Dual-Anonymous Peer Review at ESO: a dry run for P106

This document is PUBLIC

Scientific Technical Committee is invited to note this document.
Dual-Anonymous Peer Review at ESO: a dry run for P106

F. Patat – Observing Programmes Office

Introduction

Like the vast majority of scientific facilities and journals, ESO has so far used the single-anonymous peer-review (SAPR) in its time allocation process. In this paradigm, the principal investigator (and more in general the proposing team) does not know who the reviewer/s will be, while the reviewers are provided with extended information on the team. This usually consists of the applicants’ names, affiliations, publications relevant to the subject, previous usages of the given facility and so on. Implicitly and/or explicitly this implies that this information is used to judge the proposal quality and the ability of the team to carry out the proposed project. If on the one hand this may be useful to provide the reviewers with more background for evaluating the science case, on the other it may constitute a source of systematics, averting the focus of the review from the science case to the quality of the proposing team.

As a first countermeasure and following the practice recently deployed at other large ground-based and space-born facilities, as of Period 102 ESO has already:

- removed the principal investigator’s (PI) name from the front page of the proposal distributed to the referees;
- removed all the affiliations and e-mail addresses;
- listed all co-investigators (Cols) in alphabetical order on the last page of the proposal.

Despite of these actions, the reviewers are still able to identify the team, and are provided with the information regarding the publication record, previous work in the given field and previous usage of the facility. Once this information is there, it is to be expected that it is used, consciously or unconsciously, possibly diverting the discussion and the evaluation from the science case and introducing systematic effects which are not related to the effective scientific merit of the application. Gender bias is one of them (see Patat 2016 for a statistical study on time allocation at ESO).

This issue is addressed by the dual-anonymous peer review (DAPR). In this paradigm neither the applicants nor the reviewers know each other’s identity. DAPR has been already deployed by some organizations and is considered as the most effective form of peer-review (see for instance Mulligan, Hall & Raphael 2013). DAPR has been adopted by the Space Telescope Science Institute for allocating HST time as of Cycle 26 (Strolger & Reid 2019; Strolger & Natarajan 2019), and will be adopted by ALMA as of Cycle 8. A recent study by Johnson & Kirk (2020) has shown that anonymization can nearly eliminate gender systematics from application processes. In particular, their analysis shows that when the identities of the team were known, about 60% of the discussion was about the team itself, not the science being proposed. And that gives an obvious opportunity for unconscious bias to creep back into the process. In addition, as shown by the outcome of HST Cycle 27, DAPR has also the potential of levelling the playing field between new and established researchers (Reid 2019, slides #9 and #10).
In the DAPR approach, the applicants still provide all the information about the proposing team, but this is not displayed in the proposal version that is distributed to the reviewers. The information is stored by the organization that facilitates the process and disclosed to the reviewers only after the scientific evaluation and final ranking are completed.

Following the recommendations of the Time Allocation Working Group (Patat 2018), ESO is considering moving to the DAPR approach for allocating its telescope time. As a first step in this direction, the Observing Programmes Office has proposed to have a dry run in P106, which would serve the double aim of introducing the ESO community to the new paradigm and to gain experience on the procedural and implementation aspects.

The dry run was announced to the community in the P106 Call for Proposals (p.2), via a note in the Science Newsletter and with a direct e-mail message to the PIs of the last eight semesters (about 2200 scientists). All OPC and Panel members for P106 were informed.

This document presents the guidelines that were provided to the applicants for assisting them in the anonymization of their proposals for Period 106. It also provides references to relevant publications and links to some of the material provided by HST following the deployment of DAPR in their Cycle 26.

On 10 February 2020, ALMA has announced that DAPR will be deployed in the upcoming Cycle 8 and published the anonymization guidelines for proposal preparation.

References


Further reading

Publications on the dual-anonymous review and related topics can be found in the following compilation prepared by the HST Working Group on the Anonymization of Proposals.

17. Wittman, H., Hendricks, M., Straus, S., Tannenbaum, C., 2019, Are gender gaps due to evaluations of the applicant or the science? A natural experiment at a national funding agency. The Lancet 393, 531-540. doi:10.1101/232868

Additional material

1. Single-anonymous versus dual-anonymous peer review
2. Types of peer-review
3. Elsevier dual-anonymous peer-review guidelines
4. Importance of dual-anonymous reviews
5. Is dual-anonymous peer-review better?
Guidelines for applicants

