ARRow's Impossibility TheoRem

✅ Arrow's Impossibility Theorem (Plain Language)

> There is no voting system that can convert individual rankings of **three or more options** into a fair group ranking that satisfies all of the following:

0. Transitivity: the group ranking is logically consistent: $A > B$ & $B > C \Rightarrow A > C$.
1. **Unrestricted Domain**: The system works for **any combination** of individual preferences.
2. **Unanimity (Pareto Efficiency)**: If **everyone** prefers option A over B, the group must prefer A over B.
3. **Independence of Irrelevant Alternatives (IIA)**: The group's ranking between A and B should depend **only** on how people rank A vs. B — not on their views about C.
4. **Non-Dictatorship**: No single voter always decides the outcome.

> If a voting system satisfies the first fouR properties, then it must be a **dictatorship** — one person's preferences always determine the group's outcome.

COR

Majority vote with 2 candidates is the only "fully faiR" election (that satisfies these pRopeRties).

Examples

Voting methods typically give up IIA. Though IIA is the most benign of these voting conditions, we shall see that not having it can produce bad Results.

Ex. Plurality

Rule The candidate with the most top choice votes wins (even if he doesn't have a majority).

3 voters (1,2,3), 4 candidates (A, B, C, D), and consider 2 social preference profiles,

$$P_1 = \begin{pmatrix} R_1 & R_2 & R_3 \\ \hline A & A & C \\ B & B & D \\ C & C & B \\ D & D & A \end{pmatrix} \quad \text{and} \quad P_2 = \begin{pmatrix} R_1 & R_2 & R_3 \\ \hline B & B & C \\ A & A & D \\ C & C & B \\ D & D & A \end{pmatrix}.$$

For P_1:

1 votes for A
2 votes for A
3 votes for C

↓

A wins, C 2nd,
B & D tied for 3rd

For P_2:

1 votes for B
2 votes for B
3 votes for C

↓

B wins, C 2nd,
A & D tied for 3rd

Voters 1 and 2 changing their preferences of A and B led to a change in the Rankings between A and C, violating IIA.

No one changed her Relative Rankings of A versus C! Voter 1 and 2 still think A is better than C, as before, yet the rise in ranking of an irrelevant alternative (B) changed the way the social Rule (plurality) ranks A relative to C!

Why is that dependence on irrelevant alternatives a bad result? Suppose that A, B, C and D are four different candidates for a job. But we don't know whether they would really be willing to accept the job. So we'll vote, using plurality voting, and then offer the job to the candidates in order : so we'll give it to the candidate with the second–most votes if we are turned down with the candidate with the most votes.

What if candidate B isn't really interested in the job. How people rank this irrelevant candidate, who won't even accept the job, has changed the ranking of the two candidates who would accept the job, A and C. And that does not seem a very attractive property for a choice rule.

Ex. Ranked Choice Voting

Rule

1. **First-Choice Votes Counted**: All voters' first-choice selections are tallied.
2. **Majority Check**: If a candidate receives more than 50% of these votes, they win.
3. **Elimination and Redistribution**:
   - If no candidate has a majority, the candidate with the **fewest first-choice votes** is eliminated.
   - Voters who selected the eliminated candidate as their first choice will have their votes transferred to their next preferred candidate who is still in the race.
4. **Repeat Rounds**: Steps 2 and 3 are repeated until a candidate achieves a majority and is declared the winner. calrcv.org +1 rcvresources.org +4

$$P_1 = \begin{pmatrix} \underline{5} & \underline{4} & \underline{3} & \underline{1} \\ A & C & B & B \\ B & B & C & A \\ C & A & A & C \end{pmatrix}$$

$$P_2 = \left(P_1, \text{ but eliminate } C \right)$$

Rd 1: A 5, B 4, C 4
no majority
eliminate both B and C
A wins, B & C tied for 2nd

Rd 1: A 5, B 8
B > A

- Every voter's ranking of **A vs. B stayed the same**
- But removing the **irrelevant loser C** flipped the outcome from **C wins** → **B wins**

That is a **clear violation of Independence of Irrelevant Alternatives (IIA)**.

> In Ranked Choice Voting, the group's choice between A and B was **affected by the presence of C**, even though **no one changed their view of A vs. B**.

