Briefing 2:

Adjudication at the WUDC Berlin 2013

With very many thanks to the Adjudication Team from the De La Salle University WUDC 2012, who drafted the original version of this guidance. Our updates have been minor in comparison to their extensive work. for this document. Our amendments have been few and minor.

1. Deciding who wins
2. Decision-Making Process
3. Deliberation
4. Announcing the Result
5. Feedback on Adjudicators
6. Some Pitfalls to Avoid in Decision-Making and Feedback

1. Deciding who wins

Criteria

What constitutes a good/bad policy is itself debatable. Teams are allowed to debate the criteria for assessing whether a policy is good as part of arguing that it is, in fact, good. Judges should adjudicate this debate about criteria, **as opposed to just applying their own criteria.**

One common form of this mistake is to assume a **utilitarian framework,** and assess the debate in that way. If this is disconnected from a team arguing for a utilitarian framework, it is wrong. It is also wrong to disregard principled argumentation about which effects are particularly important, potentially for non-utility maximising reasons. Sometimes, of course, all teams in a debate agree on the framework to use, perhaps implicitly. In these cases, the judge should take on these consensus assumptions.

So, judges should listen to teams' arguments about what our aims and principles should be, and evaluate the claims of harms or benefits in that context. This *can* make these claims particularly vital; if in the debate 'THW Invade North Korea you show, for instance, 'war is always wrong, regardless of the practical benefits', the other side's claims about the practical benefits of invading North Korea are simply relevant until they've proved a practical calculus is relevant. Note that there is no sense in which this argument is too 'generic' to gain credit. There is an **absolute requirement** that arguments be relevant to a debate. There is **absolutely no requirement** that arguments pertain *only* to that debate.

Assessing Argumentation

The outcome of the debate should obviously depend on what the teams do and don't say. Judges must **not intervene** in the debate. Don't invent arguments for teams, don't complete arguments and don't rebut arguments.

Judges often speak about a speaker "getting at a good point" even if "(s)he didn't quite get there". This is just an excuse for a judge to invent an argument (s)he'd like to hear. Don't do it.

As we don't do teams' rebuttal for them, we don't consider claims invalid just because we disagree, or because we can see holes in their arguments. Arguments stand as 'true' once they are made until they are responded to. (By 'argument' here we mean 'a set of words

that gives us at least some reason to believe one conclusion over another'). This has an important implication: if OG, for instance, make arguments whose conclusion is 'we should do the policy' which everyone ignores, then they don't lose because 'the debate moved on from them'. Rather, their unrebutted arguments are **still true** and should be weighed as such. In fact, teams like this very often should **beat** the teams on the opposite bench.

There is one legitimate exception to this – if an argument is clearly absurd (such that you cannot conceive of anyone we would consider sane believing its logic and/or premises) **and** it was very marginal to the speech of the speaker making the claim, then it is reasonable for a team to decide to spend their time elsewhere, particularly where there is other stronger material in the round.

What is and is not rebutted is therefore of vital importance to judging debates. Note that speakers don't have to use the word "rebuttal" so respond to an argument. It is tidier if they do, but judges should not ignore material that adequately deals with an argument just because the speaker doesn't point out that it does.

Equally, this doesn't mean speakers should be 'punished' for not refuting everything: some claims do not do any harm at all to the opposite side. For example, in a debate about the legalisation of drugs, if the government say 'pink elephants are cute because they have those nice ears and are a pleasant colour', this flawed argument can be safely allowed to stand as it isn't a reason to legalise drugs. There is, therefore, no need to point out that blue elephants are obviously more tasteful. So too, if they said 'some drugs are less harmful than others', this could also be ignored. While it is clearly more related to the debate than the cute pink elephants argument, it is *pre-argumentative* – that is, it has not got sufficient surrounding words to be a reason to do or not do the policy. Opp can quite happily say 'yes, some drugs are more harmful than others' and move on. Or just ignore this argumentative non-sequitur. Often as a judge, it can be tempting to complete arguments for teams that are interesting but pre-argumentative. Don't.

Certain things **don't matter** (in themselves) in debating:

- The number of arguments a team makes,
- How clever or sophisticated the argument was,
- How interesting the argument was.

What matters, once an argument is made, is how important its conclusion seems to be in the debate, based on what people have said in this debate, and the extent to which it seems to be responded to (and how robust it is under such attacks). We should not here consider how important *we* thought a particular argument was, in the abstract, but rather how important it became in this particular debate. Similarly, we should not care how new, interesting, sophisticated or so on an argument was, except in so far as those things tacitly influence us to find those claims to be more true or important and have impacts on how well others can and do respond.

