

White Paper conference Evaluation

How far can evaluators push their “margin of discretion”?
Is consensus scoring essential? What happens if there is an impasse?

Kate Gough, Counsel, Freshfields Bruckhaus Deringer LLP

09 July 2021



Freshfields Bruckhaus Deringer

Evaluation: Why do we care?

**Key risk for
contracting
authorities**

**Basis of a
large number of
bid challenges**

**Many
practical
pit falls**

Agenda

1 Reminder of key principles

2 The thorny questions

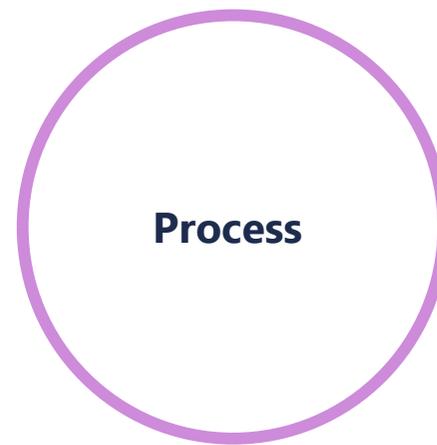
3 Where does that leave us? Some tips

1

Reminder of the key principles

Evaluation

Four key elements: risks associated with each of them



Evaluation criteria and scoring methodology

- Not the focus of this presentation BUT criteria and scoring methodology are the foundation of all evaluations (or should be)
- Contracts must be awarded on the basis of criteria set in accordance with Regulations 67 to 69
- Must disclose in advance and in full all award criteria and sub-criteria together with weightings and sub-weightings
- Principles of equal treatment and transparency apply to the evaluation criteria and the conduct of the evaluation itself (Regulation 18)

“Experience tells us that evaluating quality effectively requires time and expertise. Market testing can help to overcome any pre-existing bias and ensure that criteria are fit for purpose” (Bid Evaluation Guidance Note, May 2021)

2

The thorny questions

The questions

1

How far can evaluators push their margin of discretion?

2

Is consensus scoring essential?

3

What happens if there is an impasse?

How far can evaluators push their margin of discretion?

Do evaluators have a “margin of discretion”?

- No discretion or “margin of appreciation” when assessing a bid in terms of compliance with the obligations of equal treatment and transparency (**Firebuy**)
- BUT in relation to matters of judgment or assessment, the Court will intervene only where there has been a “manifest error” (**Varney**)
- where an evaluative judgement has been involved in arriving at the score awarded (as here), this “margin of appreciation” is brought to bear in considering whether there has been a manifest error (**EnergySolutions**)
- “no function of the court to re-evaluate the different bids absent manifest error or other breach.” (**Bechtel**)
- “the bid that is evaluated with the highest score will, assuming that it is evaluated in accordance with the ITT (and absent manifest error), win that competition (absent breaches of the principles of fairness, equal treatment and transparency).” (**Bechtel**)
- Having scoring systems which favour one bidder compared to another does not make it manifestly wrong. There must be something else wrong before the court will interfere (**Firebuy**)

What is a manifest error?

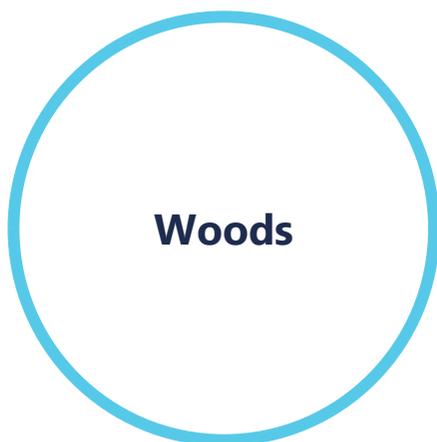
- “the word ‘manifest’ does not require any exaggerated description of obviousness. A case of ‘manifest error’ is a case where an error has clearly been made.” (**Moran J Firebuy**)
- “essentially about the nature and centrality (or materiality) of the error in question” (**Coulson J Woods referring to the observations of Green J in *Gibraltar Gaming***)

Broad equivalence with Wednesbury unreasonableness:

- “a decision so unreasonable that no reasonable person acting reasonably could have made it”
- “put another way it was not irrational” (**Coulson J Woods**)

Manifest error v margin of appreciation

- A high threshold
- BUT can evaluators be comfortable that a manifest error will never be found? **NO**
- Helpful to look at examples of where the Court has found manifest error to test how far you can push it...