Ex. Borda Count

__Rule__ With M candidates, in each profile the 1st ranked candidate gets M points, 2nd ranked candidate gets M−1 pts, ..., last gets 1 point.

$$P_1 = \begin{pmatrix} \underline{45\%} & \underline{55\%} \\ A & B \\ C & A \\ B & C \end{pmatrix} \qquad P_2 = \begin{pmatrix} \underline{45\%} & \underline{55\%} \\ A & B \\ B & C \\ C & A \end{pmatrix}$$

A gets $3 \cdot 45 + 2 \cdot 55 = 245$ pts
B gets $3 \cdot 55 + 1 \cdot 45 = 210$ pts
C gets $2 \cdot 45 + 1 \cdot 55 = 145$ pts
Social Ranking $A > B > C$

A: 210 pts
B: 245 pts
C: 110 pts
Social Ranking is $B > A > C$

Only by moving C in each ballot, the social ranking of A and B switched, violating IIA.

__Implication:__ Our Super Smash Bros tournament scoring system violates IIA. Justin's performance could impact the social rank order of Ryan and Nick.

Ex. Pairwise Majority Vote (with 3+ options)

Rule Compare each pair of candidates via majority voting one at a time

$$P_1 = \begin{pmatrix} R_1 & R_2 & R_3 \\ \hline A & B & C \\ B & C & A \\ C & A & B \end{pmatrix}$$

A > B by 2-1
B > C by 2-1
C > A by 2-1

"condorcet cycle" A > B > C > A > ...
transitivity is violated

> **Violating IIA isn't necessarily fatal.**
> In fact, **every practical voting system that avoids dictatorship and works with 3+ candidates violates IIA**. Most people accept that tradeoff.

🚨 **Why People Like IIA**

- **Prevents spoiler effects**: A third candidate (e.g., Nader, Perot) shouldn't flip the election.
- **Encourages stability**: If no one changes their opinion about A vs. B, the result between A and B shouldn't change.
- **Eliminates "irrelevant" manipulation**: You can't game the outcome of A vs. B by tweaking your ranking of C.

So on paper, it looks like a **must-have property**.

Mathematical setup for Elections

There are N voters, or individuals, denoted $[N] = \{1, 2, \dots, N\}$.
There are ≥ 3 candidates, denoted $\mathcal{C} = \{a, b, c, \dots\}$.
Each voter i has a ranking, or preference, of the candidates defined by a WEAK ORDERING, which is basically a mathematical abstraction of "$>$, $\geq$, and $=$" on the set of candidates that is a REFLEXIVE, COMPLETE, and TRANSITIVE binary relation R_i on $\mathcal{C}$. For example, $a > b > c$ is a ranking where $\mathcal{C} = \{a, b, c\}$.
Denote the set of all possible rankings, or weak orderings, by $\mathcal{R}$.
A social preference profile, or tuple of ballots, is an N-tuple containing each voter's ranking, denoted $(R_1, \dots, R_N) \in \mathcal{P} = \mathcal{R}^n$, where $\mathcal{P}$ is the set of social profiles.
An election rule or social preference function given by $E: \mathcal{P} \to \mathcal{R}$ assigns a social ranking, or weak ordering, to every possible social profile in $\mathcal{P}$.
For example, Majority Wins, Rank Choice Voting, Borda Count, etc.
A social preference function E satisfies

- Unanimity if $\forall x, y \in \mathcal{C}$, forall social preference profiles $P = (R_1, \dots, R_N) \in \mathcal{P}$ such that each voter i ranks x above y $(x >_i y \quad \forall i \in [N])$, we must have $x > y$ in the social ranking $E(P)$.

- <u>Independence of Irrelevant Alternatives</u> (IIA)

if $\forall x, y \in \mathcal{C}$, for all pairs of profiles $P = (R_1, \dots, R_N)$ and $P' = (R_1', \dots, R_N')$ in $\mathcal{P}$ such that each voter i ranks x and y the same in both P and P',

$x \geq y$ in the social ranking $E(P) \iff x \geq y$ in $E(P')$.

In other words, the social ranking between x and y won't change if you don't alter any of the relative rankings between x and y; you can freely move other candidate around.

