Evaluation criteria

**FINAL EVALUATION BY FACE-TO-FACE INTERVIEWS**

Overview

Several committees will be formed according to the number of applicants called to interviews. Each of these committees will be formed by university professors or professional experts in the disciplines assessed.

To score candidates in the face-to-face assessment process, evaluators will use a qualification grid with three aspects to assess (see below); each will have a specific weight.

Score

Each aspect evaluated must be scored with one of the number on the following scale:

<table>
<thead>
<tr>
<th>Qualification</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exceptional</td>
<td>8</td>
</tr>
<tr>
<td>Excellent</td>
<td>7</td>
</tr>
<tr>
<td>Very Good</td>
<td>6</td>
</tr>
<tr>
<td>Good</td>
<td>5</td>
</tr>
<tr>
<td>Normal</td>
<td>4</td>
</tr>
<tr>
<td>Mediocre</td>
<td>3</td>
</tr>
<tr>
<td>Poor</td>
<td>2</td>
</tr>
<tr>
<td>Very poor</td>
<td>1</td>
</tr>
</tbody>
</table>
Aspects evaluated

1. **Candidate’s Potential**

*Summary:*

In order to have a general perception of the candidate’s potential, experts will pay attention to “soft” skills, such as clear, consistent discourse and articulation of ideas, ability to present complex reasoning, team working; and capabilities such as independent reasoning, originality, entrepreneurship, leadership.

This section will mainly evaluate:

- **Originality:** ability of the proposal to think outside the box, making creative proposals or going further in depth into unexplored areas. Originality of the proposal and of the formula to achieve objectives sought.

- **Innovation:** capacity to create new knowledge or, in the case of professional proposals, to open new routes or formulas for creating wealth. Use of new technologies or innovative use of existing technologies. Use of new theoretical approaches to existing phenomena or problems or innovative use of pre-existing theoretical approaches.

- **Viability:** The project is ambitious and realistic; the capabilities of the applicant correspond to the scope of the project.

- **Expository clarity:** ability to clearly and precisely present complex reasoning and very specific matters, so that the proposal can be understood by a layman. Suitable wording, which avoids the use of specialised vocabulary.

This section will be weighted at 40%.

2. **Motivation and Statement of Purpose**

*Summary: the impact of the candidate’s statement of purpose, understood in its broadest sense: capacity to contribute to transformation and improvement in areas such as the economy and the creation of wealth, society, culture, science, the quality of life of citizens, the environment or public policy.*

Experts will assess the impact of the statement of purpose for the candidate and the society; innovation, originality and feasibility; and candidate’s capabilities with regard to the scope of the statement.
This section will mainly evaluate:

- The interests presented by the candidate should be consistent and is well-structured, and proposed studies should enroll in a career of wider scope, either professional or academic. In this sense, the candidate shall state, and the evaluator qualify, to what extent the conduct of studies for which a grant is requested is a necessary step in the direction indicated.

- The candidate must justify the interest of the studies he/she wants to pursue, fitting center or centers where it is proposed to carry out such studies and international programmes and the country or countries selected.

- The candidate must also argue the need to carry out the proposed statement of purpose to achieve the objectives presented.

- It will be evaluated the progress of science and knowledge, wealth creation and possibility of transferring to others what they have learned during the studies as well as social return, in its broadest sense.

- The ideas presented should have a new and original character. Proposals that entail risk elements and creativity, whether hundred-typhus or business, especially the proposals that raise serious venture projects will be favoured. In this regard, the potential impact of the fellowship on the future career of the candidate will also be evaluated. The following will be taken into account:
  
  - The correspondence between the level of excellence of the candidate and studies proposed, educational centre or research team where they want to carry out their project.
  
  - The cost of scientific, social or economic opportunity estimated if the candidate could not perform the proposed studies.
  
  - The fact that the candidate has not previously received similar possibilities to those facilitated with the fellowship from “la Caixa”.
  
  - In international programmes, the fact that the candidate has not previously completed post-graduate studies in the same destination country for which the fellowship is being requested.

This section will be weighted at 30%.

3. Academic and Professional Background

Summary: the academic and professional background of the candidate in relation to the stage of the career they are in and opportunities, in this respect, they may have obtained.
This section will evaluate:

- Quality and depth of curriculum in relation to applicants’ possibilities. In this respect, younger candidates, therefore, cannot be penalised for accrediting incipient curricula.

