
ISU JUDGING REFORM

PSA INTERNATIONAL CONFERENCE

May 24, 2003

Dr. George Rossano

Revised Aug, 2003

The BIG Issues

- Honesty and Integrity of Officials
- Method of Evaluating Skating Competition
- Future of the ISU

ISU Approach to Integrity Issue

- Secret Judging
- Random Selection of Judges
- Double Trimmed Mean Used to Combine Marks

Secret Judging - Benefits

- Judges less susceptible to external pressure
- *Some* deal making *may* be discouraged

Secret Judging - Weaknesses

- Some judges choose to engage in misconduct without external pressure
- Some federations will still pressure their judges anyway, and some judges will give in
- Judges from smaller federations are the federation

Secret Judging - Weaknesses

- Creates impression problems are being swept under the rug
 - Anomalies hidden from public view
 - Anonymous event review
 - Secret accountability system

Secret Judging - Weaknesses

- Public reaction is overwhelmingly negative
 - Credibility is gone!
 - Reduced event attendance?
 - Reduced TV ratings?
 - Reduced participation?
 - Reduced \$\$\$?

Random Selection of Judges

- Of 14 judges, 9 are randomly selected to compute the results at the World Championships
 - Seven of ten in smaller competitions

Random Selection of Judges

- Provides no practical deterrence against misconduct or bias
- Provides no mathematical benefit to filter out misconduct or bias

Random Selection of Judges

- Judges with their own personal agendas operate without restraint
- Judges of questionable training or ability operate without restraint

Random Selection of Judges

- On a split decision, placements will be determined by a “flip of the coin” about 25% of the time
 - 7 out of 30 places in a championship event
 - 3 out of 12 medals in a competition
 - 1 out of 4 gold medals in a competition
- This has already been confirmed this season

Voo Doo Math

(Double Trimmed Mean)

- Take the nine judges' marks. Drop the two high marks and the two low marks. Average the remaining five marks.

Voo Doo Math

(Double Trimmed Mean)

- Works moderately well to traps gross bias for 1-2 judges working together (or like minded)
- Mediocre to poor at trapping moderate bias for 1-2 judges working together
- Fails to trap bias for more than 2 judges working together

Voo Doo Math

(Double Trimmed Mean)

- High / low judges may be the best judges
- Easy for poor judges to hide in the middle marks
- One block can control the results

Voo Doo Math

(Double Trimmed Mean)

- 1-2 judges can control the results
- Enhances the power of 1 biased judge
- Reduces the numerical confidence in the results by one-third
- Systematically rewards the leader in close events

Voo Doo Math

(Double Trimmed Mean)

Several well established statistical methods all perform 2 - 8 times better than the double trimmed mean in trapping bias for up to 5 judges on a 15 judge panel!

Proposed Point Based System

- User Interface
- Hardware/Software Implementation
- New Structure of Officials
- The Point Model

User Interface

Element	Quality
Element	Quality
Element	Quality
Element	Quality
Element	Quality
Element	Quality
Element	Quality
Element	Quality
Element	Quality
Element	Quality
Element	Quality
Element	Quality
Element	Quality
Element	Quality
Element	Quality
Element	Quality

Real Time Video and Element Replay

+3	+2	+1	Complete	-1	-2	-3
Skating Skills	Transitions	Performance	Choreography	Interpretation		

User Interface

- Works well
- Easy to use
- Does not distract judges from concentrating on the performances

**Least important issue in
evaluating the proposed system!**

Hardware / Software Implementation

- Expensive and complex
- Lacks the duplication required by ISU regulations
- Rigorous testing and validation process does not exist

Hardware / Software Implementation

- A constantly moving target that will be poorly tested for many competitions
- No amount of testing, redundancy or backup systems can guarantee the system will never fail.

RISK

- What level of risk is acceptable to skaters and coaches?
 - ISU owes skaters and coaches an unambiguous, quantitative, technically sound answer
 - Skaters and coaches should have final word on this because skaters are the ones who will pay for any failures

RISK

- Proposed system has potential to require the reskate of an entire program
- Proposed system has potential to introduce an extremely lengthy delay into the midst of a skater's program.