- For a given social preference function E, an individual i is <u>decisive</u> for some $x \in \mathcal{C}$ over some $y \neq x$, $y \in \mathcal{C}$, if $x \geq_i y \implies x > y$ in the social ranking $E(P)$.

Voter i is a <u>dictator</u> if he is decisive for every x over ever $y \neq x$.

<u>Arrow's Impossibility Theorem</u> If there are at least 3 candidates and if a social preference function satisfies Unanimity and IIA, then some individual is a dictator.

Proof of Arrow's Theorem
(from Mark Fey's 2014 Paper)

Step 1 Identify voter i_*

We will find a voter i_* who we will show is a dictator!

Fix the following two profiles

$$P_{0a} = \begin{pmatrix} R_1 & \cdots & R_n \\ \hline a & \cdots & a \\ b & \cdots & b \\ \vdots & \vdots & \vdots \end{pmatrix} \quad \text{and} \quad P_{0b} = \begin{pmatrix} R_1 & \cdots & R_n \\ \hline b & \cdots & b \\ a & \cdots & a \\ \vdots & \vdots & \vdots \end{pmatrix}$$

where the dotted ranges represent the other alternative candidates in fixed but arbitrary locations.

Unanimity $\Longrightarrow$ $a > b$ in profile P_{0a}
and $b > a$ in profile P_{0b}

Now transform P_{0a} into P_{0b} by switching a and b one voter at a time, from left to right, holding all other alternative candidates fixed.

Let individual i_* be the voter for which the social preference changes from $a > b$ to something else $(b \geq a)$ for the first time.

Concretely, for the two profiles

$$P_{1a} = \begin{pmatrix} R_1 \cdots R_{i_*-1} & R_{i_*} & R_{i_*+1} \cdots R_n \\ \hline b & a & a \\ a & b & b \\ \vdots & \vdots & \vdots \end{pmatrix} \quad \text{and} \quad P_{1b} = \begin{pmatrix} R_1 \cdots R_{i_*-1} & R_{i_*} & R_{i_*+1} \cdots R_n \\ \hline b & b & a \\ a & a & b \\ \vdots & \vdots & \vdots \end{pmatrix},$$

the social preference in P_{1a} is $a > b$ and in P_{1b} is $b > a$.

Step 2 $\forall c \neq a, b$, voter i_* is decisive for b over c.

Let $c \neq a, b$. The voters in profile P_2 rank a and b (ignoring all the other candidates) the same as in profile P_{1a},

$$P_2 = \begin{pmatrix} \frac{R_1 \cdots R_{i_*-1}}{\begin{matrix} b \\ c \\ a \\ \vdots \end{matrix}} & \frac{R_{i_*}}{\begin{matrix} a \\ b \\ c \\ \vdots \end{matrix}} & \frac{R_{i_*+1} \cdots R_n}{\begin{matrix} a \\ b \\ c \\ \vdots \end{matrix}} \end{pmatrix}$$

so by IIA the social preference in P_2 is $a > b$.

By unanimity, $b > c$ in P_2, so by transitivity $a > c$ in P_2.

Now, consider the following set of profiles, where the notation b/c means that the candidates b and c can be ranked arbitrarily in the indicated spot by profiles in the set.

$$\mathcal{S}_2 = \left\{ \begin{pmatrix} \frac{R_1 \cdots R_{i_*-1}}{\begin{matrix} b/c \\ a \\ \vdots \end{matrix}} & \frac{R_{i_*}}{\begin{matrix} b \\ a \\ c \\ \vdots \end{matrix}} & \frac{R_{i_*+1} \cdots R_n}{\begin{matrix} a \\ b/c \\ \vdots \end{matrix}} \end{pmatrix} \right\} \subseteq \mathcal{P}$$

$\forall P \in \mathcal{S}_2$, voters in P rank a and b the same (ignoring all other candidates) as profile P_{1b}, and $b > a$ in P_{1b}, so by IIA $b > a$ $\forall P \in \mathcal{S}_2$.

Similarly, $\forall P \in \mathcal{S}_2$, voters in P rank a and c the same as profile P_2, so $a > c$ $\forall P \in \mathcal{S}_2$.