We've now come full circle. The team who seems to win the logical argument wins the debate. This must be assessed **solely through material provided in the debate**. Don't penalise teams for failing to make arguments. Don't consider arguments that you're aware of but which weren't made.

Mechanisms for persuasion

Many debating coaches decompose persuasiveness into different categories, which they then train teams to get better at. It is important to recognise that these are different mechanisms for persuasion, but they are means to that end.

Examples of such mechanisms are:

- Good analysis
- Good rebuttal
- Good style
- Good strategic decisions
- Using examples

If these are done well they might influence you to believe the team was winning the logical argument. That's great for the team in question. But note that you shouldn't *also* consider them as separate categories for assessment. Doing so would be double counting them. While it may be helpful in coaching to dissect debating and examine these factors separately, as judges we just see what persuasive logic seems to have been produced, much as football referees don't mark players' passing and shooting, but just count the goals.

One thing which is **certainly unpersuasive** is inconsistency. Teams should not contradict themselves or their bench partners. Internally inconsistent teams cannot simultaneously get credit for two areas of mutually exclusive argument. We're **forced** to credit them for the least effective of the two.

Judging as the ordinary intelligent voter

There is obviously a need for judges to abstract themselves from their personal positions when in debates. If we simply judged as ourselves then the result would depend strongly on what we know, how we think and what opinions we hold. We would be making personal judgments about how compelling *we* find particular arguments. This would give teams who know us a great advantage, and would make the results more about us than them. Therefore, we attempt to judge as an abstract 'ordinary intelligent voter', and minimise the impact that our knowledge and opinions have on the result. This is a difficult process, and something that bears thinking about.

It is important, however, to remember that thinking 'as' the 'ordinary intelligent voter' does not absolve us from our responsibilities to actually judge the debate – that is to evaluate the logical flow of arguments and determine the extent to which teams have seemed to win them. We should not say 'While that was clearly irrational rabble-rousing, the ordinary intelligent person would have fallen for it'. This not only leads to irrational conclusions, but also, generally, overestimates how much cleverer we are than the ordinary intelligent voter.

There is another important feature of how the ordinary intelligent voter judges, as distinct from how we do. Many of us debate a lot, and we develop aesthetic preferences about speaking as well as in-jokes and references which we find terribly funny. This is completely natural, but distracts somewhat from debating. As voters we are much less likely to credit policies for being advocated in a 'sophisticated' or 'funny' way; these things might lead us to believe them better, but we would not say we voted for someone because she was funny, but because she'd make our life better. For instance, a speaker might be very

amusing by making sarcastic remarks playing on the weakness or callousness of their case. This might lead us to find the speaker amusing, but would make the ordinary intelligent voter substantially less likely to support the policy. We as judges should remember that we are not aiming to evaluate who was cleverest, neatest or funniest, but who best used their cleverness, neatness and funniness to persuade that the policy was a good or a bad idea. The best way to do that is simply to address oneself towards debates as if there are real policies at stake and see who best persuades that the policy should or shouldn't be done. **This is the *only* reason we use the term 'voter'.**

The ordinary intelligent voter:

- Attempts to evaluate logically what the best thing to do is,
- Using as inputs the arguments made by the teams,
- Persuaded by a variety of tools which are not strictly logical, such as,
- Style, neatness, clarity, etc. and their inverses.

Critically, the ordinary intelligent voter is **actually open to persuasion**, as some actively try to avoid being!

You are probably not ordinary

Many of us (many of *you*, at any rate) know large amounts of knowledge and technical language that would be alien to many. The ordinary intelligent person has the sort of knowledge you'd expect from someone who read newspapers occasionally, but not frequently or cover-to-cover. She is not a politics geek, but has a broad understanding of major world issues and movements in high-level terms, but does not know the detail of individual conflicts, regional bodies or anything but the very most prominent countries' domestic situations.

You should certainly make reference to facts and details the ordinary intelligent person does not understand, but you should *explain* rather than *citing* these examples, facts and countries. While she may not *know* much by some of your standards, the ordinary intelligent person is genuinely intelligent, and understands complex concepts once they're explained. Where such examples are not explained beyond name-checking a country, judges should discount any material they *do* understand because the ordinary intelligent person would not (though there may be a small stylistic benefit to sounding-well-informed even as the content is lost). Judges should be bold in applying this rule: it is unfair on other teams in the room not to.

Once again, note our approach: judges should **ignore** material which is not explained, as opposed to actively punishing it.