This section provides the guidelines the applicants should follow for anonymizing their proposals, to help conceal the identity of the team and to ensure a fairer evaluation process. For the dry run of Period 106, all the team information (names, publications, previous usage, …) provided by the users via the p1 submission tool will still be displayed in the proposal version that will be distributed to the referees (while in a full DAPR deployment this would be fully obscured). Nevertheless, the applicants should start practicing with the anonymization procedures, especially when preparing their scientific rationale before uploading the PDF into the p1 tool. Here follow the main criteria the users should comply to:

- Do not include applicant names or affiliations anywhere in the scientific rationale. This refers to all parts of the document, including possible footers, headers, diagrams, figures, captions, watermarks, …
- The same applies to the proposal title, abstract and the special remarks box.
- When citing literature and providing references, especially in the case of self-referencing, third person neutral wording must be used. For instance, the sentence “as we have demonstrated in Galilei (1610)” must be rephrased as follows: “as shown by Galilei (1610)”.
- References to previous work with ESO telescopes or other facilities must not be phrased in a way that would reveal the team’s identity. For instance, a sentence like: “we observed the moons of Jupiter with the VLT programme ESO001” must be replaced by: “the moons of Jupiter were already observed by the VLT by programme ESO001”.
- Acknowledgments must be avoided, and so must references to possible grant funding.

Complying to the above guidelines requires working on the grammar and structure of the scientific rationale. As a consequence, it will not be possible to re-cycle previous material without reviewing it in the light of the DAPR requirements. The proposing teams should take this into account when planning their submission/s, because text anonymization requires some extra effort. In this process, the applicants should not be discouraged from describing and discussing the previous work in the field (theirs or by other scientists), provided they do it complying to the above guidelines.

Team Background and Expertise

As part of the DAPR implementation, the applicants will be required to fill in the new Team Background and Expertise (B&E) section in the p1 form. This is meant to be a short description of the background, expertise and roles of the various team members in the context of the science case discussed in the proposal. This section is limited in space. For small teams the applicants may wish to provide a sentence for the qualifications of each member, while for larger teams (e.g. in Large Programmes), only the leading roles need to be specified. Here follows an example.
G. Galilei has expertise in VLT data reduction for X-Shooter and FORS2 which are requested in this proposal. G. F Sagredo is an expert in planet formation and F. Salviati is leading the field of planetesimal accretion theory. V. Viviani is a student of prof. G. Galilei and is working on his PhD thesis on the moons of Jupiter. The team has a strong background in ground-based observations in UV, optical and NIR, with specific competences in the field of planetary science, both experimental and theoretical. The team has published ten papers and a book on the subject in the last five years. The same team has obtained time on the Highest Spheres Telescope (HST#0123 for a complementary study. This project is supported by the GDT Grant #0017 from the Grand Duchy of Tuscany.

Guidelines for reviewers

In the dry run of Period 106 anonymization will not be mandatory and the DAPR procedure will not be enforced on the reviewers’ side. However, the dry run will be used to train the Panel members in sight of a future full deployment. The reviewers should follow these guidelines:

- The review must aim at selecting the most promising proposals, not the best proposing teams;
- The ranking is purely based on the scientific merit of the proposals: the pre-meeting review and the panel discussions must focus on science only;
- The background and expertise of the applicants with ESO or other scientific facilities is not to be considered;
- The referees should not try to guess the PI’s or the team’s identity;
- The chairs of the panels must refocus the discussion whenever this moves to the team identity, expertise or publication record;
- The referees should flag to the Observing Programmes Office all the cases that they reckon do not comply with the anonymization rules spelled out in the Guidelines.
- In the dry run of P106 this will only be used for improving the procedures and the instructions to the users.
- In their final comments to the applicants, the panels should provide feedback in case the proposal does not comply with the anonymization rules.

Frequently Asked Questions on Dual-Anonymous Peer Review

- Why is ESO considering Dual-Anonymous Peer Review (DAPR)?
- Is it really possible to fully anonymize a proposal?
- If the review is to be anonymous, why am I asked to provide information about my team?
- How can you demonstrate that DAPR is better than single-anonymous peer review?
- Do you have evidence that DAPR reduces the systematics?
- How will ESO make sure that the reviewers do not waste time in guessing the team’s identity?
- What does DAPR mean for the applicants?
- How will ESO make sure the time is allocated to teams capable of dealing with the data?
- How will the review process change?
- How does ESO make sure the accepted proposals are feasible if the Panels cannot access the team’s past experience?
- I have obtained telescope time for the upcoming cycle at other world-class facilities for the same science case. How do I make this known to the reviewers?
- Our team has been obtaining time at this facility for many cycles. How will ESO guarantee the continuity of our successful project?
- As a referee, how can I be sure that I allocate time to the proper teams?
- Will this affect also Large Programmes?
- What happens if a proposal does not comply with the anonymization rules?
• My proposal is so unique that it is impossible to make it fully anonymous. Will it be disqualified?
• Will proposal disqualification be applied in Period 106?
• How will ESO deal with conflicts if the team identity is not accessible to the reviewers?
• When will ESO deploy Dual-Anonymous peer-review?