- However, the scope, quality and depth of the activities accredited by the applicant (courses, attendance seminars, written and audiovisual publications, professional experience, etc.) will also be taken into account, as will the intellectual curiosity demonstrated to complete their curriculum.

- Consistent, focused trajectories throughout their entire course of studies. In the event that there are changes in their trajectory, they must be justified and rationalised.

- If the professional stages and, above all, academic stages are related to the study project presented and, if not, the reasons must be adequately explained.

- Not to penalise candidates who, because they were studying while working a steady job, may have gotten lower scores than other candidates.

- Expressly take into account the efforts shown by the candidate to overcome a difficult family situation, from a socioeconomic perspective.

This section will be weighted at 30%.

Prior expert scores

The marks obtained by the candidates during the remote assessment stage of the process will be taken into account in the face-to-face assessment stage. The standardised score from the remote assessment process will be added, as an expert score, to the other scores that face-to-face evaluators grant each application.

Similarly, all committee members will have access to the scores and evaluations assigned to each application by the evaluators having examined it in the remote assessment process. Likewise, they will have additional information on the position of the application within its remote assessment group, the number of applications evaluated within such group and the number of applications who pass the remote assessment process and any other information that the Fellowships Programme Office may consider relevant to the evaluation of applications.

To the extent they consider it appropriate, committee members may take into consideration such information when evaluating and scoring the candidates interviewed.

Content of the interview

The face-to-face interview will make it possible to detect, from more subjective, refined and subtle findings, the quality and consistency of the candidate evaluated.
During the interview evaluators may test the theoretical knowledge of the candidate, although it is not its main purpose. What is primarily sought is to check the soundness of the application, taking into account the above-mentioned aspects.

The face-to-face interview will seek to:

- Enlarge the information provided in the application, especially in relation to the candidate's academic and/or professional project.
- Ask about issues not mentioned in the application and which the committee considers relevant for evaluating the candidate's suitability to continue with the studies proposed.
- Evaluate the candidate's academic and/or professional potential.
- Evaluate their overall training, interests, concerns and curiosities for social, scientific, economic, cultural or artistic contexts, although not directly related to their studies.
- Assess their personal and academic maturity, motivation for carrying out the studies and the project proposed and the ability to express themselves clearly and convincingly defend the ideas expressed.

Formal aspects

- Interviews will last approximately 20 minutes.
- The candidate will not be told who the evaluators are that form part of the committee.
- The representative of the "la Caixa" Foundation will open the interview, which will begin in every case, by giving the floor to the applicant so that, within a couple of minutes, they can give a summary of their personal project. Then the other committee members will ask the questions they believe are relevant to properly assess the application.
- Interviewers should not ask questions about topics already reported in the application, unless they wish to clarify some aspect.
- Language: Interviews can be conducted in English or Spanish.
- Committees should endeavour to follow established schedules and be as punctual as possible with the candidates called.
- There is no established protocols for address (formal/informal) or if it is necessary or not to shake hands before or after the interview. These aspects are at the discretion of each committee or the spontaneity of the candidates. Nevertheless, the interviews should obviously be characterised by their seriousness, propriety and pertinence of the questions.
Final evaluation by Face-to-face interviews

PROCEDURES

Table of contents

1. Introduction ................................................................................................................................. 2
2. Grading ........................................................................................................................................ 2
3. Standardization ............................................................................................................................ 3
4. Detection of discrepancies and an alarms .................................................................................. 4
5. Weighting according to expertise with the discipline assessed .................................................. 4
   Computation of the final grade ...................................................................................................... 6
6. Presentation of the scores ............................................................................................................ 6
7. Reserve lists .................................................................................................................................. 7
8. Feedback for the candidates ......................................................................................................... 8
   Normalization (for each criteria) .................................................................................................. 8
   Quartile division ........................................................................................................................... 9
9. Annex 1 ......................................................................................................................................... 10
   About normalization of scores and its effects ............................................................................... 10
   Example ......................................................................................................................................... 10
1. Introduction

This document explains the procedure for evaluating and selecting candidates in the face-to-face stage by assessment committees.

