

# Intuitive Preference Aggregation: Tests of Independence and Consistency

Todd Davies  
Symbolic Systems Program  
Stanford University

Raja Shah  
Google

# Overview

- 1) Can we construct preference profiles for which the aggregation judgments of large majorities of neutral survey-takers robustly violate Arrow's criteria of IIA and collective rationality?
- 2) If such violations occur, are they the result of robust and widespread agreement among survey-takers about what preference aggregation procedure (or family thereof) should be applied to profiles generally?
- 3) What principles characterize neutral survey-takers' revealed social preferences?
- 4) What is the proper role of experiments in solving the practical problem of selecting a social choice rule for a given population?

DEFINITION: A social welfare function  $f$  is an **Arrow social welfare function** if

- the domain  $\mathbf{D}$  of  $f$  contains all possible preference profiles  $\mathbf{R}^N$  (**universal domain**),  
and
- the range  $\mathbf{R}$  of  $f$  is the set of weakly ordered social preference relations  $R_N$  (**collective rationality**).

THEOREM: **Arrow's Impossibility Theorem**

(Arrow, 1951/1963). There can be no Arrow social welfare function  $f$  satisfying all of the following:

- for all social outcomes  $x$ , and  $y$  in  $X$ , if a preference profile  $\mathbf{R}^N$  obeys  $x P_i y$  for each individual  $i$  in  $N$ , then  $f(\mathbf{R}^N)$  yields  $x R_N y$  (**weak Pareto efficiency**);
- for all preference profiles  $\mathbf{R}^N$  and  $\mathbf{R}^{N'}$  in  $\mathbf{D}$ , and all social outcomes  $x$  and  $y$  in  $X$ , if  $\mathbf{R}^N$  and  $\mathbf{R}^{N'}$  obey  $x R_i y$  iff  $x R_i' y$  for all individuals  $i$  in  $N$ , then  $f(\mathbf{R}^N)$  and  $f(\mathbf{R}^{N'})$  yield  $x R_N y$  iff  $x R_{N'} y$  (**independence of irrelevant alternatives**);

and

- there is no individual  $d$  in  $N$  such that for all preference profiles  $\mathbf{R}^N$  in  $\mathbf{D}$ , and all social outcomes  $x$  and  $y$  in  $X$ , if  $x P_d y$  then  $f(\mathbf{R}^N)$  yields  $x P_N y$  (**non-dictatorship**).

## Arrow on IIA

“The essential argument in favor of this principle is its direct appeal to intuition.” (1952)

“Stricter than desirable” but has practical benefits, necessary to keep the lid on the need to gather limitless information on unavailable options (1967)

Dead candidate example:

If an election were held and one of the candidates then died: "Surely the social choice should be made by taking each of the individual's preference lists, blotting out the dead candidate's name, and considering only the orderings of the remaining names..." (1951/1963).

DEFINITION: Given a set  $X$  of social outcomes, a preference profile  $\mathbf{R}^N$ , and a social choice rule  $C: x R_N^* y$  (meaning there is a **revealed social preference** for  $x$  in relation to  $y$ ) iff there is some environment  $S \subseteq X$ , such that  $x$  and  $y$  are in  $S$  and  $x$  is in  $C(S, \mathbf{R}^N)$ .

DEFINITION: **Weak axiom of revealed social preference.** Given a set  $X$  of social outcomes, a preference profile  $\mathbf{R}^N$ , and a social choice rule  $C$ , a social preference relation  $R_N$  satisfies **inter-menu consistency**<sup>1</sup> (**IMC**) iff it is a revealed social preference relation under  $C$  for all pairs of social outcomes in  $X$ , and the following condition holds: For all environments  $S \subseteq X$ , if  $x$  and  $y$  are in  $S$  and  $x$  is in  $C(S, \mathbf{R}^N)$ , then for all environments  $S' \subseteq X$  such that  $x$  and  $y$  are in  $S'$ , if  $y$  is in  $C(S', \mathbf{R}^N)$ , then  $x$  is in  $C(S', \mathbf{R}^N)$ .

THEOREM: If  $R_N^*$  is a revealed social preference relation associated with a set  $X$  of social outcomes, a preference profile  $\mathbf{R}^N$ , and a social choice rule  $C$ , and  $R_N^*$  violates inter-menu consistency, then it violates collective rationality.