Competition Officials

PROPOSED

- Judges
- Event Referee
- Judges Coordinator
- Technical Controller
- Technical Specialist
(caller/spotter)

CURRENT

- Judges
- Referee
- Referee
- (New)
- (New)

Technical Specialist (Caller)

- Identifies the elements
 - can be overruled by the Technical Controller
- Judges the difficulty of all elements other than jumps, throws and twist lifts
- Events will be pre-judged

Technical Specialist (Caller)

- Can assign a special bonus of 2 points for an innovative element
- One opinion!
 - Judging of levels can move skaters several places
 - Bonus can move skater one place
 - Other judging responsibilities can move skaters several places

Technical Specialist (Caller)

- Requires a perfection of judgement unlikely to be found in a mortal being
- No recourse to correct errors after an event
- One error can skew results several places
 - In current system, chance of misidentification skewing the results is less than 1 in 100,000

Technical Specialist (Caller)

- No formalized/defined training and appointment structure
- No accountability procedure

Who Will Guard the Guards?

- Misconduct
- External pressure
- Deal making
- National bias
- Conflict of interest
- Software integrity

Point Model

- Individual elements receive a base mark value for difficulty, plus quality points and deductions for execution
- Two additional technical merit marks
- Three presentation marks (four in compulsory dance)

Point Model

- Jumps in second half get an extra 10% credit
- Jump Combinations - Sum of base marks plus quality of most difficult jump
- Jump Sequences - 80% of sum of base marks of two most difficult jumps plus quality of most difficult jump

Point Model

- Goal is to rate performances on an absolute point scale, consistent and directly comparable for all eternity.

Is this humanly possible?

Absolute Vs. Relative Assessment

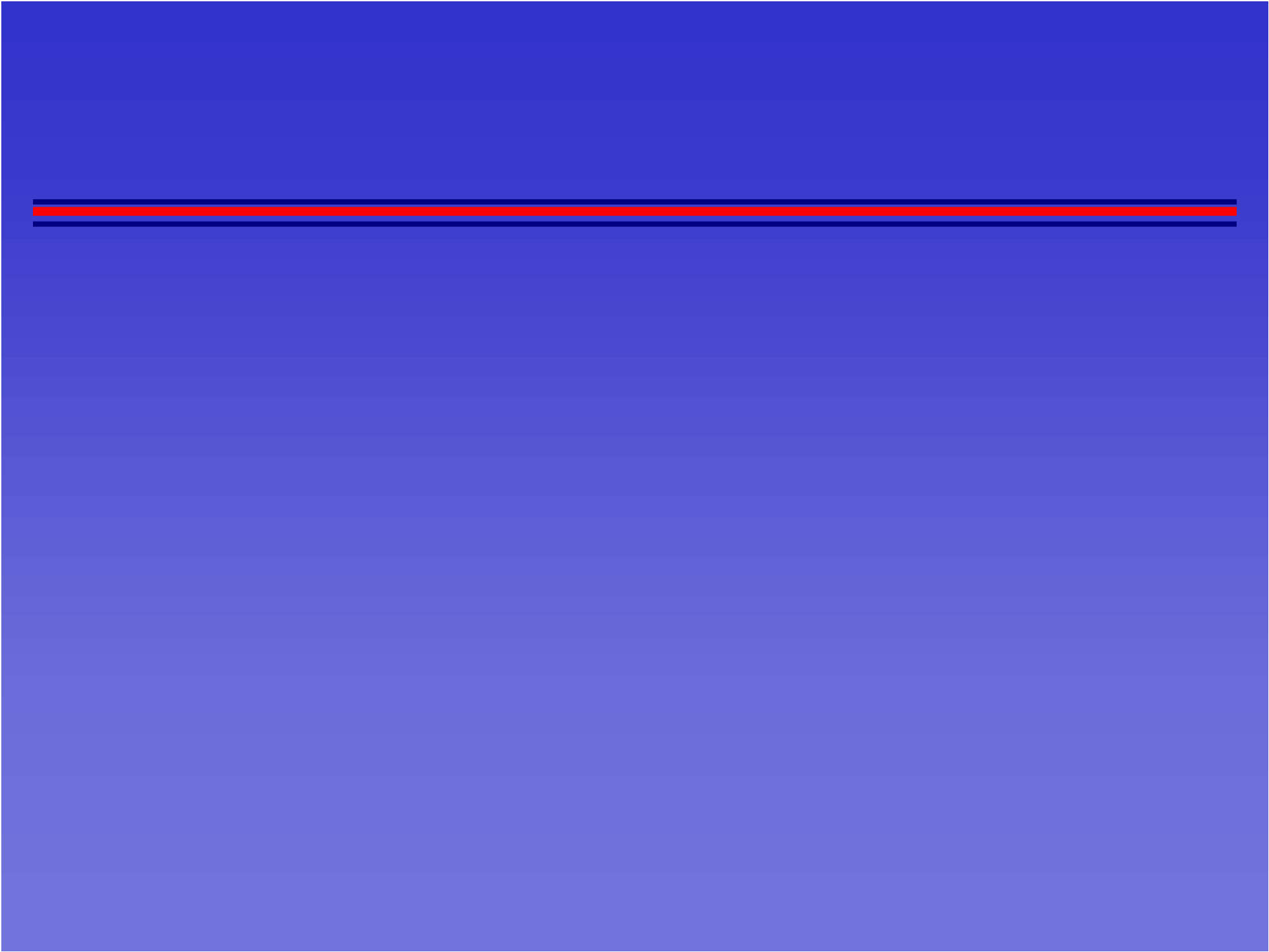
- Currently - Goal of judge is to place the skaters in order of skating skill demonstrated by intercomparison (relative assessment).
- Proposed - Mark on an absolute scale in the absence of any intercomparison or other markers with 99% accuracy (absolute assessment).

