

Supporting information for the ‘Methodology’ section of the KEF options survey

For the detailed background to these questions, please refer to section 6 – Metrics and methodology of the [KEF review report](#).

Question 4 – Perspective level calculation methodology

We are proposing to change the underlying methodology for the KEF perspective calculations to address a data problem that arises as a result of the diversity of the providers represented. Due to this diversity, the data representing providers’ diverse activities may contain significant outliers and differences in orders of magnitude within an individual metric.

The current use of scaling

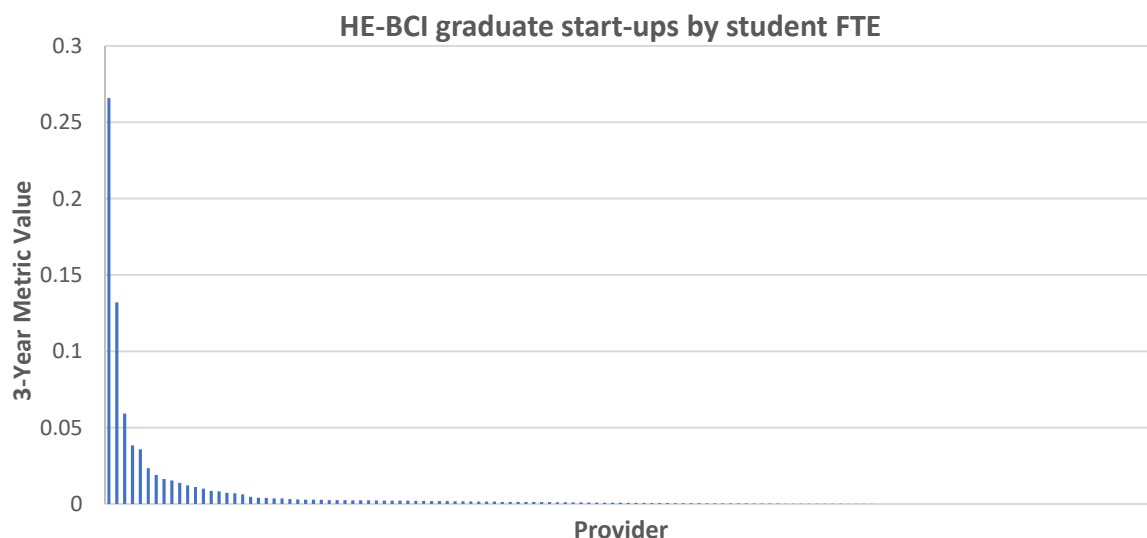
In the current KEF method, the largest 3-year value across the sector for each metric is used to normalise all other sector values to a 0-1 scale before calculating the perspective deciles. Therefore, if the largest 3-year value is a significantly large outlier then the remaining data values are scaled to significantly smaller values than if the outlier were not there. This results in metrics with large outliers effectively contributing less towards the overall perspective decile, and relatively high or low performance in this metric may not be represented in the overall perspective decile appropriately.

The use of scaling, particularly for metrics where data outliers are present, can therefore result in the unequal representation of different metrics in the overall perspective decile.

Worked example:

Although there are a number of metrics that exhibit this issue, one example is the *HE-BCI graduate start-ups by student FTE* metric where there is one provider has a significantly greater 3-year value than the rest of the sector which forms a long tail (see Figure 1 below).