The Ordinary Intelligent Voter comes from nowhere

Our ordinary intelligent voter does not come from where you do; there are no 'domestic examples' requiring less explanation, **even where everyone in the room comes from that country**. Wherever you are from, assume your audience is from somewhere else.

This is an English-language competition, so our voter understands English. There are, however, multiple Englishes, and one should not use terms one does not expect international English-speakers to understand. This does not mean you must pitch your speech to someone who has only *some* English, but rather that you should recognise that an Austra-

lian English speaker has a different vocabulary to an English English speaker, who has a still different English to a Chinese English speaker. It is a good thing to make your English accessible. It is also part of the rules of this tournament that domestic colloquialisms do not count as comprehensible, unless our Ordinary Intelligent Voter would be able to figure them out without trouble from the context. Consequently, if you make an argument using them, you may not in fact count as having made it. Our advice: don't risk it.

Following on from the above, the ordinary intelligent person **does not know technical terms** that one would require a particular university degree to understand. She can be assumed to possess the sort of generalist vocabulary that comes from *a* university education of some sort, but probably not yours. He does not have the sort of halfway-there economic or legal jargon that we as debaters have become familiar with either. Saying 'Laffer curve' to most people is equivalent to making some clever-sounding noises. Similarly, using terms like 'economic efficiency' will lead to their being understood only as a layperson would grasp them, losing any technical specificity. Judges should judge accordingly.

Giving credit where it is due

We'd like to say something about awarding credit to teams for the material they bring to the debate. In particular, we want to answer the question "what counts as new material?"

Anything teams say that has not already been said in the *same way* is a new contribution and worthy of credit (or, if it is harmful, debit). This means they get credit for **everything** beneficial they do: arguments, rebuttals, framing, examples, etc. Their contribution is valuable insofar as what they say benefits their side more than what has already been said. So, a whole new argument that is entirely theirs is valuable to the extent that argument helps their side; an improvement on an existing argument is valuable insofar as *their addition to the argument* helps their side; rebuttal is valuable insofar as *defeating that specific material* helps their side. All this is determined by the current state-of-play.

One conclusion of this is that it is **not enough** for a closing team to make an argument *somewhat* better than their opening: they gain credit only for the improvement on the previous argument. They must make improvements and additions to the opening *case* that outweighs the totality of that case.

Comparing contributions

Where teams have a chance to rebut each other, assessing relative contribution is easy. Judges should track the argument and assess, given their responses to each other, which team's contribution was more significant in furthering their cause to logically persuade us that we should do the policy, or that we should not.

Where teams don't get a chance to rebut others, this is trickier. This happens fairly often:

- Teams on diagonals
- O4 explains something in a new way
- Opening teams shut out of POIs

In these circumstances judges are forced to perform some more independent assessment of the 'robustness' of the arguments teams made. The question to ask is 'how robust to

potential rebuttals was a team's contribution?' More 'robust' material is a greater contribution than less 'robust' material.

Ideally, assessing robustness will involve a comparison with material on the table, or very minor extensions thereof. For instance, when comparing with a team diagonally in front, judges should first ask whether anything in that team's case is relevant or (ideally) rebuttive. Did the team being assessed deal with this material? Check whether they allowed the diagonal team in on POIs, to give them an opportunity to engage. Deliberately shutting out engagement from a team whose material is relevant is often obvious and deeply unpersuasive.

2. Decision-Making Process

Once the debate has finished, the judges should decide the rankings of the teams according to the following rules:

- Please remember the criterion: *Judges assess which team did best to persuade them by logical argument that the motion ought to be carried or rejected. The judges do so as the 'ordinary intelligent voter', applying a very small set of technical rules, which are there to facilitate, not displace, an analysis of arguments.*
- The 15 minute time limit

The aim of the discussion is consensus. Decide the ranking first, with no consideration of speaker marks until this has been established. This reflects the fact that **teams win debates**, not speakers, and they win based on their aggregate contribution. There is **no problem at all** with the vast majority of that contribution coming in one speech. We are not evaluating our aesthetic appreciation of the speeches (or proxy-marking 'team balance'): we're assessing the team's aggregate contribution. Imbalance within a team should be reflected by giving the speakers different speaker marks.

Note that the **time limit is a rule of the tournament.** That is all the time you have. Chairs of panels must manage their time accordingly, and recognise that the rules require a vote if no consensus has been reached early enough for the adjudication to **complete** in 15 minutes. Taking into account the time taken to decide on speaker marks, this means you should consider a vote around **12 minutes into a discussion.**

Votes aim to establish a majority, with the chair holding a casting vote in case of an even split. In the case that there is absolutely no agreement, vote rank-by-rank from first to fourth.