---

<sup>1</sup> This term is from Sen (1993).

DEFINITION: A social choice rule  $C$ , defined on a set  $X$  of social outcomes and a domain  $\mathbf{D}$  of preference profiles satisfies **independence of unavailable alternatives (IUA)** iff for all environments  $S \subseteq X$ , and all preference profiles  $\mathbf{R}^N$  and  $\mathbf{R}^{N'}$  in  $\mathbf{D}$ , the following condition holds: If for all  $x, y$  in  $S$ ,  $\mathbf{R}^N$  and  $\mathbf{R}^{N'}$  obey  $x R_i y$  iff  $x R_i' y$  for all individuals  $i$  in  $N$ , then  $C(S, \mathbf{R}^N) = C(S, \mathbf{R}^{N'})$ .

**Figure 1 – Profile triplets (Rank-form presentation)**

**In each triplet, picking X in top profile and Y in middle profile violates IMC. Picking Y in middle profile and X in bottom profile violates IIA.**

5-2		<b>Voter 1</b>	<b>Voter 2</b>	<b>Voter 3</b>	<b>Voter 4</b>	<b>Voter 5</b>
	<b>Ranked Alternatives</b>	X	X	X	Y	Y
		Y	Y	Y	X	X
5-3		<b>Voter 1</b>	<b>Voter 2</b>	<b>Voter 3</b>	<b>Voter 4</b>	<b>Voter 5</b>
	<b>Ranked Alternatives</b>	Z	X	X	Y	Y
		X	Y	Y	Z	Z
		Y	Z	Z	X	X
5-3'		<b>Voter 1</b>	<b>Voter 2</b>	<b>Voter 3</b>	<b>Voter 4</b>	<b>Voter 5</b>
	<b>Ranked Alternatives</b>	X	X	X	Z	Y
		Z	Z	Y	Y	X
		Y	Y	Z	X	Z

