

Evaluation in Regulated Procurements: To what extent are contracting authorities obliged to moderate scores as part of evaluation of bids, and how should contracting authorities approach moderation?

consistently held (as set out in this essay) that moderation of scores is closely linked to the maintenance of adequate documentation to justify evaluation decisions. In cases where contracting authorities have moderated scores without documenting reasons for adjustments, the courts have held that it is arguable that those adjustments were made unlawfully.⁵ In *Bristol Missing Link Limited v Bristol City Council*, Coulson J held that “in cases where there are clear issues arising out of individual scores, it will be difficult for the court to conclude that there is no serious issue to be tried”.⁶ Whilst it does not necessarily follow from this that a contracting authority which fails entirely to moderate scores has done so unlawfully, it is at least arguable that in a complex procurement, an evaluation process which does not consider scores beyond an initial evaluation exercise will not be able to show adequate justification of those scores, and therefore will not be compliant with regulation 84 PCR 2015. This gives further weight to the view that contracting authorities are under a legal obligation to moderate scores in an evaluation process.

How should contracting authorities approach moderation?

The extent to which scores should be moderated, and the method of that moderation, is open to the discretion of contracting authorities to choose an approach. Recent decisions of the UK courts in considering various methods of moderation provide useful guidance and should inform the planning stages of procurement exercises.

i. Applying a calculation to initial scores

A simple moderation method is to apply a cumulative calculation to the scores of all evaluators, such as the average, mean or mode, to achieve an overall score. Whilst this may be an appropriate and time-efficient method of evaluating the simplest procurement exercises, this method has attracted criticism from the courts. McCloskey J stated in *Resource NI v Northern Ireland Courts and Tribunals Services* that moderation exercises should be:

*...something considerably greater than merely formal events. They are solemn exercises of critical importance to economic operators and the public and must be designed, constructed and transacted in such a manner to ensure that full effect is given to the overarching procurement rules and principles.*⁷

Such calculations would not identify typographical mistakes or errors arising from misinterpretation of criteria or bids. As discussed in *Bristol Missing Link*⁸, an erroneous score may unfairly impact upon the overall contract award if such a calculation were applied without closer consideration. In that case, the panel of five evaluators moderated

⁵ *Bristol Missing Link Limited v Bristol City Council* [2015] EWHC 876 (TCC)

⁶ *Bristol Missing Link Limited v Bristol City Council* [2015] EWHC 876 (TCC) [34]

⁷ *Resource NI v Northern Ireland Courts and Tribunal Service* [2011] NIQB 121 [35]

⁸ *Bristol Missing Link Limited v Bristol City Council* [2015] EWHC 876 (TCC)

Evaluation in Regulated Procurements: To what extent are contracting authorities obliged to moderate scores as part of evaluation of bids, and how should contracting authorities approach moderation?

scores together and made adjustments on a consensus basis. Some adjustments appeared unusual, such as individual scores of 4,4,4,4,4 and 3 which were moderated to an overall score of 3. As the Council had not kept adequate records of its reasons, it struggled to justify these unusual adjustments. Significantly, the Council admitted in a letter to the Claimant that:

...there were a number of instances whereby the panel members had not marked the bid in accordance with the advertised criteria. Had the moderation process not have taken place, and a mere average or median of the individual marks awarded, we would have been scoring according to incorrect criteria and would have left ourselves open to a legitimate challenge from an aggrieved bidder.⁹

This admission indicated that the original evaluation process may have been flawed, and thus Coulson J held that this highlighted the importance of conducting a moderation exercise (rather than simply performing a calculation). Moderation by way of group discussion thereby increases the chance of evaluators rectifying erroneous scores, and should therefore be preferred above any simple calculation.

ii. Group moderation

Group discussion should not, however, mean conducting the initial evaluation and moderation of bids simultaneously by way of group marking. In *Counted4 Community Interest Company v Sunderland City Council*¹⁰, five panel members initially read and evaluated all bids together as a group. They agreed and recorded scores, noting contemporaneously the features and shortcomings of each response. Unfortunately for the Council, one of the five panel members had an alleged conflict of interest. Due to the evaluation/moderation having been performed as a group, the court held that the conflicted evaluator may have unduly influenced the decisions of all other evaluators, and therefore held that it was properly arguable that the Council had failed to effectively prevent, identify and remedy conflicts of interest in the evaluation process. This highlights the importance of evaluators initially considering bids separately so as to preserve their independence within the initial review.