Why is ESO considering Dual-Anonymous Peer Review (DAPR)?

The goal is to minimize potential sources of systematics and put the referees in a better condition for focusing on the science case of a proposal. Conscious or unconscious biases can play an important role in creating systematic effects in the process. For instance, they are likely at least partially responsible for the observed difference in the success rates of proposal submitted to ESO by female and male PIs (see Patat 2016). While it is understood that this is a complex problem, there are clear indications that DAPR can reduce the systematics that affect gender, scientific seniority, affiliation, ethnicity and so on. In addition, the case of HST has shown that proposal anonymization helps a lot the panels to concentrate on the science. In this respect, it is also worth noting that HST TAC members reported that this type of review is actually a lot less stressful and more collegial, which is an important consideration given the increasing number of large-scale proposal reviews.

Is it really possible to fully anonymize a proposal?

Certainly not in all cases. However, with the exception of very special and unique instances, guessing the identity of a proposing team is more difficult than one would think. Statistical studies conducted in other fields show that even when the overall team identity is correctly guessed, the PI identity remains unknown in the majority of the cases. It is important to emphasize that the ultimate aim is NOT to make it impossible for anyone to guess who is on the proposal, but rather to change the tenor of the discussion so the focus is on the science, not the scientist. The main goal is to level the playing field for everyone, not address specific imbalances within particular sub-groups.

If the review is to be anonymous, why am I asked to provide information about my team?

Anonymous review does not mean that the proposals are anonymous, but that the team information is kept separated from the scientific evaluation process. The Observatory, which facilitates the process, has full access to all the information as the Director General is the ultimate responsible for the time allocation process. The information you provide on your team is initially concealed to the reviewers and is disclosed only at the end of the process. Therefore, indeed, you have to provide it when you prepare your proposal.
How can you demonstrate that DAPR is better than single-anonymous peer review?

The deployment of DAPR is one of the actions that are being taken in order to reduce systematics that are known to affect single-anonymous peer-review (SAPR). The purpose is to improve the equity and integrity of the review process, in the attempt of making the selection fairer. ESO will keep monitoring the success rates (by country, gender, professional seniority, …), and this will allow to quantify the effect and take further corrective measures in case of need. After the recent requirement for providing more detailed information about applicants and referees in the ESO User Portal (see this announcement), more accurate studies will soon be possible. And, of course, one can also ask the same question for SAPR.

Do you have evidence that DAPR reduces the systematics?

DAPR has already been deployed for more than ten years in different fields outside astronomy and the results have been documented in a number of publications. The conclusion is that the obfuscation of team information turns into a reduction of the systematics in the review process. The HST Working Group on the anonymization of proposals has compiled a list of relevant publications. See Johnson & Kirk (2020) for the specific case of HST.

How will ESO make sure that the reviewers do not waste time in guessing the team’s identity?

During the pre-meeting phase when the proposals are reviewed online by the single referees there is no way of doing that. But this applies to any of the review guidelines and it is not specific to the single- vs. dual-anonymous implementation. In the case of HST, the Space Telescope Science Institute has introduced the role of “levelers”, who are in charge of making sure the discussions focus only on science. The levelers are selected from the STScI staff members. The implementation details and the possibility of having a similar entity in the ESO panels is being discussed.

What does DAPR mean for the applicants?

As described in the Guidelines for applicants, they will have to comply to some simple rules when preparing the title, the abstract, the special remarks and especially the scientific rationale of their proposals. This requires a change of style (grammar and structure) in the way their previous work (both publications and usage of scientific facilities) is described and used in the text to support the proposed case. This implies that previous versions of the proposal will need to be reviewed and adapted to the new anonymization requirements set by the DAPR paradigm. The applicants are also requested to fill in the Background and Expertise field in the p1 web form. This is supposed to provide a concise summary of the applicants’ profiles and expertise relevant for the proposed science case. This field will not be included in
the material distributed to the referees during the review and will only be accessible to them after the ranking phase is completed.

How will ESO make sure the time is allocated to teams capable of dealing with the data?

During the review process the Panels are instructed to take note of the cases which require special competences or expertise levels. When the team information will be made available to them, they will be in the position of making a recommendation to the Observatory, expressing their possible concerns on the qualifications of the team with respect to the specific science case. However, they will not be in the position of changing the ranking, which is purely based on the scientific merit of the proposal. The can only make a recommendation to the Director General for possible actions.