Hence, by transitivity, $b > c$ $\forall P \in \mathcal{S}_2$.

Now, let $P \in \mathcal{P}$ be any profile where $b >_{i_*} c$.

Voters in P must rank b and c the same as some profile P' in $\mathcal{S}_2$ (ignoring all other candidates). Since $b > c$ $\forall P \in \mathcal{S}_2$, $b > c$ in P', so by IIA we must have $b > c$ in P. Since P was arbitrary, $\forall P \in \mathcal{P}$ $b >_{i_*} c \Rightarrow b > c$, so i_* is decisive for b over c.

$\square$

Step 3 $\forall c \neq a,b$, voter i_* is decisive for a over c.

Let $c \neq a,b$ and consider the following set of profiles,

$$\mathcal{S}_3 = \left\{ \begin{pmatrix} \frac{R_1 \cdots R_{i_*-1}}{\begin{matrix} a/c \\ b \\ \vdots \end{matrix}} & \frac{R_{i_*}}{\begin{matrix} a \\ b \\ c \\ \vdots \end{matrix}} & \frac{R_{i_*+1} \cdots R_n}{\begin{matrix} a/c \\ b \\ \vdots \end{matrix}} \end{pmatrix} \right\} \subseteq \mathcal{P}.$$

By Step 2, i_* is decisive for b over c, so $b > c$ $\forall P \in \mathcal{S}_3$.

By unanimity, $a > b$ $\forall P \in \mathcal{S}_3$.

Hence, by transitivity, $a > c$ $\forall P \in \mathcal{S}_3$.

Now, let $P \in \mathcal{P}$ be any profile where $a >_{i_*} c$.

Voters in P must rank a and c the same (ignoring all other candidates) as some profile P' in $\mathcal{S}_3$.

Since $a > c$ $\forall P \in \mathcal{S}_3$, $a > c$ in P', so by IIA we must have $a > c$ in P. Since P was arbitrary,

$\forall P \in \mathcal{P}$ $a >_{i_*} c \implies a > c$, so i_* is decisive for a over c. $\square$

Step 4 $\forall c \neq a, b$, voter i_* is decisive for c over a.

Let $c \neq a, b$. The voters in profile P_4 rank a and b the same (ignoring all other candidates) as in profile P_{1a}, and

$$P_4 = \begin{pmatrix} R_1 \cdots R_{i_*-1} & R_{i_*} & R_{i_*+1} \cdots R_n \\ \hline b & c & c \\ c & a & a \\ a & b & b \\ \vdots & \vdots & \vdots \end{pmatrix}$$

$a > b$ in P_{1a}, so by IIA $a > b$ in P_4.
By unanimity, $c > a$ in P_4. Hence, by transitivity, $c > b$ in P_4.

Now, consider the following set of profiles,

$$\mathcal{S}_4 = \left\{ \begin{pmatrix} R_1 \cdots R_{i_*-1} & R_{i_*} & R_{i_*+1} \cdots R_n \\ \hline b & c & a/c \\ a/c & b & b \\ \vdots & a & \vdots \\ & \vdots & \end{pmatrix} \right\} \subseteq \mathcal{P}.$$

$\forall P \in \mathcal{S}_4$, voters in P rank a and b the same (ignoring all other candidates) as profile P_{1b}, and $b > a$ in P_{1b}, so by IIA $b > a$ $\forall P \in \mathcal{S}_4$.
Similarly, $\forall P \in \mathcal{S}_4$, voters in P rank b and c the same as profile P_4, so $c > b$ $\forall P \in \mathcal{S}_4$.
Hence, by transitivity, $c > a$ $\forall P \in \mathcal{S}_4$.

Now, once again, let $P \in \mathcal{P}$ be any profile where $c >_{i_*} a$.
Voters in P must rank a and c the same as some profile P' in $\mathcal{S}_4$ (ignoring all other candidates). Since $c > a$ $\forall P \in \mathcal{S}_4$, $c > a$ in P', so by IIA we must have $c > a$ in P. Since P was arbitrary,
$\forall P \in \mathcal{P}$ $c >_{i_*} a \Rightarrow c > a$, so i_* is decisive for c over a.

$\square$

Step 5 $\forall c \neq a, b$, voter i_* is decisive for c over b.