Absolute Vs. Relative Assessment

- A long history of studies of human perception show humans are better at relative judgements than absolute judgements
 - Absolute - wrong 15 to 25% of the time
 - Relative - wrong 1 to 2% of the time

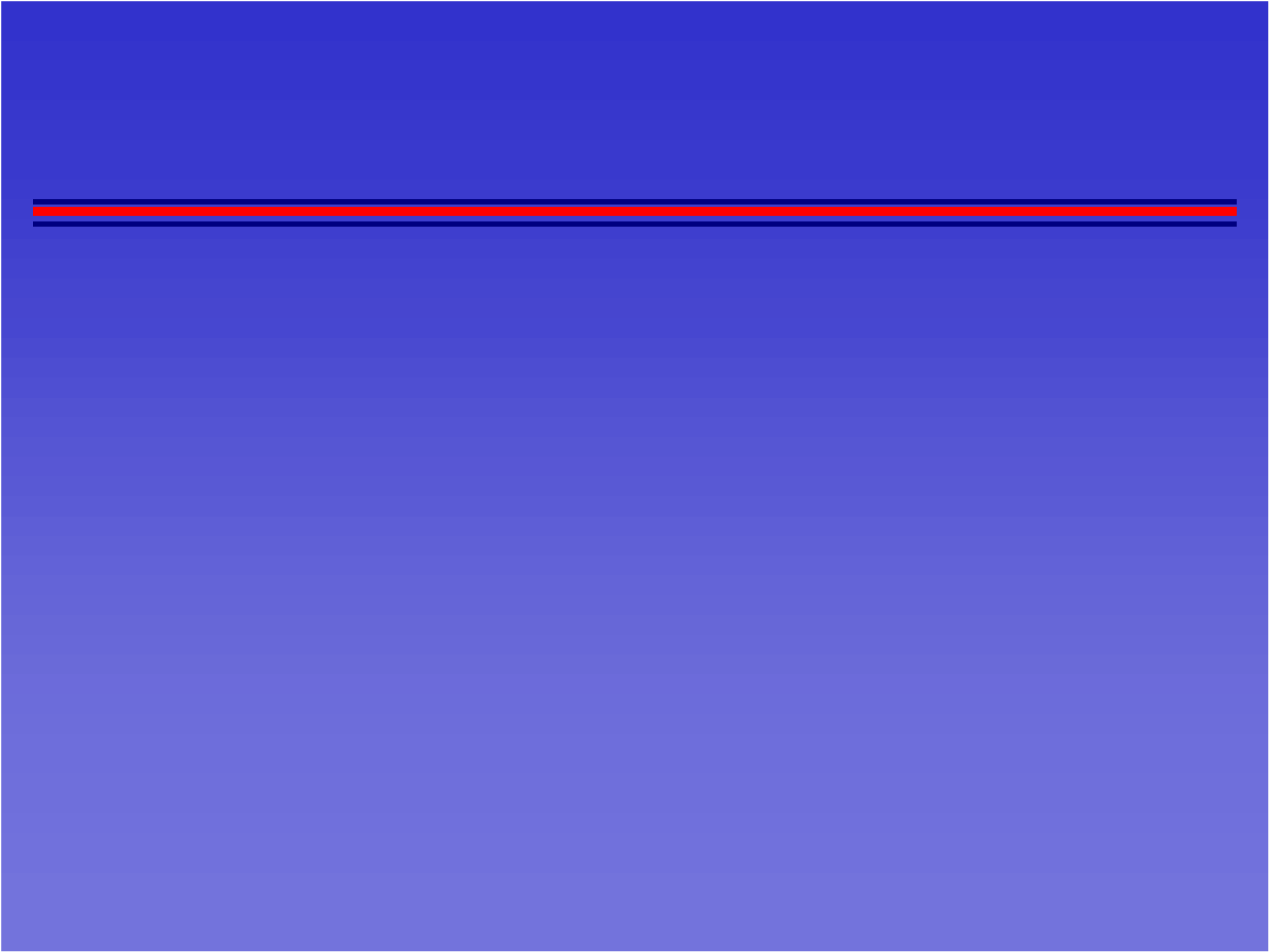
Which





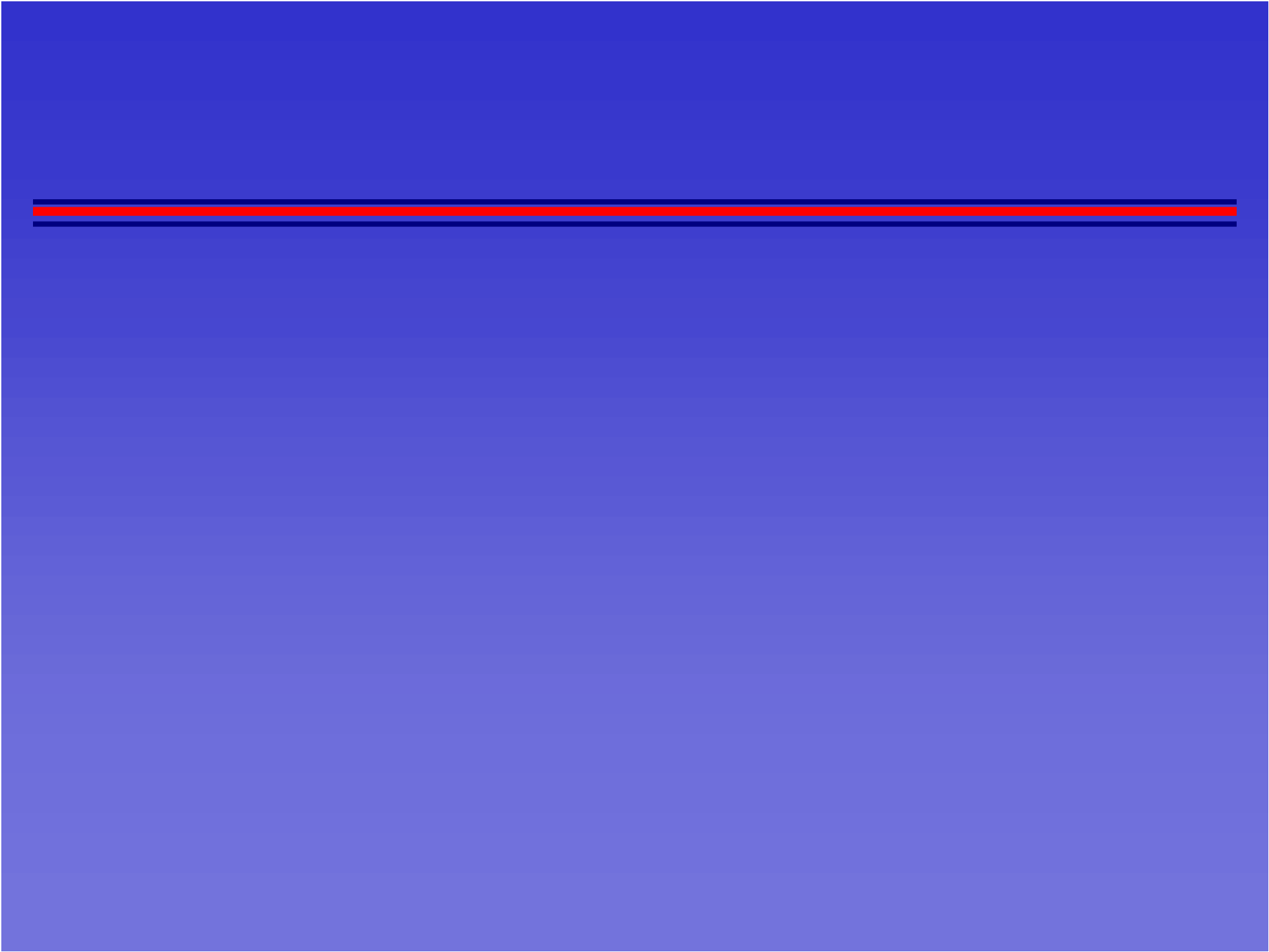
Line





Is





Longest?



Which Line Is Longest?



Absolute Vs. Relative Assessment

- Experiments in human perception show that humans can at best quantify observations into no more than 4 to 7 bins
 - Equivalent to 75 to 85% accuracy
- In skating competition, skaters' point totals differ by 1-2% between places

Example

- Judge a competition for speed by assigning an absolute number for the average speed of each performance

Hard to do!

Example

- Bin the speed into seven bins
 - Excruciatingly Slow
 - Very Slow
 - Slow
 - Moderate
 - Fast
 - Very Fast
 - A Rocket

Much Easier to do.

Binning

- Element execution points are divided into 7 bins (+3, +2, +1, 0, -1, -2, -3)
 - Works OK
- Subjective marks
 - 41 bins each
 - well beyond human capability
 - values for skaters of similar skill are meaningless if actually assigned on an independent absolute scale

Point Values

- Point values define the sport
- Proposed values are not based on current judging/performance standards
- Not based on previous ISU guidelines for element difficulty