*European Dynamics SA v HM Treasury*¹¹ used a similar consensus scoring methodology, and despite failing to maintain sufficient records of the evaluation process, HM Treasury's two-stage moderation approach contained some exemplary features. The scores of two initial evaluators were checked by a consensus scorer. If any scores appeared too high or low, the court heard that the consensus scorer (in consultation with the initial evaluators) adjusted scores accordingly. The consensus scorer identified areas of variation and focussed the group's discussions on those areas, thus increasing the time-efficiency of the

⁹ *Bristol Missing Link Limited v Bristol City Council* [2015] EWHC 876 (TCC)

¹⁰ *Counted4 Community Interest Group v Sunderland City Council* [2015] EWHC 3898 (TCC)

¹¹ *European Dynamics SA v HM Treasury* [2009] EWHC 3419 (TCC)

Evaluation in Regulated Procurements: To what extent are contracting authorities obliged to moderate scores as part of evaluation of bids, and how should contracting authorities approach moderation?

moderation process. The initial evaluators scored bids independently whilst anomalies were discussed and resolved as a group. Unfortunately, the lack of adequate records meant that the contracting authority was unable to demonstrate its reasons for making the adjustments. The court therefore was unable to determine on the evidence whether the adjustments were justified or the result of flaws in the process. Had it recorded its reasons fully, and if those reasons had been justifiable, the court may have found that there was no serious issue to be tried on this basis. Under regulation 67(1) PCR 2015, contracting authorities have the discretion to assess the MEAT from their own point of view. This is a significant point within the PCR 2015, yet this is intrinsically linked to compliance with regulation 84 PCR 2015. If the courts have no documentation containing evidence of decisions, an unusual decision that is compliant with regulation 67(1) may be found to be ultimately unlawful for lack of such evidence. *European Dynamics* therefore emphasises the importance of contracting authorities complying with regulation 84 PCR 2015.

iii. Super-moderators

Contracting authorities conducting large procurement exercises may choose to moderate scores without further consultation with original evaluators. This was the chosen methodology for at least a portion of the Legal Aid Agency's (LAA) procurement of legal aid contracts for duty solicitors. Award letters and early disclosure given in the judicial review and Part 7 claims made against the LAA by unsuccessful bidders¹² revealed that the LAA developed a "supra-moderation" or "consistency review" technique, whereby so-called "super-moderators" reviewed generic answers provided by firms bidding across multiple procurement areas.¹³ Little information is publicly available on the role of the super-moderators, as the litigation concluded before such detailed information came to light. Information made available to the media indicates that they assumed responsibility for the overall score of any generic answers. This could have included adjusting scores at their own discretion and without further discussion with initial evaluators. The LAA maintained that "*the entire process was subject to careful moderation and management.*"¹⁴

Although it is difficult to evaluate this approach on such little information, one assumes that this approach was chosen in order to increase time-efficiency and consistency across this large evaluation exercise. The concern for contracting authorities in adopting a similar

¹² Case management directions were set out in the Court Order in claim HT-2015-000373 ((1) Kaim Todner Solicitors & Ord. (2) Other Claimants identified in the Schedule to this Order v The Lord Chancellor) in the Technology & Construction Court before Mr Justice Stuart-Smith dated 18 November 2015

¹³ "Super-moderators" were discussed at the interim hearing on 21 December 2015 in which the LAA were ordered to disclose "*assessment and moderation documents underlying the summary of reasons document given with the award notices to all Claimants*" including (Schedule A) "*all guidance and training materials provided to assessors, moderators and/ super-moderators*" and "*all notes of meetings and memoranda in relation to the scoring process for bids, including the core documents relating to the development of the so-called 'supra-moderation' or 'consistency review' process.*" Court Order in claim HT-2015-000373 ((1) Kaim Todner Solicitors & Ord. (2) Other Claimants identified in the Schedule to this Order v The Lord Chancellor) in the Administrative Court before Honourable Sir Kenneth Parker dated 21 December 2015

¹⁴ Monidipa Fouzder, "LAA committed to 11 January start for new legal aid contracts" *Law Society Gazette* (London, 4 November 2015)