3. Deliberation

In close rounds, we expect that the judges on the panel will have different views on the debate. Therefore, achieving consensus in 10-12 minutes is a difficult task, requiring careful management by the chair judge. Here we sketch some suggestions for how this could be managed.

The chair should begin by asking each wing to give a ranking of the teams. This is not binding, it is a working hypothesis which will evolve as the discussion progresses. Wing

[redacted]

judges should not feel any pressure to agree with one another or the chair in their initial call, as there is no negative implication for changing your call. Judges should have *some* opinion of the debate as soon as it is done, and should share their leanings along with their uncertainties. Having *no idea whatsoever* until five minutes of note-reading has passed does not bespeak an active following of the debate.

The chair should then assess the level of consensus which exists. There are thousands of possible combinations, but thankfully a few scenarios crop up fairly often.

- Everyone has exactly the same rankings – celebrate (but briefly). Have a brief discussion to ensure rankings are the same for similar reasons. Move on to scoring.
- Everyone has the same except 1 person – ask him to defend his position. Be specific, tailoring the requested defence to the difference between the minority and majority opinion. If it is a difference of one team, focus on that team, etc.
- There is similarity in rankings (you agree on where 1 team ranked or some relative rankings – everyone agrees OG is better than CG) but also some crucial differences. Begin by establishing which discussions need to happen (i.e. there is disagreement about whether OO beat OG). Begin by consolidating the consensus that exists, and use this as a platform to break deadlocks.
- Chaos. There is no similarity between the rankings. Guide a discussion of each team's arguments, or, depending on what makes sense to you and in context, of the clashes between particular pairs of teams.
    - These debates often hinge on how one argument was evaluated, so your aim is to detect such differences in interpretation. The initial discussion is intended to inform each other of your perspectives and find some level of common understanding.
    - If two judges believe different arguments are central, frame a discussion about their relative priority. Get each judge to explain their position, and attempt to establish a metric for the importance of arguments in the debate.
    - After this brief discussion, rank the teams and compare again. If you have achieved some overlap, move on to the suggestions under (c) above. Vote if necessary.

Once a ranking has been decided upon, the chair should lead a quick determination of speaker scores. These should reflect the majority decision, and **not** be a compromise between various opinions. i.e. don't say "We all think OG wins, but we can make sure the speaks reflect your view". If the majority doesn't think a relative ranking is close, there is no reason that the speaker marks should reflect such.

Again, please remember that, while achieving a consensus is ideal, it is not an ideal that is to be placed above justice. If you believe that a team placed first and the other judges disagree, try to establish their reasons. If those reasons are unconvincing, you must do your best to ensure that the correct team wins the debate, though that may mean simply putting the issue to a vote. <u>Change your mind if you are persuaded- otherwise cling to what you believe to be right. The honours that a debating round dispenses belong to those who have earned them. They are not yours to give away in shabby compromises.</u>

4. Announcing the Result

The chair of the panel delivers the adjudication speech. In the case that the chair loses a vote, she may, if she chooses, retire from this position and ask one of the wing judges who voted in the majority to deliver the adjudication. If he does give the adjudication, this *must* be to defend the *majority* position, although the chair should overtly state they disagreed with the majority.

The adjudication speech should distinguish between *the reasons for the decision* and *advice for teams*. The reasons should be about what *did* happen, while advice is about what *didn't* happen. The latter cannot be a basis for the former.

The primary aim of an adjudication speech is to convey to the teams the reasoning of the panel in ranking the teams as they did. The speech should therefore present a logical argument for the ranking, using as evidence the arguments made in the debate and how they influenced the judges. Debates shouldn't be judged according to coaching models (either prescriptive models like 'problem/solution' or decompositions of persuasiveness like 'content, style, strategy').

The speech should be structured as followed:

- Announce the ranking of the teams,
- Go through the teams in an order that makes sense, comparing pairs of teams and explaining why one beat the other,
- Summarise and invite teams for more detailed feedback.

Comparing teams involves more than making isolated statements about Team X and Team Y, and saying 'so X clearly beat Y'. It requires that you explain the interaction between the teams to establish who had the better argument.

Be specific and be detailed – the vague application of adjectives is not sufficient judging. Identify arguments, whether and how they were responded to, and what the impact of the remainder was. Identify which teams get credit for what, and how this influenced your decision about whether or not we should do the policy.