How will the review process change?

In terms of the various steps (asynchronous online grading, triage, face-to-face discussion and grading) nothing changes up to the final compilation of the merged ranking lists per telescope, which are produced by the Observing Programmes Office once the review is completed. At that point, the information on the proposing teams, which was concealed to the reviewers, is made accessible to the Panel Chairs. This is examined and the cases judged as critical during the review are discussed and flagged to the Observatory for possible disqualification. Reasons for disqualification include violations of the anonymization rules, unethical behavior (e.g. deliberate misrepresentation of the team’s expertise and/or available resources) and justified concerns on the applicant’s competences for the specific science case. The disqualification flag is only a recommendation, and the final decision will be taken by the Director General, who will also evaluate the possibility of flagging obvious misbehaviors to the host institutes of the applicants.

How does ESO make sure the accepted proposals are feasible if the Panels cannot access the team’s past experience?

From the technical point of view, and as in the case of the current SAPR paradigm, the accepted proposals are subject to a feasibility assessment by the instrument specialists from the Observatory. From the scientific point of view, the reviewers have access to this information AFTER having expressed their opinion on the scientific merit of the proposal, which was formulated without being influenced by possible concerns on the team’s profile. At that point the panels can flag their concerns about the reliability/suitability of the team and make a recommendation to that effect.
I have obtained telescope time for the upcoming cycle at other world-class facilities for the same science case. How do I make this known to the reviewers?

This can be done in a fully non-anonymous way in the Background & Expertise field, which is made available to the panels only when the ranking is completed. In the anonymized proposal text this has to be mentioned in an anonymous way.

Our team has been obtaining time at this facility for many cycles. How will ESO guarantee the continuity of our successful project?

If your project has a strong approval record it means that it is based on a very robust science case, and hence it has nothing to fear from the DAPR process. You can discuss and present all the previous work in the scientific rationale in an anonymous way, hence making clear to the reviewers that this is an important topic, that has been systematically and successfully followed in the past. This will make the panel aware of the fact that a significant amount of time and effort was invested on the specific topic, properly setting the stage for the discussion as to whether the proposal makes a sufficiently compelling case for the need of new data.

As a referee, how can I be sure that I allocate time to the proper teams?

Technically, the single referees and, more in general, the Panels do not allocate time: they only make a recommendation to the Director General in the form of a ranked list. The final allocation is then the result of a complex process which involves a number of stakeholders and constraints on which the Panels do not have any control (this includes technical activities, atmospheric conditions, instrument loads, right ascension distributions, carry-over from previous periods, ongoing Large Programmes, etc.). In any case, both in SAPR and DAPR the panels have the possibility of expressing their concerns about the teams. In the DAPR paradigm this happens after the rankings are compiled, but the possibility of flagging critical cases is still present, without the negative drawback of influencing the scientific judgement. It is the Observatory which, based on the Panel recommendations, is responsible for allocating the time so that the scientific return is maximized. Finally, there is one further aspect. If the case is judged to be strong, it means that, scientifically speaking, the Panels reckon it is worth that ESO secures those data.

Will this affect also Large Programmes?

Yes. All proposals submitted for ESO telescopes will be subject to DAPR, and all applicants will have to comply with the anonymization rules. This will also apply to Director’s Discretionary Time (DDT) proposals. For Large Programmes, the Background & Expertise section may be longer.
What happens if a proposal does not comply with the anonymization rules?

Alleged violations to the anonymization rules will be flagged to the Director General and discussed internally. In case of major abuses, the corresponding proposals will be disqualified and excluded from any subsequent review step. Less serious infringements (e.g. usage of first person instead of neutral third person) will be flagged and discussed but will not be removed from the review in the panels. A final decision for those cases will be taken at the end of the process. Proper feedback on possible violations of the anonymization rules will be sent to the applicants.

My proposal is so unique that it is impossible to make it fully anonymous. Will it be disqualified?

No, if you have followed the anonymization rules. You should not obfuscate your science case and your models (and, by doing that, render your case weaker) to make you and/or your team anonymous. But you must use the neutral third person and cite your work as indicated in the anonymization guidelines, so that your identity is not deliberately revealed to the reviewers.

Will proposal disqualification be applied in Period 106?

No. Period 106 will only be used for testing and information purposes. Failing to properly anonymize the proposal will have no consequences on the review. The proposals distributed to the panels will not be anonymized, and the review process will not differ from the past.

How will ESO deal with conflicts if the team identity is not accessible to the reviewers?