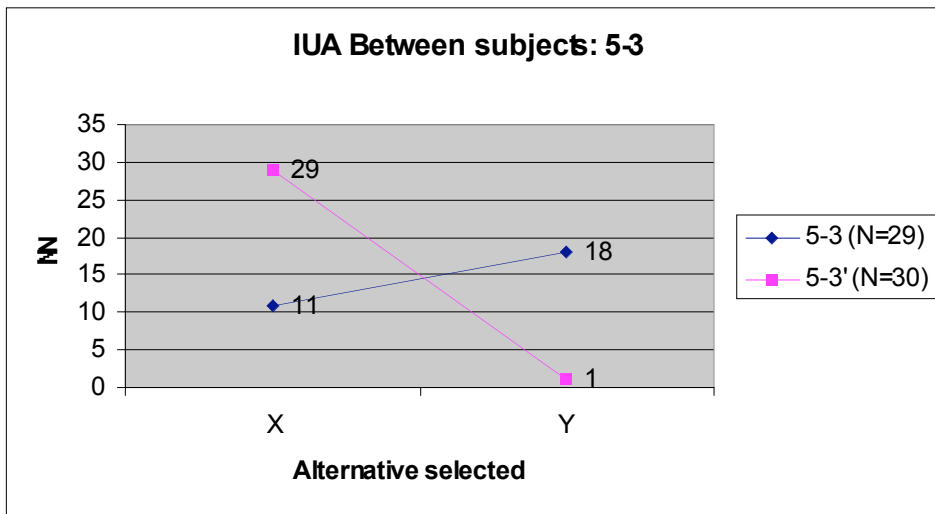
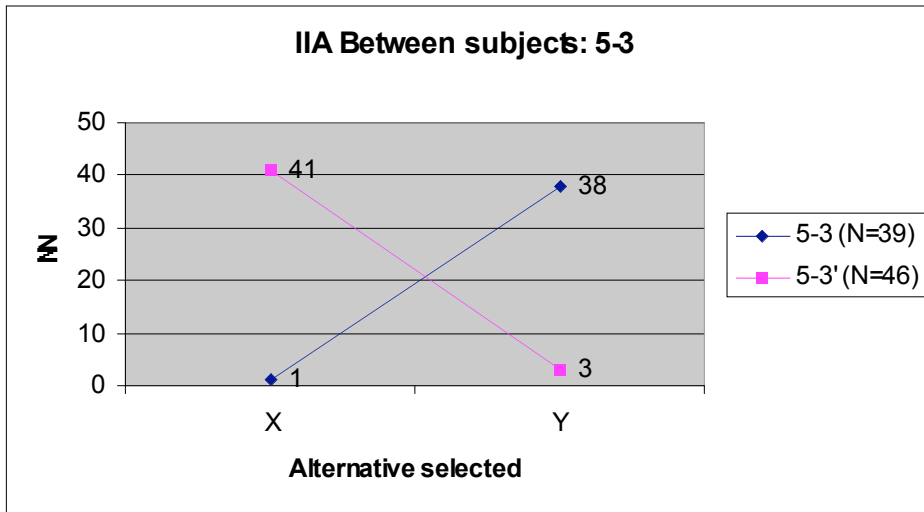
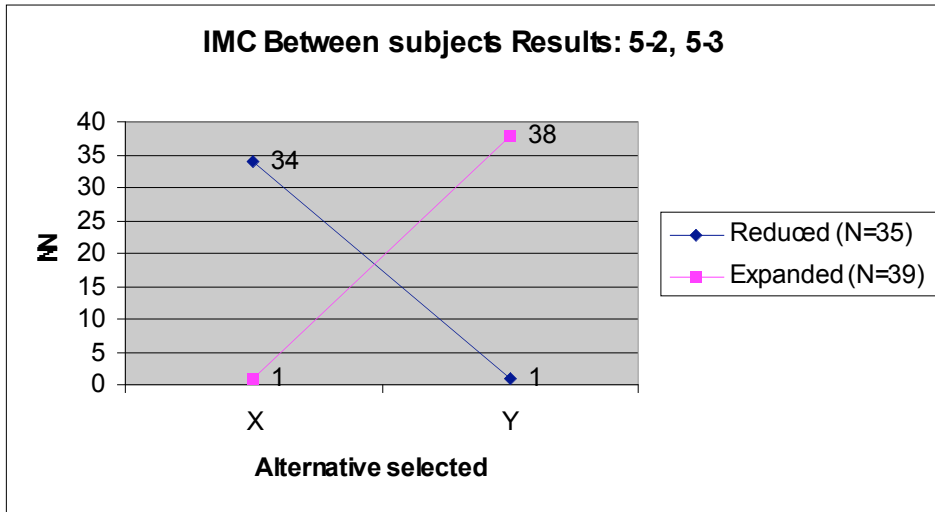
3-2		<b>Voter 1</b>	<b>Voter 2</b>	<b>Voter 3</b>
	<b>Ranked Alternatives</b>	X	X	Y
		Y	Y	X
3-4		<b>Voter 1</b>	<b>Voter 2</b>	<b>Voter 3</b>
	<b>Ranked Alternatives</b>	X	X	Y
		Y	Y	Z
		Z	W	W
		W	Z	X
3-4'		<b>Voter 1</b>	<b>Voter 2</b>	<b>Voter 3</b>
	<b>Ranked Alternatives</b>	X	X	Z
		Z	W	Y
		W	Y	X
		Y	Z	W

4-3		<b>Voter 1</b>	<b>Voter 2</b>	<b>Voter 3</b>	<b>Voter 4</b>
	<b>Ranked Alternatives</b>	X	X	Y	Y
		Y	Z	X	X
		Z	Y	Z	Z
4-6		<b>Voter 1</b>	<b>Voter 2</b>	<b>Voter 3</b>	<b>Voter 4</b>
	<b>Ranked Alternatives</b>	U	X	Y	Y
		X	Z	W	V
		Y	Y	V	U
		Z	U	U	W
		V	V	X	X
		W	W	Z	Z
4-6'		<b>Voter 1</b>	<b>Voter 2</b>	<b>Voter 3</b>	<b>Voter 4</b>
	<b>Ranked Alternatives</b>	U	X	Y	Y
		X	Z	X	X
		Z	U	W	V
		V	V	V	U
		W	W	U	W
		Y	Y	Z	Z