Evaluation in Regulated Procurements: To what extent are contracting authorities obliged to moderate scores as part of evaluation of bids, and how should contracting authorities approach moderation?

approach is in affording the super-moderator such a large amount of discretion. If concerns were to arise as to a potential conflict of the super-moderator, this could prove extremely damaging to the validity of the evaluation process. Contracting authorities using this approach should firstly ensure that any such conflicts on the part of the super-moderator are addressed at the outset of the evaluation. Equally, whilst it is important to train all evaluators to score bids in accordance with the PCR 2015, it is essential for the super-moderator to have an excellent understanding of their obligations. All decisions of the super-moderator should be meticulously recorded in compliance with regulation 84 PCR 2015. Finally, it would be prudent to consider methods of moderating the scores of the super-moderator. This could be approached by way of approval of the super-moderator's recommendations by committee, or if multiple super-moderators are used, by holding a moderation meeting between the super-moderators. However this is approached, contracting authorities should be wary of a moderation method which allows one person's decisions to go unchecked.

iv. Multi-stage approach

The recent investigation conducted by Monitor into the handling of a procurement exercise by NHS Northern, Eastern and Western Devon Clinical Commissioning Group (NEW Devon CCG) provides an excellent example of an effective evaluation and moderation process under a different procurement regime (the NHS (Procurement, Patient Choice and Competition) Regulations 2013¹⁵ (the NHS Regulations)). Whilst the NHS Regulations operate within a different framework to the PCR 2015, there are similarities between the general requirements of the NHS Regulations (which refer to "best value for money"¹⁶) and the MEAT in the PCR 2015 ("most economically advantageous"). The NHS Regulations, like the PCR 2015, also require compliance with principles derived from the TFEU¹⁷.

New Devon CCG's evaluation had four stages. Firstly, each bid was evaluated separately by individual evaluators thus preserving the independence of the initial evaluations. Secondly, the initial scores and comments were reviewed by a separate panel of moderators. The panel was led by a non-scoring chair whose role was to conduct the moderation and assist the panel in reaching a consensus. In utilising a non-scoring chair, this reduced concerns of undue control of the process by a single moderator. Thirdly, the moderation panel's recommended scores were presented to an executive committee for approval. Finally, the recommended scores were presented to the governing body for approval. This dual-level approval system significantly weakened any challenge on grounds of lack of justification of scores. The process was designed so that a different group of people made up of representatives from various interested bodies were involved

¹⁵ National Health Service (Procurement, Patient Choice and Competition) Regulations 2013, SI 2013/257

¹⁶ National Health Service (Procurement, Patient Choice and Competition) Regulations 2013, SI 2013/257 1 3(3)(b)

¹⁷ National Health Service (Procurement, Patient Choice and Competition) Regulations 2013, SI 2013/257 1 (3)(2)

Evaluation in Regulated Procurements: To what extent are contracting authorities obliged to moderate scores as part of evaluation of bids, and how should contracting authorities approach moderation?

at each stage, thus ensuring "*independence and challenge*"¹⁸ throughout. Monitor also praised the training given to all evaluators in the importance of evaluating bids in a non-discriminatory, proportionate and transparent manner.¹⁹

Monitor's investigation into New Devon CCG's evaluation process highlights the merits of a multi-stage moderation approach. Whilst not a judicial decision (and whilst unlikely to influence the decisions of the court), Monitor's decision is useful in considering how other bodies have approached the issue of moderation and how this approach has stood up to close scrutiny much like the UK courts would perform.

When is moderation not enough?

In *Bristol Missing Link*, Counsel for both parties debated whether all that mattered in respect of moderation was the final scores produced as a result of a moderation exercise, or whether the manner of the exercise itself should come under scrutiny.²⁰ Coulson J did not reach a conclusion on this point as he considered that this was not at the heart of the case before him. His conclusion in *Woods Building Services v Milton Keynes Council*²¹, however, indicates that the mere formality of conducting a moderation exercise does not necessarily satisfy the courts that an evaluation is compliant. In that case, the Council's initial evaluation panel of two, one of whom was found to have a potential conflict of interest, awarded the lowest-priced tender 30 marks less than the highest-scoring tender. Significantly, the Council admitted that it suspected that something had gone wrong in the initial evaluation exercise.²² The Head of Major Works moderated the initial scores and, upon review, adjusted the scores to a difference of 16 marks. As this did not change the outcome of the evaluation, a third and final moderation exercise was conducted in which all three evaluators reviewed the scores again. Coulson J stated that, having suspected that something was wrong with the initial scores, the Council should not have moderated those scores.²³ The erroneous scores may have prejudiced the next two stages of moderation by unduly influencing the decisions of the evaluators. Coulson J stated that the Council should have disregarded the initial scores and re-evaluated the tenders from scratch. By continuing to use the initial scores, the court held that, whilst not amounting to a material breach of the PCR 2015, the Council's approach was insufficiently clear.²⁴