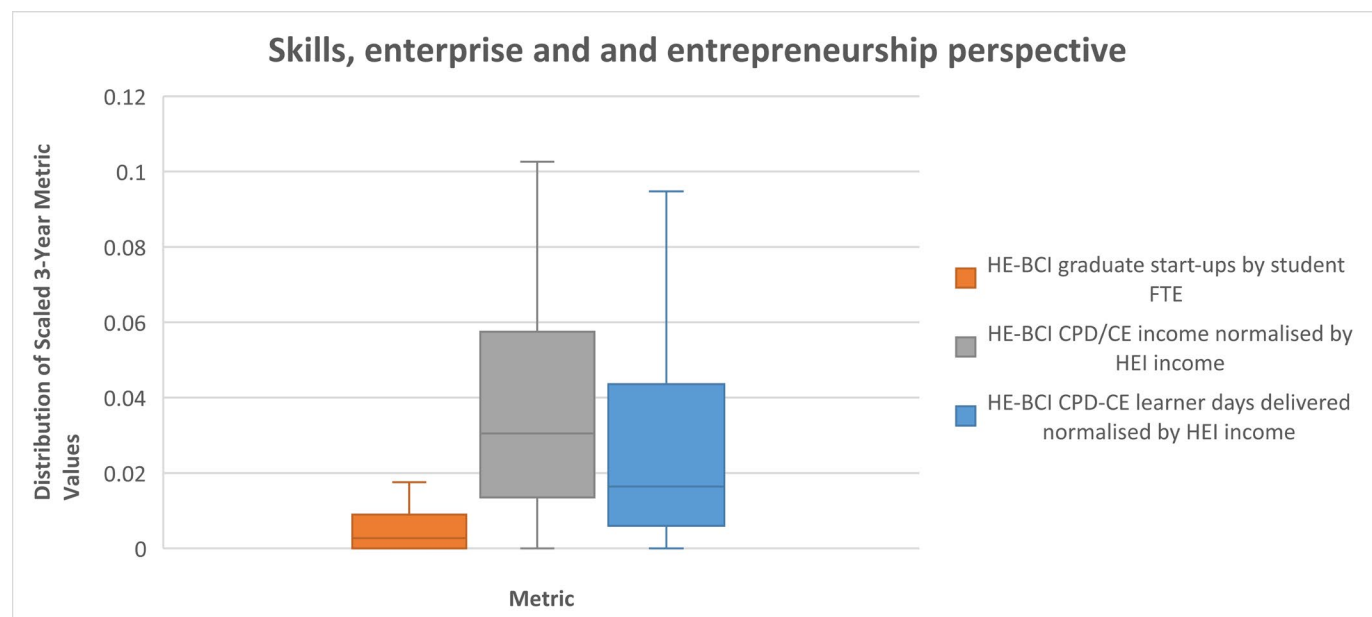
Figure 1: 3-year metric value for each provider for the HE-BCI graduate start-ups by student FTE metric in the Skills, enterprise and entrepreneurship perspective.



As a result, the scaled values of this metric are significantly smaller than that of the other metrics contributing to the overall perspective decile, potentially regardless of a provider’s performance relative to the rest of the wider sector. This results in the possibility that the relatively high or low performance of a provider in this metric is then not appropriately represented in the overall perspective decile.

The relative size, and therefore contribution, of scaled values for this metric compared to the other metrics in the *Skills, enterprise and entrepreneurship* perspective is highlighted by Figure 2 below.

Figure 2: Box and whisker of plots of the distribution of the scaled 3-year values for each of the Skills, enterprise and entrepreneurship perspective metrics.



Proposed alternative method

The alternative method proposed here seeks to improve on the current KEF method primarily through the removal of the scaling step and the use of deciles. This method does not require a scaling step to combine metric performance to calculate perspective performance, but instead totals a provider's position relative to the rest of the sector for each metric, rather than requiring a total of the absolute values for each metric.

The proposed new method involves the following steps:

- A three-year mean average value is calculated for each metric.
- All providers in the whole sector are then ordered by their three-year averages to give a metric position (1st-134th). Note that ties are unlikely in the metric values. All providers reporting a zero value would be given an equal lowest position (134th).
- For each provider the total perspective value is calculated by summing the positions of each contributing metric.
- Providers are then ordered across the sector by their total perspective values to give a perspective position (1st-134th) across the sector.
- The sector is divided into quantiles based on their perspective positions, and their performance in each perspective given a label using an involvement level corresponding to their quantile (very high to very low). We are proposing to only display this involvement level, rather than the rank or position relative to the rest of sector.
- Cluster average involvement levels are calculated by taking the mean average of the perspective positions of providers belonging to that cluster, and reporting the involvement level of the quantile in which the cluster average lies.

This method is similar to the current deciling method used, but positions are being used rather than the absolute metric values so the need for a scaling step is removed. Each metric is then equally represented in the perspective, and relative performance within a metric is also still reflected. Similar to

the current method, metric quantiles will be found by the same technique as for perspective quantiles – in the new method by dividing the sector based on metric positions.