The objective of this system is to ensure maximum efficiency and objectivity in the assessment of candidates, eliminating, as much as possible, the effect of bias in scorings and minimizing the possibility that an expert's assessment is determined by not strictly technical factors.

In this regard, the procedure is made up of the following steps:

1. **Grading.** Collecting the candidates' scores given by each of the experts and weighted according to the evaluation criteria explained in the document “2. Evaluation Criteria”

2. **Normalization or standardization.** Transforming the scores to mitigate the differences in scale and dispersion between different experts.

3. **Detection of discrepancies.** Reviewing the candidates' scores for whom there are significative differences between expert assessments. In case of significative discrepancies, the affected experts have the option of reconsidering their grades.

4. **Weighing.** Adding the (normalized) scores from all evaluators, together with the grade obtained from the preselection procedure, each of them weighted according to the expert's level of expertise with the discipline of the candidate. This level will be declared by the experts previously to the face-to-face interviews.

5. **Presentation.** Rescaling of the scores to present them, after the procedure above, following a scale from 1 to 8.

6. **Reserve list.** Creating a unique reserve list for each area.

7. **Feedback for the candidates.** Separating the candidates, for each evaluation criterium, into for quartiles according to their scores. Inform the candidates of the quartile to which belong.

2. Grading

Experts shall rate the various aspects of a candidate's application during the face-to-face assessment sessions.

The representative of “la Caixa” Foundation will be in charge of entering all experts' scores for each candidate into a database. The final marks generated in the remote assessment (preselection) process will be included. From this moment on, these marks will be considered as if they came from one additional expert in the committee, who is assumed to possess level
of expertise 1 (see step 4). The system will consider the various levels of weighting of each aspect evaluated and arrive at an initial score for each candidate from each expert

\[ \text{mark}_{c,e} = \text{mark given to candidate } c \text{ by expert } e. \]

Assuming that there are \( n \) candidates and \( m \) evaluators (including the remote assessment), then \( c \in \{1, \ldots, n\} \) and \( e \in \{1, \ldots, m\} \).

Consequently, every candidate has \( m \) marks: one from each expert, plus the one coming from the preselection phase or remote assessment. These marks take values from 1 to 8.

For later use, at this stage we calculate the quantity

\[ \text{mark}_c = \frac{1}{m-1} \sum_{e=1}^{m-1} \text{mark}_{c,e}, \]

which is the average of all experts’ marks obtained by candidate \( c \), excluding the preselection score. This number will not be used until step 5 (Presentation).

3. Standardization

The normalization or standardization of scores given by one expert in relation to all candidates he or she has evaluated will be performed according to the following procedure. For every evaluator \( e \in \{1, \ldots, m\} \):

- The expert’s mean score is calculated
  \[ \text{mean}_e = \frac{1}{n} \sum_{c=1}^{n} \text{mark}_{c,e}, \]

- The standard deviation of this same set of scores is also obtained from
  \[ \text{stdev}_e = \frac{\sqrt{\frac{\sum_{c=1}^{n} (\text{mark}_{c,e} - \text{mean}_e)^2}{n-1}}}{n-1} \]

- Finally, the set of scores is of every evaluator (also the ones coming from preselection) are normalized by
  \[ \text{mark}^\text{norm}_{c,e} = \frac{\text{mark}_{c,e} - \text{mean}_e}{\text{stdev}_e} \]

Observations

With this procedure, the original marks

\[ \text{mark}_{c,e} \in [1,8] \]

are converted in new quantities
\[ \text{mark}^{\text{norm}}_{c,e} \in (-\infty, \infty). \]

The mean of the new scores of each expert is 0 and its standard deviation is 1. In this way, the possible different tendencies of the evaluators (giving higher or lower scores in general, for example) are eliminated. The new marks will be higher or lower, depending on how far they are from the average of the original grades, and how frequent this distance is. (See the annex for further details about the effects of normalization).

4. Detection of discrepancies and an alarms

In the event that the standardized scores given by the different experts to the same candidate show a significant divergence (see below), the system displays an alarm.

To be considered as a significant divergence, the difference between the maximum and the minimum score among all those of the same candidate must be larger than 2. More precisely,

\[ \max_e (\text{mark}^{\text{norm}}_{c,e}) - \min_e (\text{mark}^{\text{norm}}_{c,e}) > 2. \]

The scores coming from the preselection process are not taken in this part of the procedure.