Considering Coulson J's decision in *Woods*, it is arguable that his view in *Bristol Missing Link* (whilst not articulated) was that it is the manner of the moderation and the final scores

¹⁸ Monitor, "Investigation into NEW Devon CCG's commissioning of community services for adults with complex care needs in eastern Devon: final report" (August 2015) CCD01/15 <<https://www.gov.uk/government/publications/case-investigation-into-the-provision-of-complex-adult-community-services-in-east-devon#history>> accessed 30 March 2016

¹⁹ Monitor, Investigation into NEW Devon CCG's commissioning of community services for adults with complex care needs in eastern Devon: final report" (August 2015) CCD01/15 [57]

²⁰ *Bristol Missing Link Limited v Bristol City Council* [2015] EWHC 876 (TCC) [40]

²¹ *Woods Building Services v Milton Keynes Council* [2015] EWHC 2011 (TCC)

²² *Woods Building Services v Milton Keynes Council* [2015] EWHC 2011 (TCC) [26]

²³ *Woods Building Services v Milton Keynes Council* [2015] EWHC 2011 (TCC) [41]

²⁴ *Woods Building Services v Milton Keynes Council* [2015] EWHC 2011 (TCC) [44]

Evaluation in Regulated Procurements: To what extent are contracting authorities obliged to moderate scores as part of evaluation of bids, and how should contracting authorities approach moderation?

produced as a result of that process that matter in an evaluation exercise. *Woods* highlights the need to evaluate from scratch in the event that moderation identifies a compromised evaluation process. This may be a difficult line to identify, as one could question at what point errors should trigger the need for an entirely fresh evaluation. The fact that the unusual difference in marks created an overall suspicion that an error may have occurred in *Woods*, but that the errors could not be traced or rectified, should have alerted the Council to the need to start the process again. A cautious approach would be to say that if errors cannot be easily traced and yet there is a pervading sense of suspicion, a fresh evaluation should be used to avoid the initial evaluation prejudicing the moderation process.

Conclusion

Whilst the PCR 2015 does not explicitly require a contracting authority to conduct a moderation exercise as part of its evaluation process, there is a legal obligation to conduct evaluations in which all bidders are treated equally, and which identify the MEAT in a transparent, proportionate, and non-discriminatory way. In considering whether a contracting authority has complied with these principles, case law set out in this essay shows that the courts closely scrutinise evaluation exercises. Indeed, McCloskey J stated in *Resource NI* that "*the activities and deliberations of an evaluation panel, [and] the evidence bearing thereon, will, inevitably, be carefully and objectively scrutinised by the court.*"²⁵ It is difficult for a contracting authority to demonstrate that the PCR 2015 has been complied with if its evaluation process does not include, firstly, a robust moderation process, and secondly, documentation of decisions made during that process. Contracting authorities should therefore factor these considerations into procurement exercises. The most effective moderation methods, as shown in the approach of New Devon CCG and *European Dynamics*, include an evaluation by an independent panel followed by moderation, ideally by a new panel of moderators or in coordination with a non-scoring chair. Whilst contracting authorities may well be reluctant to include several stages of approval at senior levels of their organisations on grounds of time-efficiency, this adds further weight to moderation exercises. Crucially, the extent to which a contracting authority complies with regulation 84 PCR 2015 in relation to recording decisions throughout the evaluation will be the most important factor in the event of a legal challenge from an unsuccessful bidder, and the evidence contained therein will inform the court's decision in such a challenge. It is therefore imperative that contracting authorities make every effort to record detailed reasons for decisions made at every stage of an evaluation process.

²⁵

Resource NI v Northern Ireland Courts and Tribunals Service [2011] NIQB 121 [35]