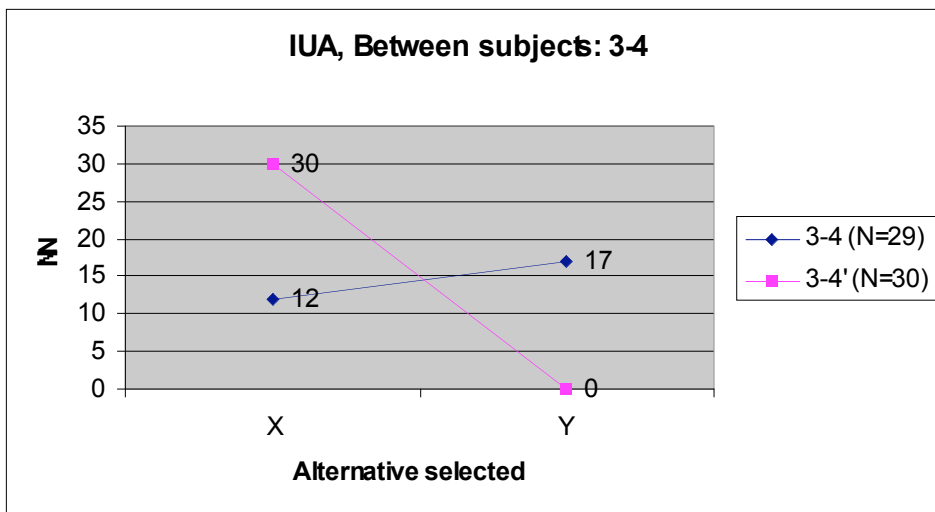
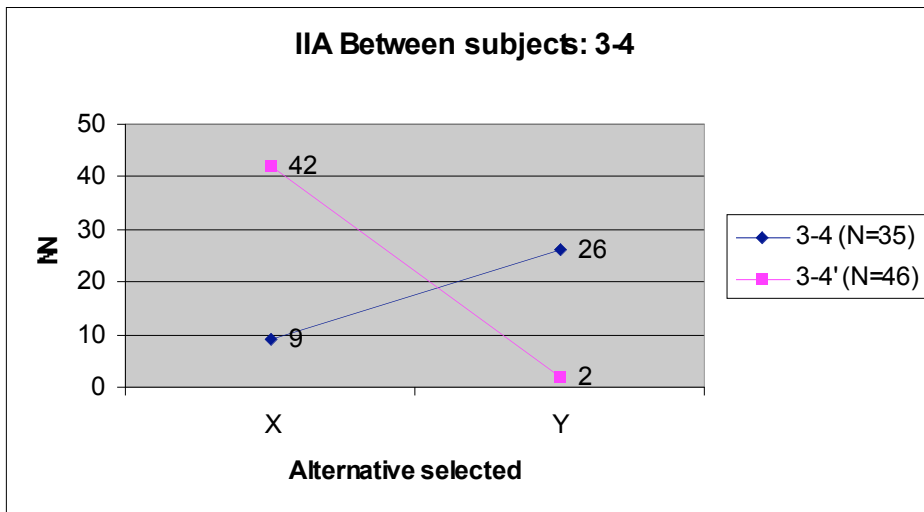
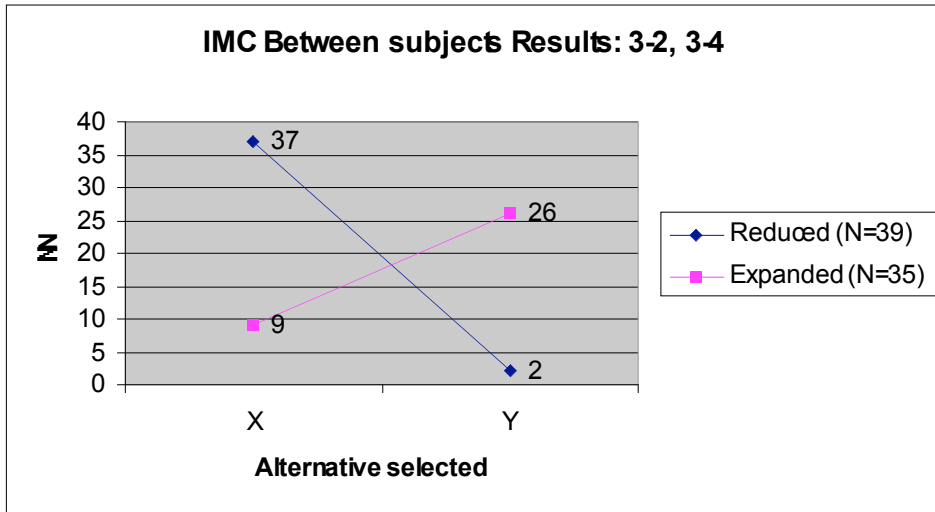
Question 1: Can we construct preference profiles for which the aggregation judgments of large majorities of neutral survey-takers robustly violate Arrow's criteria of IIA and collective rationality?