The system will only indicate the alarms of those candidates in the last positions with a scholarship or the first positions without fellowship. The exact number of positions to be considered will be proportional to the number of fellowships to be granted by the committee.

The candidates of this restricted group with significant divergences will be discussed by the committee at its final meeting and evaluated again by each expert, who can then maintain or change their original score. The new scores will be entered into the system again and they will be final.

5. Weighting according to expertise with the discipline assessed

The experts of each committee may indicate, via the online application that gives them access to applications, their level of expertise with the discipline of the candidacy evaluated.

Each expert can choose between two possible levels of expertise:

LEVEL 1: Their knowledge corresponds, generically, with the field of the discipline evaluated and their evaluation can therefore be considered that of an expert. The remote assessment mark is always considered that of an expert.
LEVEL 2: Their knowledge does not correspond, generically, with the field of the discipline evaluated; therefore, their evaluation cannot be considered strictly that of an expert, but it is sufficient taking into account the characteristics of the call.

The levels of expertise of the different evaluators for a given candidate will result in different weights of the expert’s score at the time of adding all the scores of this candidate. Supposing we have \( m \) evaluators (we are including here the score of the remote assessment), the weights would be distributed in the following way:

- every expert has an ensured weight of \( \frac{1}{m+1} \) and moreover
- there is an additional weight of \( \frac{1}{k(m+1)} \) to be uniformly distributed among those experts with level 1 of expertise, among which we always find the remote assessment score.

Hence, if \( k \) evaluators (including the preselection score) declare level 1 of expertise with the candidate \( c \), the expert \( e \) for the candidate \( c \) will have a weight of

\[
weight_{c,e} = \frac{1}{m+1}, \text{ if the expert } e \text{ has declared level 2, and}
\]

\[
weight_{c,e} = \frac{1}{m+1} + \frac{1}{k(m+1)}, \text{ if the expert } e \text{ has declared level 1.}
\]

Example 1:

A panel is formed by 5 experts: \( e \in \{1,2,3,4,5\} \).
The experts 3 and 5 have declared a level 1 of expertise for a certain candidate \( c \) (hence \( k = 3 \)).
As a result, the weights are distributed as follows:

<table>
<thead>
<tr>
<th>Expert ( e )</th>
<th>Level of expertise</th>
<th>( weight_{c,e} )</th>
<th>( weight_{c,e} ) (num)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>( \frac{1}{7} )</td>
<td>0,1429</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>( \frac{1}{7} )</td>
<td>0,1429</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>( \frac{1}{7} + \frac{1}{21} )</td>
<td>0,19</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>( \frac{1}{7} + \frac{1}{21} )</td>
<td>0,19</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>( \frac{1}{7} )</td>
<td>0,1429</td>
</tr>
<tr>
<td>preselection</td>
<td>1</td>
<td>( \frac{1}{7} + \frac{1}{21} )</td>
<td>0,19</td>
</tr>
<tr>
<td>TOTAL</td>
<td></td>
<td>( \frac{6}{7} + \frac{3}{21} )</td>
<td>1</td>
</tr>
</tbody>
</table>

Example 2:

A panel is formed by 5 experts: \( e \in \{1,2,3,4,5\} \).
No expert has declared a level 1 of expertise for a certain candidate \( c \) (hence \( k = 1 \)). As a result, the weights are distributed as follows:

<table>
<thead>
<tr>
<th>Expert ( e )</th>
<th>Level of expertise</th>
<th>( weight_{c,e} )</th>
<th>( weight_{c,e} ) (num)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>( \frac{1}{7} )</td>
<td>0,1429</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>( \frac{1}{7} )</td>
<td>0,1429</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>( \frac{1}{7} )</td>
<td>0,1429</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>( \frac{1}{7} )</td>
<td>0,1429</td>
</tr>
</tbody>
</table>
**Computation of the final grade**

Once the weights of each expert for every candidate have been calculated, the final score of each candidate will be computed adding for the first time the $m$ existing scores (one from each expert and one from the remote assessment), all normalized and weighted according to the expert’s level of expertise. In other words,

$$\text{final mark}_c = \sum_{e=1}^{m} \text{mark}_c^{\text{norm}} \times \text{weight}_{c,e}.$$  

This final score can take any value in $(-\infty, \infty)$, and it is the one being used to rank the candidates. This ranking will not be affected by any of the later steps, which serve only aesthetic purposes.