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

- MEAT: cost (at 60%) and quality (at 40%)
- Quality: 12 questions scored 0, 2, 4, 6, 8 or 10
 - 10 = Response meets requirements to a very high standard with clear and credible added value and/or innovation.
 - 0 = Response does not meet requirements and/or is unacceptable. Insufficient information to demonstrate Tenderer's ability to deliver the services.
- Woods' offer was the cheapest. EAS scored significantly higher on quality had highest overall score
- Woods challenged the decision inc alleging the Council's evaluation was fundamentally flawed

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

(Cont'd)

Example of manifest error [para 51] (see also [para 63]): *you must apply the stated scoring criteria*

- Requirement: method statement setting out proposals to meet the service requirements. The only part of the tender response that was to form part of the contract and had a significant weighting
 - EAS scored 10. Argued should be 0. Alleged it failed to deal with the reinstatement works at all. There was undisputed evidence that the reinstatement works would be worth approximately 40% of the total contract value
- “scoring criteria are a matter of law and that, if a response did not meet the Council’s requirements and/or was unacceptable in a significant or substantive way, then it required to be given a zero score. That would make any other score a manifest error.”

In evidence the evaluator had accepted:

- EAS did not deal with reinstatement works at all, despite the fact that this was a major part of the project; their failure so to do was “unacceptable”; there were a range of other omissions from the EAS tender, including a lack of any explanation for how the asbestos removal works themselves might have been performed

EnergySolutions v NDA [2016] EWHC 1988 (TCC)

- Procurement of a very sizeable contract for the decommissioning of 12 different Magnox nuclear facilities in the United Kingdom
- RSS lost by 1.06%
- ES alleged inter alia NDA had made a number of manifest errors of evaluation

Example manifest error: *be careful to apply the correct test*

- Evaluators found that RSS had failed to identify key critical assets or critical assets sufficiently or adequately in the Tender Response, and that as a result the submission did not qualify for a score of 5
- There was detailed evidence on the test that had been applied by the NDA to evaluate what was a critical asset
- Judge found there had been a manifest error
- failed to apply the correct industry-wide test when they were considering the identification of critical assets. They applied the wrong test when they were evaluating the RSS Tender Submission in this respect



- The NDA’s evaluation of all of these Requirements was therefore, in my judgment, manifestly erroneous
- The test applied by Mr Grey was a different one to the “risk based approach” he himself had included in the Gate C document that he had drafted

***EnergySolutions v NDA* [2016] EWHC 1988 (TCC)**

Remember be careful of pass/fail criteria [para 871]

- NDA made it clear in their Statement of Response Requirements that certain failures would lead to disqualification
- Fraser J found that where that is the case “there is no discretion upon the NDA [the tenderer] to disregard such a failure”
- He referred to a number of EU cases including the statement of M. Jaeger in *T-40/01 Scan Office Design SA v European Commission* [2002] ECR II-5046 at para 76 that “even if the contracting authority has a certain margin of discretion in the context of a negotiated procedure, it is always bound to ensure observance of the terms and conditions of the tender specifications, which they have freely chosen to make mandatory”.
- Fraser J noted:
 - “The NDA’s bid evaluation team were not permitted to “lean against disqualification” and to apply the scoring criteria differently to how they were set out in the tender documentation, simply because they feared that to apply them correctly would result in the drastic consequences of a bid being disqualified”

How far can evaluators push their margin of discretion?

Careful!! Breach of principles of fairness, equal treatment and transparency

Transparency

- RWIND test (**SIAC Construction**)
 - Understand their significance and interpret them in the same way
 - The contracting authority can ascertain whether tenders satisfy the relevant criteria
- Undisclosed criteria
- Record keeping
 - Failure to retain notes/documentation is a breach of transparency (**Energy Solutions**)
 - “a procurement in which the contracting authority cannot explain why it awarded the scores which it did fails the most basic standard of transparency.” (**Lancashire**)

Equal treatment

- “contracting authority must treat both parties in the same way.” (**Woods**)
- Award criteria and weighting cannot be changed once bids received (**C226/09**)
- A failure to comply with award criteria is a breach of duty (**Easycoach Ltd**)
- Careful of clarifications
 - Any clarification should not appear to have unduly favoured or disadvantaged the tenderer to whom it is addressed (**Oriel Training Services**)

Is consensus scoring essential?

What happens if there is an impasse?

