relatively easy to navigate and the guides for assessors really helped to identify what to do and which buttons to locate/press.
- ✓ For the majority of the video-based exercises that we observed i.e. the interview and the briefing exercise, the candidates came across clearly, the audio was clear, and the video quality was good. It was difficult to see one candidate's face as the video was quite dark, though her audio and answers to questions were perfectly clear. Sufficient guidance is provided to candidates about how to set up the camera/lighting angle, so we did not feel that there was any more that the College could do about this.
- ✓ One observation of a written exercise revealed that the candidate had appeared to have started the exercise, typed in a paragraph but did not continue any further. Potentially his/her internet could have gone down. When we queried what had happened and what provisions had been made for this candidate, it was evident that the College had offered the candidate a retake of this exercise, however, the candidate did not take up this offer. Retakes (based on this small sample for review) appear to be offered appropriately and for logical/evident reasons.
- ✓ The majority of candidates for the interview and briefing exercise videos appeared calm, and appeared to know how to follow the instructions and know what to do. Only one candidate from this small sample appeared to find the process daunting and struggled with her interview. She still appeared to know how to do the recordings - it was more that she found the questions difficult and appeared to struggle thinking of examples for the questions. Again, we felt that the College provides sufficient time for the interview so perhaps any material about competency-based interviews, to think of one specific example and the STAR approach would be useful to send to candidates as part of their preparation. It was evident that some had come across the STAR technique before whilst others had not so a guide may help to level the playing field.
- ✓ Overall, we were in agreement with the assessor ratings given for this small sample, but we cannot generalise any further. A look at the assessor's ratings provided for each of the candidates we observed appeared to be justifiable, (with the caveat that we have not undertaken exercise specific assessor training). Our ratings either matched or were 1 point different to the assessors for all ratings except for one.