- Not based on a bio-mechanical determination of intrinsic physical difficulty

Point Values

- The sport is not defining the point model.
- The point model is arbitrarily redefining the sport.

The Tail Wags the Dog!

Point Values

- Relative values of elements are not consistent
- Relative quality factors are not consistent
- Gaming the system is essential to maximize points and exploit all the quirks!

Point Values

- Jumps, throws and twist lifts have strictly defined values for each trick
- All other elements divided into 3 levels judged by the caller (includes all dance elements)

Some Point Model Omissions

- Walley
- Toe-Walley
- Inside Axel
- One-foot Axel
- Delayed Axel
- Toeless Lutz
- One-foot Salchow
- Split loop, flip, Lutz
- Half loop
- Throw Lutz
- Lateral Twist
- True complexity of spins
- True complexity of jump combinations and sequences

Other Issues and Quirks

- Falls can earn nearly the full base mark of the element (even after a deduction of 3 points)
- A poorly executed triple can earn more points than a well executed double
- Can't do a triple? Eleven double Axels are permitted!

Other Issues and Quirks

- A badly cheated jump (more than $1/4$ turn) counts as a successful jump of one less rotation
- Spins with difficult changes of foot or position get no additional credit
- Definitions of combinations and sequences are extremely complex and confusing. Beware!
- Point values for combinations and sequences leads to absurd situations

Other Issues and Quirks

- Sequences with difficult connecting moves get no additional credit
- Sequences with three (or more) jumps only get credit for two
- Combinations/sequences get less credit than the same jumps executed as solo jumps
- Combinations/sequences get the same credit regardless of jump order

Jump Combination Example

Triple Lutz - Triple Toe Loop

Error	None	3T	3L	Both
Solo	16.6	10.6	10.6	4.6
Comb.	13.6	13.6	7.6	7.6
Seq.	10.88	10.88	6.08	6.08

Subjective Marks

- Technical Merit
 - Skating Skills
 - Transitions (Connecting Moves)
- Presentation
 - Performance / Execution
 - Choreography
 - Interpretation
 - Timing - Compulsory Dance Only

Subjective Marks

- Marked 0 to 10 in steps of 0.25
- Cannot be humanly marked on an absolute scale
- Plenty of flexibility to manipulate the marks

Marks must still be saved!

Apples and Oranges Scoring

- Elements are judged on an open ended scale
- Subjective marks are judged on a fixed scale and marks must be saved

Apples and Oranges Scoring

- Quality assessments are divided into 7 bins. Consistent with limits of human perception.
- Subjective marks are divided into 41 bins. Far beyond human ability, to the point of being meaningless.