- No prescribed method for moderation
- Different practices include moderation
 - By consensus (either via sequential discussion with individual evaluators or simultaneous discussion collectively in one room)
 - By review alone (rather than through discussion)
- Variation in practice also occurs in relation to how the authority arrives at the final score
 - Consensus
 - Averaging scores
 - Removing outliers and arriving at a median score
 - Casting vote (to moderator or designated SME)
- Each is fraught with difficulty but move toward consensus scoring
 - The detail of the process must be set out in advance otherwise risk serious allegation of unequal treatment
 - Process should not allow for an impasse

Is consensus scoring essential?

What happens if there is an impasse

Bid evaluation guidance note May 2021 (some 34 pages)

The guiding principle is that:

“Whether for a complex procurement, or a more straightforward contract, a well-planned and implemented bid evaluation process is critical to deliver value for money”

- Bids should be independently scored by at least two evaluators, who have sufficient knowledge and understanding of the criteria they are evaluating (para 3.3.2)
- Best practice to have the same evaluator evaluate all responses to a particular question (para 3.3.2) (consistency)
- Independent moderator, skilled in moderation, to support evaluators in reaching a consensus score. Evaluators should not act as moderators
- Have a separate note taker in addition to the moderator (para 3.3.4)
- Critical that the scores awarded are supported by records which provide justification (9.1.2)
- Once individual scoring is complete, evaluators should send their scores and evaluation reports to the moderator for review (9.3.1)
 - Once the completed evaluation reports have been reviewed a meeting should be held between all evaluators and the moderator
 - The purpose of the moderation meeting is to **agree a single consensus score (should not be an average)**
- **NB** There is no discussion on what should happen should the parties be unable to reach this ‘single consensus score’ – the evaluation does not complete until that occurs

Is consensus scoring essential?

What happens if there is an impasse

The judgment in **Lancashire Care NHS Foundation Trust v Lancashire County Council [2018] EWHC 1589 (TCC)** provides guidance for moderating scores:

- There should be a clear record of reasoning that leads each evaluator to approving the final score, including the decisive factors. In the record, there should be an attempt to attribute conflicting points to individual evaluators, *“to reconcile them or to indicate whether or to what extent the panel reached agreement, even if that agreement was an agreement to disagree on the significance of a point while agreeing on the ultimate consensus score.”* (para 33)
- Moderation notes must contain a *“full, transparent and fair summary of the discussion that led to the consensus score”* (para 58). The notes must contain the ultimate reason for awarding the consensus score to each bidder.
- Individual evaluators score sheets are not adequate evidence, particularly when an evaluator’s view shifts based on the discussion in the moderation meeting. The reasons for the award of a score must reflect the discussions at a moderation meeting.

Is consensus scoring essential? What happens if there is an impasse

Some limited guidance in the case law –
see **Wordperfect Translation v Minister for Public expenditure [2018] IECA (1)**

- In the High Court Barrett J. rejected the argument that the evaluators were required to explain the progression of marks as between the various evaluation meetings. I think that he was perfectly correct in so holding. As he put it, the test of manifest error is a ground of objection targeted ultimately at an impugned decision, not at the notes of the meetings which precede the making of such decision.
- Just as importantly, evaluators should have the freedom to explore, consider and reflect on the strengths and weaknesses of the various tenders.
- The evaluators must be prepared to stand or fall by a review of the final published evaluation for manifest error. But short of that they cannot be expected to have to defend what are, at best, tentative or provisional views expressed during the course of evaluation process.