The proposed change of methodology is also related to the number of quantiles that are used to display provider performance. Where there are multiple providers with zero metric values, and so multiple providers sharing the joint lowest position, the metric decile range will effectively be reduced (in the worked example, a metric decile range is 3-10) and all zero-reporting providers will be placed in a higher decile. As discussed in question 5 below, if fewer quantiles are used then all providers with zero metric values will be placed in the lowest quantile. A similar approach will be taken if there are multiple providers with zero values for *all* contributing metrics and therefore have the joint lowest perspective position, when placing them in perspective quantiles.

An ongoing issue with this proposed method is that the summing of metric values to give a perspective value effectively encourages consistent average performance across metrics rather than excelling in specific areas of KE based on strategic decisions or institutional capacity. However, this issue exists in the current method where the scaled metric values are mean averaged to calculate the perspective value, so the alternative method is not exacerbating this issue.

Worked example:

In the *Skills, Enterprise and Entrepreneurship* perspective the *graduate start-ups by student FTE* metric for many providers contributes relatively less to the perspective decile and therefore high performance in this metric may not be being appropriately represented in the final perspective decile.

Table 1 below details the scaled metric values for some exemplar providers and illustrates where their relative performance in a given metric does not appear to be aligned with their final perspective decile. The table also gives the new perspective deciles calculated using proposed method described above.

Table 1: Scaled 3-year metric values, metric deciles and perspective deciles for 3 exemplar provider.

Provider	Metric	3-Year Metric Value Scaled	Metric Decile	Existing method Perspective Decile	New method Perspective Decile
Provider 1 <i>Very large CPD/CE income is relative outlier</i>	HE-BCI CPD/CE income normalised by HEI income	0.3	10	10	6
	HE-BCI CPD-CE learner days delivered normalised by HEI income	0.03	6		
	HE-BCI graduate start-ups by student FTE	0	3 [#]		
Provider 2 <i>Strong performance in graduate start-ups</i>	HE-BCI CPD/CE income normalised by HEI income	0.008	1	1	3
	HE-BCI CPD-CE learner days delivered normalised by HEI income	0.003	3		
	HE-BCI graduate start-ups by student FTE	0.007	7 [#]		
Provider 3 <i>Very large graduate start-ups is relative outlier</i>	HE-BCI CPD/CE income normalised by HEI income	0.0009	1	8	4
	HE-BCI CPD-CE learner days delivered normalised by HEI income	0.002	3		
	HE-BCI graduate start-ups by student FTE	0.2	10 [#]		

[#]Due to sparse data in this metric, the full range of metric deciles has not been used and the lowest metric decile is 3.

The examples in Table 1 illustrate a number of scenarios where metrics may not be best represented in the perspective decile using the current methodology, in particular the graduate start-ups metric is consistently under-represented due to its scaled values, which has the following impacts:

- **Provider 1 - perspective performance dictated by outlying metric performance**
The perspective decile mostly reflects the exceptionally high performance in the CPD/CE income metric as this scaled value is relatively very high, and the perspective decile does not represent the performance in the other metrics.
- **Provider 2 - high performance in under-represented metric not reflected**
The perspective decile is equal to the lowest performing metric decile and the high performance in the graduate start-ups metric does not appear to be reflected (due to the low metric values after scaling).
- **Provider 3 - very high performance in under-represented metric only reflected as relative outlier**
The graduate start-ups metric is having a significant effect on the perspective decile as this provider is one of the relative outliers in this metric and so has a very high scaled metric score.

KEF Options Survey question 4* – Perspective level calculation methodology

We are proposing to change the underlying methodology for the KEF perspective calculations to remove the need for scaling by presenting perspective level results that relate to relative positions rather than absolute metric values.

Are you in agreement with RE making the proposed methodology change for KEF2?

- a) No change - continue with the same methodology used in KEF1 in 2021.
- b) Yes, make the proposed change – Use the proposed alternative methodology for KEF2.

**Survey question placed here for reference, please use the online [KEF Options Survey](#) to provide your response.*

Question 5 & 6 – Proposal to move to five quintile levels and labelling options

Background

Feedback gathered through both the KEF survey and focus groups suggested that the current 'top/bottom' nomenclature used to label deciles and final KEF results invited a competitive view of the KEF which was beyond the original objective of the exercise (presenting inter-cluster comparisons).