**6. Presentation of the scores**

With the goal of presenting the candidate's marks in a range from 1 to 8, the following procedure will be followed:

1. All final marks are first rescaled to obtain a temporary score from 0 to

$$\text{temp mark}_c = \frac{\text{final mark}_c - \min_c(\text{final mark}_c)}{\max_c(\text{final mark}_c) - \min_c(\text{final mark}_c)}.$$  

That is, to the candidate's final score, we subtract the minimum mark among all candidates and divide by the difference between the maximum and the minimum score, again among all candidates. Every mark is now between 0 and 1 but the ordering remains the same it was.

2. All scores are translated to the interval

$$\left[ \min_c(\text{mark}_c), \max_c(\text{mark}_c) \right]$$

where we recall that $\text{mark}_c$ is the direct average of the expert's scores for the candidate $c$ (see step 1), before normalizing or applying the weights according to expertise, and without taking into account the preselection results.

After this rescaling we obtain the definitive final score of each candidate:
The reserve list of each committee will be formed by the candidates who did not obtain a fellowship, in the order established in step 4 (or 5, since it is the same one).

In the event that the number of candidates to be interviewed is too high to be handled by only one committee, two or more subcommittees will be created in the same area. Candidates will be distributed between among the different subcommittees in a sequential manner according to their score. Therefore, the candidate with the highest score will be assigned to subcommittee 1, the next one to subcommittee 2, and so on until starting over, and until exhausting the candidate list.

In this situation, every subcommittee will be preassigned a number of fellowships to award and this will be done accordingly to the procedure described in the sections above.

Once the candidates to be awarded have been determined in each subcommittee, the remaining ones will be joined in a unique reserve list, common to all subcommittees (of the same committee). This list will be ordered according to the definitive score of the candidates $mark_c^{def}$, after a new normalization has been applied. More precisely, if $N$ subcommittees have been created, the second normalization and final reserve list are done as follows:

1. For every subcommittee $S$, let us say with $n$ candidates (in total), we normalize the definitive scores $mark_c^{def}$ by calculating first their average

$$mean_S = \frac{1}{n} \sum_{c=1}^{n} mark_c^{def},$$

and then their standard deviation

$$stdev_S = \sqrt{\frac{\sum_{c=1}^{n} (mark_c^{def} - mean_S)^2}{n - 1}}$$

and finally computing the normalized grade

$$mark_c^{def,norm} = \frac{mark_c^{def} - mean_S}{stdev_S}.$$  

Ordering by this new mark, does not change the previous ranking of the subcommittee. Hence fellowships will be awarded to the same candidates, but a comparison with the other subcommittee’s lists will be meaningful.

2. A unique reserve list is created by joining all candidates with no fellowship and ordering by the new normalized score of step 1, $mark_c^{def,norm}$. 

$$mark_c^{def} = temp mark_c \times (max(mark_c) - min(mark_c)) + min(mark_c)$$
In case a resignation occurs, the fellowship will be awarded to the best ranked candidate in the reserve list. In case of tie between reserve candidates, this will be resolved based on the preselection score.

8. Feedback for the candidates

In order to provide adequate feedback to the candidates on the assessment of their candidacy during the face-to-face assessment process, their qualifications for each criterion will be analyzed for comparison with the qualifications of the other candidates of their panel.

This feedback will not take into account the remote assessment mark, since interviewed candidates will have already obtained the corresponding feedback from the remote assessment process.

Candidates will not be informed about the exact marked obtained in each of the criteria, but only about the quartile to which their grade belongs, once again normalized.

Normalization (for each criteria)

A new normalization will take place although in this occasion it will be done separately for each of the criteria used for the candidate’s evaluation (for example potential, background, ...). Indeed, given a criterium \( k \in \{1,2,3\} \), and an expert \( e \in \{1,...,m\} \), the normalization will be carried out by first calculating the average and the standard deviation in the set of \( n \) candidates

\[
\text{mean}_{e,k} = \frac{1}{n} \sum_{c=1}^{n} \text{mark}_{c,e,k}
\]

\[
\text{stdev}_{e,k} = \sqrt{\frac{\sum_{c=1}^{n} (\text{mark}_{c,e,k} - \text{mean}_{e,k})^2}{n-1}}
\]

where \( \text{mark}_{c,e,k} \) denotes the score of the candidate \( c \) obtained from the expert \( e \) for the criterium \( k \).

