

Director's Perspective: Small Differences

Steven Beckwith, svwb@stsci.edu

Fritz Strobl of Austria won the men's downhill event at the 2002 Winter Olympics in 1:39.13. Kjetil Andre Aamodt of Norway posted 1:39.78 to come in fourth. His performance lagged Strobl's by just 0.7%, but that was enough to knock him out of the medals. Fifteenth place went to Franco Cavegn (Switzerland) with a time of 1:40.81, 1.7% off the best mark, and 1.4% away from a medal. At 85 miles per hour, the 0.65 seconds separating gold and fourth place corresponds to 25 meters, an unmistakable margin if the races were run side-by-side, yet almost imperceptible as a percentage of performance. A small burr on a ski's edge, a slight flare on a turn, too much airtime over a bump, and a racer could plummet from first to eighteenth. Only 1/5 of the skiers between Cavegn and Strobl would medal, about the same as the fraction of Hubble proposals that are granted telescope time.

Electronic timing eliminates controversy from the downhill race. Not so for women's figure skating. Sarah Hughes won the gold with a stunning free skate that brought her from fourth to first, dropped the leader, Michelle Kwan, to third, and knocked Sasha Cohen out of the medals. Nine judges watched the athletes go through their routines and graded them on their performance. The winners were decided by a rating system that is confusing to the untrained eye. A slight 'double foot' when landing a triple Axel, insufficient stretch during a spin, or a poor extension on a wide spiral merits a deduction from the perfect mark.

The final score in figure skating is decided in two parts: a mark for technical merit, the product of the degree of difficulty of the routine and the perfection of its performance, and a mark for artistic presentation, a measure of how much the judges liked the combination of jumps, spins, and footwork chosen by the skater. There is considerable latitude in assigning this second mark: witness the controversy over the pairs competition between the Russian and Canadian couples. As with downhill skiing, the separation between the winner and fourth place is tiny.

The rating of proposals for Hubble time closely parallels the judging of figure skating. Proposals are rated first on technical merit and then on the scientific equivalent of artistic performance. A stumble on technical merit, like a fall in figure skating, is usually enough to knock the proposal out of the competition for a time award. Most proposals pass muster on technical merit, and the medalists are selected on the scientific equivalent of artistic performance: is the question really interesting? Is the technique clever? Most importantly, do the proposers demonstrate a clear understanding of current norms in scientific thinking, showing they will advance knowledge in predictable ways that their peers can understand? This latter demonstration is a lot like choosing music and choreography to suit the judges. What works for one Telescope Allocation Committee (TAC) may fail for another.

Just as in Olympic events, the separation between the best proposal and the cutoff is exceedingly small. Almost all members of the TAC and the subject panels have

expressed frustration that they differentiate between proposals on unquantifiable subtleties. We cannot make sound judgments about the likelihood of great science near the margin between successful and unsuccessful proposals. Nevertheless, just as there are only 3 medals for which the top 15 ice skaters compete, there is only enough Hubble time to award to about 1 of every 5 proposals. The TAC must find reasons to make those judgments, even though the reasons are small differences in perception.

Just as there are protests about Olympic judging, there are protests about the decisions made by the TAC. After every cycle, I receive correspondence from unlucky proposers who insist that their proposals were rated unfairly. Usually, poorly worded comments in the notification letters stimulate the protests—comments suggesting the TAC misunderstood the technical merit or did not value what is truly chic in science. Indeed, the TAC often fails to ‘get’ the proposal, usually because the message is unclear or unpersuasive. We assume that the burden is on the proposer to write cogently for non-specialists. Not all proposers feel that way. My job is to explain why the decision will not be reversed and why an appeal is almost always futile: because we cannot run the TAC process twice.

In truth, the problem is one of small differences. We get about three times as many interesting, technically meritorious proposals as we can approve. There is nothing wrong with the 2/3 of these proposals that do not get time. A brilliant proposal may get time in one cycle and be turned down in the next only because the panel uses slightly different criteria for ‘artistic presentation’. As with Olympic judging, there is some consistency from year to year, with the known favorites being looked upon favorably and some corporate memory of long term programs. There is, nevertheless, a large element of chance; past performance is not a good indicator of future promise, a concern voiced bitterly by a number of luminaries this cycle, who believed the very fact that they did not get time indicated a problem with the process. Michelle Kwan and Picabo Street probably feel the same way.

One of the harder parts of my job is to reply to people who cannot understand why their proposals are turned down. I take no pleasure in these replies. Luck is a factor in any selection process with such a high rejection ratio. I generally sympathize with the protest, but there is rarely a practical way to change the outcome.

There is one obvious solution to my dilemma. We need three Hubble Space Telescopes. Three Hubbles would produce three times as many great discoveries as the current one does. The ideas are there. Our ability to predict which ideas are best is modest or worse. If chance is at work, the increase in science productivity will be proportional to the resources we devote up to the limit of meritorious ideas.

So, the next time you suffer at the hands of your peers and feel inclined to write a letter of protest to me, put your energy to more productive use by writing your government representatives. Tell them that we need three Hubble Space Telescopes. Please don’t tell me your proposal really deserved time. I already know it was deserving. It was edged out of the medal competition by small differences.