3

Where does that leave us? Some tips

What can you do? Some tips

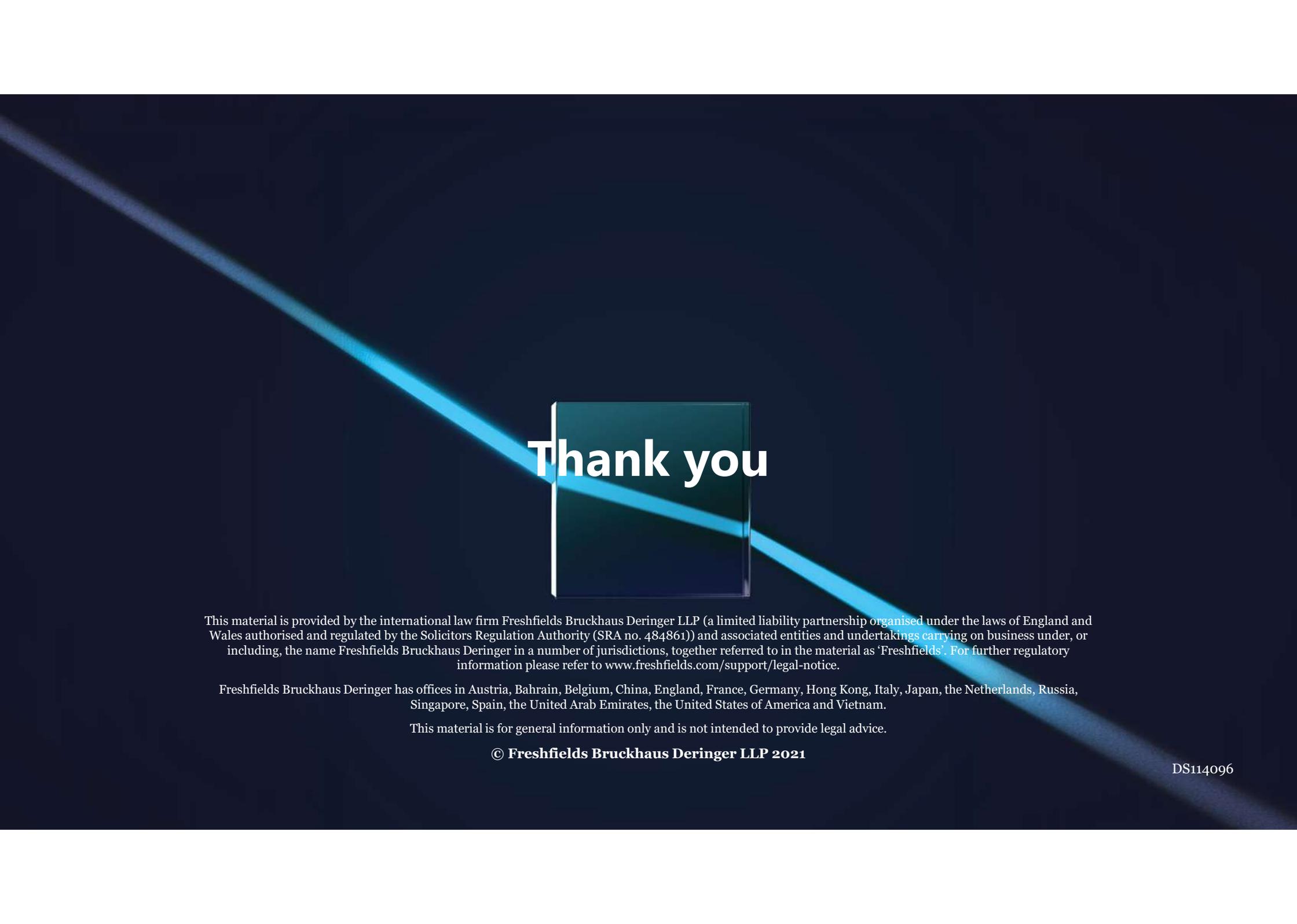
Focus on getting the evaluation criteria and scoring right at the outset

- Always have an eye on the rules/strict parameters
- What does the CA actually want/what does good look like
- Are the criteria clear and easily capable of consistent interpretation – think about the RWIND tenderer
- Be careful of pass/fail criteria
- Model different scenarios
- Be aware of not unreasonably favouring one bidder inc the incumbent
- Consider risk of scores bunching (close score of themselves not necessarily an issue but creates risk)
- Does the scoring/weighting match what is actually important

What can you do? Some tips

The evaluation

- Select the right evaluators: do they have subject matter expertise, do they understand the procurement process, criteria, scoring and rules
- Allow adequate time and identify and agree appropriate resources
- Training
- Design a clear, simple process: over complications helps no one and leads to errors
- Must be clear on the process at the outset and stick to it!
 - Set out what procedure is to be followed including the scoring and any moderation
 - Address how, in practice, the bids will be marked and by whom
- Record keeping



Thank you

This material is provided by the international law firm Freshfields Bruckhaus Deringer LLP (a limited liability partnership organised under the laws of England and Wales authorised and regulated by the Solicitors Regulation Authority (SRA no. 484861)) and associated entities and undertakings carrying on business under, or including, the name Freshfields Bruckhaus Deringer in a number of jurisdictions, together referred to in the material as 'Freshfields'. For further regulatory information please refer to www.freshfields.com/support/legal-notice.

Freshfields Bruckhaus Deringer has offices in Austria, Bahrain, Belgium, China, England, France, Germany, Hong Kong, Italy, Japan, the Netherlands, Russia, Singapore, Spain, the United Arab Emirates, the United States of America and Vietnam.

This material is for general information only and is not intended to provide legal advice.

© Freshfields Bruckhaus Deringer LLP 2021

DS114096