Institutional conflicts will be identified and handled based on the User Portal profiles of the applicants and the referees, as is the case in the current implementation. With DAPR, personal conflicts become, to some extent, simpler to manage. If a referee feels in conflict with a given proposal because s/he has a conflict with one of the applicants, the conflict itself (which would translate into the inability of expressing an objective judgement) is created by the fact that the names are disclosed. If they are not, the conflict simply does not exist.

For the scientific conflicts, these are generated by the science case and not by the team who wrote it. In this case, the referee is in the position of declaring that the science case is too close to her/his research field, and hence s/he may not be able to express an objective opinion (regardless of the fact that the proposal may have been written by a competing team). In this case there is no difference with SAPR, in that this is fully in the hands of the reviewer, as in both cases s/he may or may not declare the conflict.
When will ESO deploy Dual-Anonymous peer-review?

Period 106 will be used as a dry run for testing various aspects of the process and giving the community the possibility of getting familiar with the anonymization procedures. The implementation of the full process in production is being discussed with the relevant advisory boards (Observing Programmes Committee, Users Committee, Scientific Technical Committee). The community will be informed in due time about the implementation timeline.

An example of text anonymization

Non anonymized text

Since 1609 our group has been accumulating evidence that earth rotates around the sun (Galilei 1610). In particular, our observations obtained at the Earliest Sidereal Observatory (under programmes ESO001, ESO002; Principal Investigator: Galilei) have demonstrated that Venus shows phases with a pattern that cannot be reconciled with the geocentric theory (Ptolemy 0150). In addition, we have discovered that at least four satellites orbit around Jupiter (Galilei 1610), hence demonstrating that not all celestial bodies rotate around the sun.

We have also provided further evidence of the incorrectness of the Ptolemaic model (Ptolemy 0150) with the observations of the sun we carried out with ESO telescopes (programmes ESO003 and ESO004; PI Galileo). As we discussed in Galilei (1613), we concluded that the surface of the sun is disturbed by topological defects which are not stationary and evolve with time (see also Scheiner 1612 for an alternative interpretation by a competing team). This is in evident conflict with the claims that all celestial bodies are incorruptible and composed by perfect matter (Aristoteles 0350 BC; Ptolemy 0150 and references therein). With this proposal we plan to finally demonstrate that our preliminary conclusion that earth orbits around the sun (Galileo, Salviati & Sagredo 1632) is correct, by measuring the parallax of a sub-set of the stars included in the HSC (Hipparchus 0135 BC) accessible from the southern hemisphere. With these data we will be able to settle once and forever the matter on the chief world systems (Copernicus 1543 and references therein). We have been granted 30 orbits at the Highest Spheres Telescope (programme HST#0123) to cover the part of the HSC sample with parallaxes below the limit that can be achieved from the ground. This project is supported by the GDT Grant #0017 from the Grand Duchy of Tuscany.

The same text after anonymization

Evidence that earth rotates around the sun has been accumulating since 1609 (Galilei 1610). In particular, observations obtained at the Earliest Sidereal Observatory (under programmes ESO001, ESO002) have demonstrated that Venus shows phases with a pattern that cannot be reconciled with the geocentric theory (Ptolemy 0150). In addition, it has been shown that
at least four satellites orbit around Jupiter (Galilei 1610), hence demonstrating that not all celestial bodies rotate around the sun.

Further evidence of the incorrectness of the Ptolemaic model (Ptolemy 0150) was provided by the observations of the sun carried out with ESO telescopes (programmes ESO003 and ESO004). Galilei (1613) has concluded that the surface of the sun is disturbed by topological defects which are not stationary and evolve with time. This is in evident conflict with the claims that all celestial bodies are incorruptible and composed by perfect matter (Aristoteles 0350 BC; Ptolemy 0150 and references therein). With this proposal we plan to finally demonstrate that the preliminary conclusion that earth orbits around the sun (Galileo, Salviati & Sagredo 1632) is correct, by measuring the parallax of a sub-set of the stars included in the HSC (Hipparchus 0135 BC) accessible from the southern hemisphere. With these data we will be able to settle once and forever the matter on the chief world systems (Copernicus 1543 and references therein). Thirty orbits have been granted at the Highest Spheres Telescope to cover the part of the HSC sample with parallaxes below the limit that can be achieved from the ground.

Acknowledgements

These guidelines were put together based on the experience gained at the Space Telescope Science Institute. The discussions with Neill Reid and Lou Strolger during the Dual Anonymous Proposal Review Workshop (Baltimore, September 25, 2019) and the help provided during the preparation of these guidelines are kindly acknowledged.