Deductions

- Time Violation - 1.0 for every 5 sec.
 - (equivalent to 0.067 on 6.0 scale)
- Music Violation - 1.0
 - (equivalent to 0.067 on 6.0 scale)
- Illegal element - 2.0 each
 - (equivalent to 0.15 on 6.0 scale)
- Costume Violation - 1.0
 - (equivalent to 0.067 on 6.0 scale)

Well Balanced Program

Senior Singles

- Current

- Unlimited Solo Jumps (restrictions on repeated triple and quad jumps)
- 1-3 Jump Combinations or Sequences
- 4 Spins Minimum
- 2 Step Sequences

- Proposed (Maximums)

- 8 Jump Elements (7 for ladies)
 - Up to two combinations or sequences permitted, but not required.
 - Two jumps in one combination. Up to three jumps in the second.
- 4 Spin Elements
- 2 Step Sequences

Well Balanced Program

Senior Pairs

- Current

- 3-5 Lifts (1 Twist lift min., 2 max.)
- 1-2 Throws
- 1-2 Solo Jumps
- 1 Jump Combination or Sequence
- 1 Solo Spin
- 1-2 Pair Spins
- 1-2 Death Spirals
- 1 Step Sequence
- 1 Spiral Sequence

- Proposed (Maximums)

- 3 Lifts
- 1 Twist Lift
- 2 Throws
- 1 Solo Jump
- 1 Jump Combination or Sequence
- 1 Solo Spin Combination
- 1 Pair Spin Combination
- 2 Death Spirals
- 1 Step Sequence
- 1 Spiral Sequence

Singles Point Distribution (Contemporary Senior Men)

- Jumps 42%
- Spins 8%
- Footwork 6%
- Basic Skating Skills and
Connecting Moves 18%
- Presentation 26%

Pairs Point Distribution (Contemporary Senior Pairs)

- Lifts 20%
- Throws and Jumps 16%
- Spins 6%
- Death Spirals 7%
- Footwork 7%
- Basic Skating Skills and
Connecting Moves 18%
- Presentation 26%

Combination of Points

- **Program Points = Element Points + Multiplier * Subjective points**
- **Total Points = Short Program Points + Long Program Points**
- **Subjective points Multiplier**
 - **Men: 1.0 in short program, 2.0 in long program**
 - **Ladies & Pairs: 0.8 in short program, 1.6 in long program**

Combination of Segments (Singles)

- SP and LP
 - SP will become 33% of total for men, 34% for ladies
 - LP will become 67% of total for men, 66% for ladies
- QR and SP and LP
 - QR will become 21% of total for men and ladies
 - SP will become 26% of total for men, 27% for ladies
 - LP will become 53% of total for men, 52% for ladies

Combination of Segments

- A skater can place first in the long program and second in the short program and lose the competition
- With a qualifying round, a skater can place first in the short program, first in the long program, and still lose the competition

Combination of Segments

	QR	SP	LP
You	160 (2nd)	112(1st)	172(1st)
Me	165(1st)	110(2nd)	170(2nd)

	Current	Proposed
You	TFP 1.2 - You win	444 points - You lose
Me	TFP 1.8	445 points

Open Questions

- Definitions of the element levels?
- Deductions for falls when base mark plus quality points are less than 3 points? For spins? For footwork?
- Bonus point definition / criteria?
- How will *significant* new or rare elements be accommodated?

Open Questions

- Details for handling errors in jump combinations and sequences?
- How will errors that interrupt combination spins be handled? When can they be retried?
- How will errors and differences in elements be handled for side-by-side jumps and spins in pairs?
- Do pops of half rotation count as attempts? Can they be retried?

Strategy

- **Leave NO points on the table!**
 - Attempt all elements permitted
 - Attempt all 11 jumps
 - Attempt the 11 most difficult jumps you can get all the way around on
 - Attempt something clever for your bonus points
- Take great care in how one improvises during a performance

Strategy

- Attempt as many of the most difficult jumps as possible just after the half way point in the program
- Take care in choosing combinations and sequences
- Limit sequences to two multi-revolution jumps

Strategy

- Concentrate on the three most important types of elements
 - Jumps
 - Jumps
 - Jumps
- Limit spins and footwork to the minimum effort needed to get to +3 of Level 3

Strategy

- Be extremely careful counting spin rotations
- Performing complete programs in practice is essential for new programs
- Optimize the program for maximum points, not best presentation
- Game the system thoroughly