Let $c \neq a, b$ and consider the following set of profiles,

$$\mathcal{S}_5 = \left\{ \begin{pmatrix} R_1 \cdots R_{i_*-1} & R_{i_*} & R_{i_*+1} \cdots R_n \\ \hline a & c & a \\ b/c & a & b/c \\ \vdots & b & \vdots \\ & \vdots & \end{pmatrix} \right\} \subseteq \mathcal{P}.$$

By Step 4, i_* is decisive for c over a, so $c > a$ $\forall P \in \mathcal{S}_5$.

By unanimity, $a > b$ $\forall P \in \mathcal{S}_5$.

Hence, by transitivity, $c > b$ $\forall P \in \mathcal{S}_5$.

Now, let $P \in \mathcal{P}$ be any profile where $c >_{i_*} b$.

Voters in P must rank b and c the same (ignoring all other candidates) as some profile P' in $\mathcal{S}_5$.

Since $c > b$ $\forall P \in \mathcal{S}_5$, $c > b$ in P', so by IIA we must have $c > b$ in P. Since P was arbitrary,

$\forall P \in \mathcal{P}$ $c >_{i_*} b \implies c > b$, so i_* is decisive for c over b.

$\square$

Step 6 voter i_* is decisive for a over b & b over a.

Consider the following set of profiles,

$$\mathcal{S}_6 = \left\{ \begin{pmatrix} R_1 \cdots R_{i_*-1} & R_{i_*} & R_{i_*+1} \cdots R_n \\ \hline a/b & a & a/b \\ \vdots & c & \vdots \\ \vdots & b & \vdots \\ \vdots & \vdots & \vdots \end{pmatrix} \right\} \subseteq \mathcal{P}.$$

By step 3, i_* is decisive for a over c, and by step 5, i_* is decisive for c over b.

Hence, $\forall P \in \mathcal{S}_6$, $a > c$ and $c > b$, so by transitivity $a > b$ $\forall P \in \mathcal{S}_6$.

Now, let $P \in \mathcal{P}$ be any profile where $a >_{i_*} b$.

Voters in P must rank a and b the same (ignoring all other candidates) as some profile P' in $\mathcal{S}_6$.

Since $a > b$ $\forall P \in \mathcal{S}_6$, $a > b$ in P', so by IIA we must have $a > b$ in P. Since P was arbitrary,

$\forall P \in \mathcal{P}$ $a >_{i_*} b \implies a > b$, so i_* is decisive for a over b.

The argument for b over a is the same, swapping the position of a and b. $\square$

Step 7 Individual i_* is a dictator.

We must show i_* is decisive for every x over every $y \neq x$. We have shown i_* is decisive
for b over y $\forall y \neq b$ (Step 2 and 6),
for x over b $\forall x \neq b$ (Step 5 and 6),
for a over y $\forall y \neq a$ (Step 3 and 6),
for x over a $\forall x \neq a$ (Step 4 and 6).

The only case remaining is $x \neq a, b$ and $y \neq a, b$.
Let $x \neq a, b$ and $y \neq a, b$. Consider the following set of profiles,

$$\mathcal{S}_7 = \left\{ \begin{pmatrix} \underline{R_1 \cdots R_{i_*-1}} & \underline{R_{i_*}} & \underline{R_{i_*+1} \cdots R_n} \\ x/y & x & x/y \\ a/b & a & a/b \\ \vdots & b & \\ & y & \\ & \vdots & \end{pmatrix} \right\} \subseteq \mathcal{P}.$$

Since i_* is decisive for x over a, a over b, and b over y, $\forall P \in \mathcal{S}_7$ we have $x > a > b > y$, so by transitivity $x > y$ $\forall P \in \mathcal{S}_7$.

Now, let $P \in \mathcal{P}$ be any profile where $x >_{i_*} y$.
Voters in P must rank x and y the same (ignoring all other candidates) as some profile P' in $\mathcal{S}_7$.
Since $x > y$ $\forall P \in \mathcal{S}_7$, $x > y$ in P', so by IIA we must have $x > y$ in P. Since P was arbitrary,
$\forall P \in \mathcal{P}$ $x >_{i_*} y \implies x > y$, so i_* is decisive for x over y. □