With this two quantities, we normalize each of the score lists by calculating

\[
\text{mark}_{c,e,k}^{\text{norm}} = \frac{\text{mark}_{c,e,k} - \text{mean}_{e,k}}{\text{stdev}_{e,k}}
\]

Finally, the normalized score of each candidate \( c \) with respect to the criterium \( k \) is the average of all the normalized scores obtained by each of the evaluators, that is

\[
\text{mark}_{c,k}^{\text{norm}} = \frac{1}{m} \sum_{e=1}^{m} \text{mark}_{c,e,k}^{\text{norm}}
\]

Every candidate obtains in this way three normalized scores, one for each of the criteria considered.
Quartile division

For every criterion $k$ the list of candidates will be ordered according to the normalized scores $mark^{norm}_{c,k}$, and divided into four equal parts or quartiles $Q_1, Q_2, Q_3$ and $Q_4$, where $Q_1$ corresponds to the top group of the $n/4$ highest scores, and $Q_4$ to the bottom group with the $n/4$ lowest ones.

The candidate will be informed of the quartile assigned for each of the criteria.

In case several subcommittees had been created, the quartiles will be computed separately in each subcommittee.
9. Annex 1

About normalization of scores and its effects

The goal of this annex is to comment on the process of normalization (or standardization) which will be applied to the scores given by the experts in the face-to-face assessment phase, as well as the effects of this action.

The objective of normalizing each expert's scores before adding them to the others' or comparing between them is to ensure that the grade of every evaluator has a similar weight in the final candidate's mark, mitigating the differences in scale and dispersion that might exist between them.

The experts' scores in each of the evaluation criteria can take values in between 1 and 8, and so does the weighted average of these grades computed for every candidate and which we denote by $mark_{c,e}$ (where $c$ is the candidate and $e$ the evaluator).

To normalize the scores of the expert $e$, the average ($mean_e$) and the standard deviation ($stdev_e$) of all of his or her scores are calculated (see Section 2 of the document for more details). With these two quantities a new score for each candidate is obtained by

$$mark_{c,e}^{norm} = \frac{mark_{c,e} - mean_e}{stdev_e}.$$  

This new score takes values in $(-\infty, \infty)$, and is the one that will be used (after being weighted by the level of expertise of the evaluator for the given candidate) to compute the average grade of all the experts' scores for the given candidate.

The performed normalization has the following effects:

- The mean of the scores of each expert is equal to 0. This cancels the possible (natural) tendencies to "grade high" or "grade low" of the different experts.
- The standard deviation of the scores of each expert is equal to 1. This means that, in average, the distance (squared) to the new mean (0) is equal to 1. Approximately 95% of the new scores of each evaluator are between -2 and 2. Scores that were given in a very narrow range ($stdev \ll 1$) will now be more dispersed, while marks given in a large range ($stdev \gg 1$) will now become closer to the mean.
- Outliers will still be outliers (and sometimes even more than before). If a grade was much further from the average than the others, the new grade will have the same property. If the deviation of the list was small, this can even be enhanced.

**Example**

Suppose there are 44 candidates and, for the purpose of this example, two evaluators. The graph below shows in blue the scores of one of the experts ($e=1$) and in orange the scores of the second one ($e=2$), quite more scattered than the blue ones and with a clear outlier, with a grade of 3.
The values computed for this set of scores are:

\[
\text{mean}_1 = 7.6; \quad \text{stdev}_1 = 0.27;
\]

\[
\text{mean}_2 = 6.85; \quad \text{stdev}_2 = 0.941.
\]

The next figure shows the distribution of the new scores after normalization.

The yellow grades show a similar distribution around the mean to the one they had before (their standard deviation was close to 1), and the outlier is still there. Instead, the blue scores are now more scattered than before (even more than the yellow ones!), because their standard deviation was originally very small (they distributed tightly around their mean).