Understanding your Situational Judgement Test (SJT) score

SJT as a measure of meeting the person specification
The SJT is designed to assess your approach to work as a Foundation doctor as defined by the FP 2019 person specification.

The SJT measures your understanding of situations that arise for a foundation doctor, your judgement in differentiating between appropriate and inappropriate responses, and your ability to recognise the most important concerns in any situation. It focuses on appropriate behaviour with respect to interacting with patients and colleagues and in managing your own workload. The test covers five core domains:
- commitment to professionalism
- coping with pressure
- effective communication
- patient focus
- working effectively as part of a team

How the papers are marked
There are 70 items in a test paper, 10 of which are trial items which are being piloted for future use. All of the 60 live items have been reviewed by experts and piloted to ensure that they are appropriate for use. Around two thirds of the items are ranking items and one third of the items are multiple choice. There is no negative marking. Please see the scoring grid on the UKFPO's FAQs for more information.

- **Ranking items**: The items in Part One of the paper require you to rank a series of options in response to a given situation. Your answers are marked by comparing your response to the model response determined by an expert panel. The closer your response is to the model response, the more marks you are awarded. A perfect match receives 20 marks. If you tie two options in your answer you will receive no credit for either option. This marking convention means that even if you were to answer a ranking item completely out of order, you would score a minimum of 8 marks.

- **Multiple choice items:** The items in Part Two of the paper require you to select three from eight possible responses to a given situation. As with ranking items, your answers are marked by comparing your response to the model response determined by an expert panel. These items are worth a maximum of 12 marks. Each of the three individual responses is worth 4 marks. If you select more than three options, you will receive no credit for the question.

Test equating

There were three test papers used for selection to FP 2019. The papers were balanced for content and difficulty; items were selected taking into account discrimination[1] and facility[2] to ensure that each version of the test was as close to equal quality and equal difficulty as possible. The three test papers covered an equal number of items assessing a similar number of each of the core domains, with the papers also designed to include a similar number of items according to a theme (i.e. colleague-related or patient-related). It is not, however, possible to give an SJT 'mark' by domain as one item could assess several different domains, and other domains are integral to all items, for example decision making.

Statistical procedures were used to make minor adjustments to marks on the different papers to ensure that all scores are equivalent – this is known as test equating[3].

How SJT points are determined

SJT scores are reported on a 0 to 50 scale with the lowest possible score receiving 0.00 points and the highest possible score receiving 50.00 points. The distribution of the scale is set to match the mean and distribution of the EPM scores. This ensures that when the SJT and EPM scores are combined they each have an equal weighting. Table 1 below shows how SJT scores were distributed in selection to FP 2019.

**Table 1: Distribution of Scaled SJT scores for FP 2019**

SJT points	Percentage of applicants
0.00–10.00	0.0%
10.01–20.00	0.2%
20.01–30.00	2.0%
30.01–35.00	8.3%
35.01–40.00	35.3%
40.01–45.00	47.2%
45.01–50.00	6.9%

---

[1] By examining the correlation between item score and overall test score, not including the item itself, this shows how well an item differentiates between high and low scorers.

[2] The facility value is the mean score achieved for the item. If the facility value is very low, then the item may be too difficult and may not yield useful information. If the facility value is very high, then the item may be too easy and may not provide useful information or discriminate between applicants.

[3] Statistical process of determining comparable scores on different forms of an assessment which takes into account the relative difficulty of items.

Figure 1: Distribution of Scaled SJT Scores for FP 2019
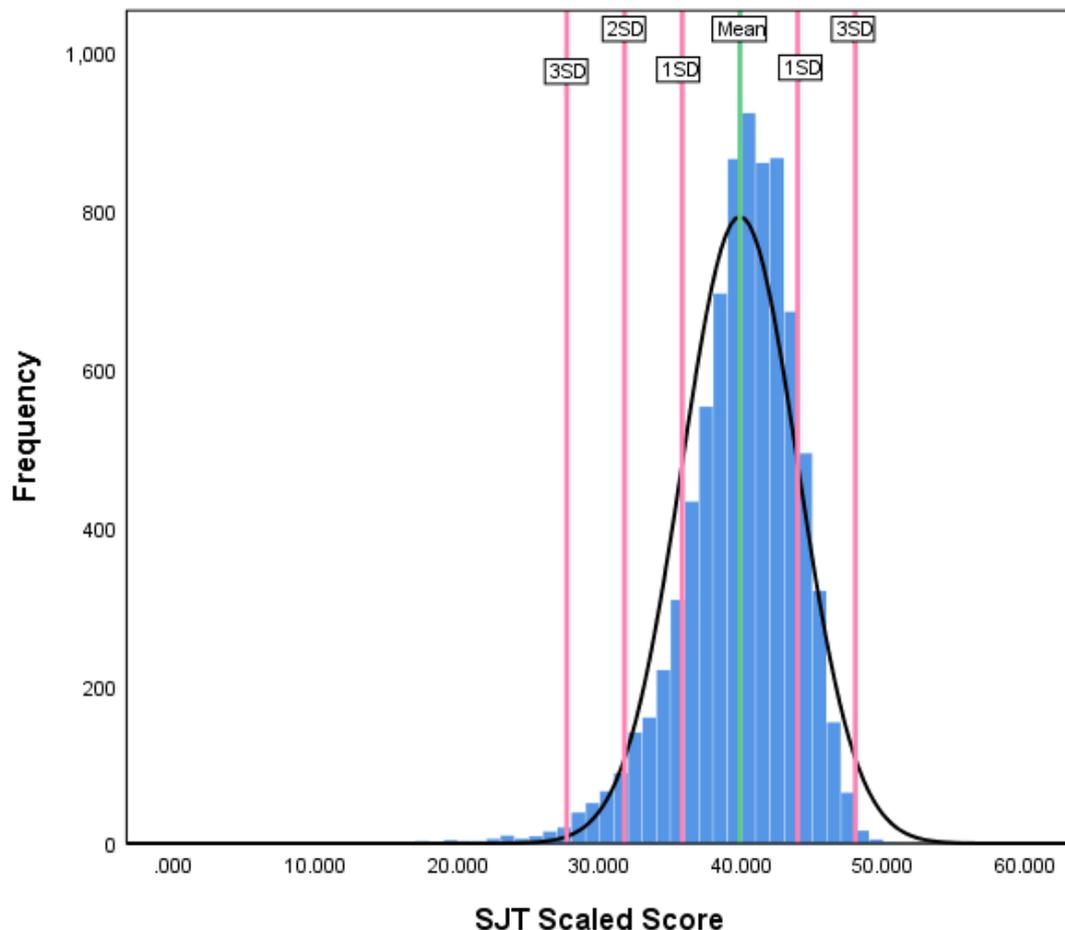
Mean: 39.89
Median: 40.45
Standard Deviation: 4.07

Standard Deviation (SD) provides an indication of the spread of scores. In a normal distribution, the majority (68%) of applicants would score within 1 SD from the mean; a further 27% would score 1–2 SDs away from the mean; 4% would score 2–3 SDs away from the mean; less than 1% would score more than 3 SDs from the mean. Figure 1 shows that the distribution of SJT scores was a near normal distribution with a slight negative skew, i.e. a longer tail of low-scoring applicants.

More information
The SJT has been developed over a number of years, with extensive piloting, clinician and psychometric review, to ensure that the test is fair, reliable and robust. For more information about the Department of Health funded *Improving Selection to the Foundation Programme (ISFP) Project,* the results and analysis of the pilot SJTs and further reading about the SJT, please visit www.isfp.org.uk.