We **strongly** suggest that the best way to give feedback is to speak to the teams and their contribution in chronological order. In other words, discuss first the contribution made by first proposition, then the contribution made by first opposition and why one was more persuasive than the other and the factors that went into that decision. Judges are not required to follow this format, but they are required to be comparative and specific.

Advice should be separated from the reasons for your decision, lest you confuse teams about which is which. There are a number of broad areas of advice you may want to give as a judge:

- General advice on how to improve
- Suggestions of reasons why things identified in the adjudication happened
- What might have been run (minimise this unless asked!)

5. Feedback on Adjudicators

We want to know how judges are doing, for two reasons. Firstly, to ensure we provide the fairest possible competition by putting the best judges in charge of panels, secondly, because judges care about their success in the tournament and try to improve their standings.

There are three types of feedback:

- Team feedback on the chair judge's adjudication speech,
- Chair judge's feedback on wing judges,
- Wing judges' feedback on chairs.

Each type is important. At this tournament we have developed an entirely new system for tracking feedback, in order to allow us to use it more efficiently. The only way we can use this system is if you participate.

6. Some Pitfalls to Avoid in Decision-Making and Feedback

What follows is a common set of mistakes that judges may make in determining results and giving feedback.

A) Dealing in generalities rather than specifics

"We thought that second opposition really brought the case home for us, so they won the debate."
"First proposition talked about rights, but I really didn't find it persuasive."
"First opposition had some interesting things to say, but the analysis didn't get better until second opposition."

It's perfectly fine for adjudicators to use general language to introduce their reasons, provided that each general statement is supported by examples of what actually happened. No statement of the sorts that we've listed above should ever go unsupported by specific examples of the claim being made, either during the deliberation or during feedback.

B) Failing to judge the debate as it happened.

"Proposition never talked about rights in this debate."
"It took until the summation speaker until we heard anything about rights."
"I really wouldn't have propped it like that."

Judges may have their own opinion as to what the best arguments for each side in the debate will be, but these are not the criteria on which the debate is to be judged. Judges may advise teams that there were interesting avenues of analysis left unexplored, but they may **not** penalise teams for their approach to the motion, or the things that each team decided to emphasize.

C) Granting certain "classes" of arguments undue priority

"Second opposition were the only team who told us about the importance of principles."
"Only first proposition knew the names of major Brazilian cities."

"Second prop won because their arguments were moral rather than practical."
"You didn't name any actual footballers in your speech."

This judging pitfall takes a number of forms. One of which is the fetishisation of the use of specific knowledge in the making of arguments. Teams which make strong arguments buttressed by good knowledge should be rewarded, but **not** because of the total amount of facts they named, but because of the strength of the arguments which those facts were marshalled in support of. A clever use facts makes an argument stronger and better, it does not make an argument.

A second form of this pitfall is according improper priority to arguments that are of various types (moral/philosophical/economic/practical). A 'principled' argument, for example, is not necessarily better or worse than a practical one. It depends what each argument seeks to prove and how well it does so.

D) 'Penalty judging'

"Your last point came after six minutes in your speech, so that really hurt your team."
"You didn't take any points of information, so there was no way you could come first."
"We had questions about the mechanism, so we put you last."

A good referee is not one who incessantly blows their whistle and stops play. Similarly, a good judge isn't one who tries to find as many reasons as possible to exclude consideration of a team's arguments and speak instead about the form- rather than the content- of their contribution. If a team violates the duties of role fulfilment, they should be penalised only up to the point of removing any harm they caused to the debate through failure to fulfil their role.

Some Examples:

- Not taking any POIs means that a speaker's material is to be viewed as less persuasive, not excluded from consideration.
- Lack of clarity in a mechanism should be resolved by allowing the opposition teams to make any reasonable assumptions of their own and letting the debate carry on from there.
- If a speaker introduces new arguments in an opposition summation speech, these are to be discounted, as though the speaker had said absolutely nothing.

E) Judging on Format Rather than Content

"You only spoke for one minute about the effect of this policy on Wales."
"You should have put your argument about rights first."
"Your team was unbalanced- all the good points came from the first speaker."
"You only spoke for five minutes."

Speaking for a certain length of time or placing arguments in a certain order is irrelevant (in and of itself) to which team won the debate. Naturally, speakers and teams who spend all their time on good arguments and spend more time explaining more important and more complex arguments will do better at being persuasive, but they succeed because they have made good arguments and have explained those arguments well, not because they 'spent time on them'. A speaker can win a debate with a one minute speech (but it's very, very hard to do so).