Further to this, the 'bottom' language can read quite negatively and may not be accurately reflecting the performance of providers in that decile. The juxtaposition of language between being in the 'top 50%' and the 'bottom 50%' also implied a greater difference in performance than may be present in reality.

Feedback from the focus groups and discussions with the metrics expert group also highlighted issues with using deciles to present final KEF results. The deciles used to present performance can result in final provider results seeming more dissimilar than in reality due to there being 10 different quantiles to divide a relatively low number of providers between. This can result in providers being presented in different final deciles even if their underlying data is very similar because their results sit very close to a threshold between different deciles. This issue occurs particularly where the underlying data is quite 'bunched', including for the public and community engagement self-assessment scores. Decreasing the number of quantiles reduces the frequency of this issue occurring.

In addition, data used in the KEF can be sparse and as a result a large number of providers report a zero value for the 3-year average in some metrics. This issue is particularly apparent in the IP and commercialisation perspective where a count of spin-outs is used rather than HEI income for the denominator and consequently a large proportion of the sector report a zero value. In order to place these providers in a decile, the current method reduces the decile range and all zero-reporting providers are all placed in a higher decile. As a result, relative performance across metrics is not necessarily consistent or accurately reflected, and the cluster average can appear high when comparing across over metrics and perspectives

The use of fewer quantiles would allow the placement of all providers with a zero metric value in the lowest quantile, without there being a potentially inaccurate gap to the quantile of the next non-zero reporting provider. In addition, relative performance across metrics will then be more consistent.

For the reasons set out above, we are proposing to move away from the use of ten deciles and instead move to the use of five quintile indicating level of involvement through labelling of 'very high to very low'. As shown in **Option 3** of the below table.

Table 2: Labelling options for the KEF dashboard.

Option 1 – no change	Option 2 – deciles with number labelling only*	Option 3 Research England proposed change – Quintiles of involvement level with word labelling	Option 4 – Quintiles with number labelling [#]	Option 5 – Quartiles with word labelling	Option 6 – Quartiles with number labelling*
Top 10%	Decile 10	Very high involvement	Quintile 5	Higher involvement	Quartile 4
Top 20%	Decile 9				
Top 30%	Decile 8	Higher Involvement	Quintile 4	Higher medium involvement	Quartile 3
Top 40%	Decile 7				
Top 50%	Decile 6	Medium involvement	Quintile 3	Lower medium involvement	Quartile 2
Bottom 50%	Decile 5				
Bottom 60%	Decile 4	Lower involvement	Quintile 2	Lower involvement	Quartile 1
Bottom 70%	Decile 3				
Bottom 80%	Decile 2	Very low involvement	Quintile 1	Lower involvement	Quartile 1
Bottom 90%	Decile 1				

[#] We would provide a label to indicate that the higher the quantile number the higher the performance and quantile '1' represents the lowest level of performance.

Question 5* – Proposal to move to five quintile levels

We are proposing to move from presenting results as ten 'decile' levels to five 'quintile' levels of involvement. There is also an alternative option to move to four 'quartile' levels of involvement.

Please rank from the following options where 1 = most preferred and 3 = least preferred

- a) **No change** – retain ten deciles (options 1 or 2)
- b) **Proposed change** – use five 'quintile' levels (options 3 or 4)
- c) **Alternative option for change** – use four 'quartile' levels (options 5 or 6)

**Survey question placed here for reference, please use the online [KEF Options Survey](#) to provide your response.*

Question 6* – Perspective level labelling options

We are also proposing to change the nomenclature used to label results from 'top/bottom X%' to 'involvement levels'. These involvements levels could either be labelled with qualitative descriptions to indicate a provider's higher or lower level of involvement, or with numerical labels where '1' represents the lowest level of involvement. We are proposing to use qualitative descriptions as set out in table 2, option 3 above.

Please state your preference for labelling perspective level outcomes.

- a) **Words** – use words to label the levels of involvement (Research England's proposed outcome)
- b) **Numbers** – use numbers only to indicate whether an involvement level is high or low

**Survey question placed here for reference, please use the online [KEF Options Survey](#) to provide your response.*