Graphs 1.1-1.3: IMC, IIA, and IUA Between subjects results for 5-voter profile pair



Graphs 2.1-2.3: IMC, IIA, and IUA Between subjects results for 3-voter profile pair



Graphs 3.1-3.3: IMC, IIA, and IUA Between subjects results for 4-voter profile pair

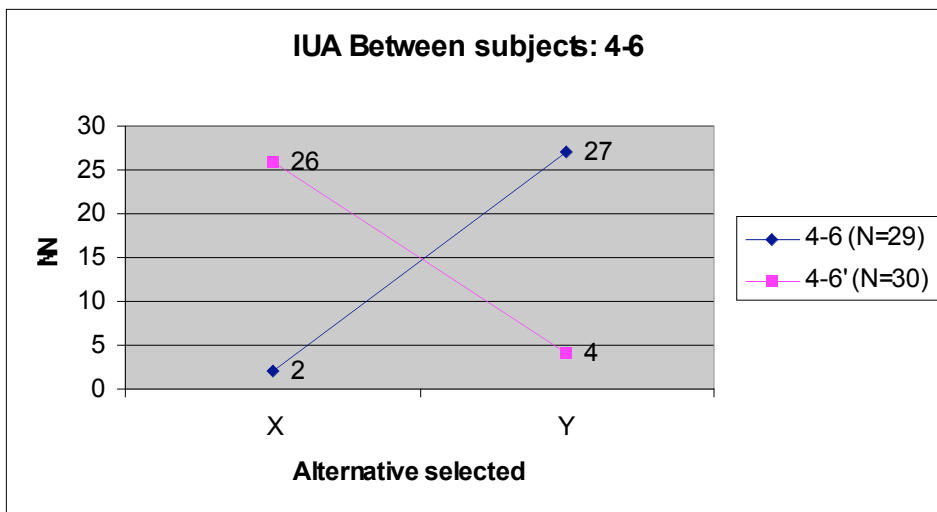
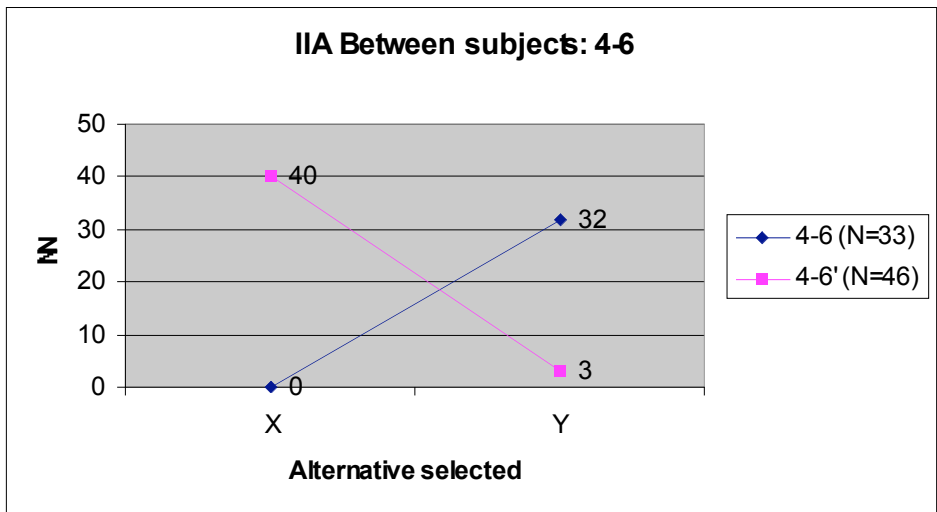
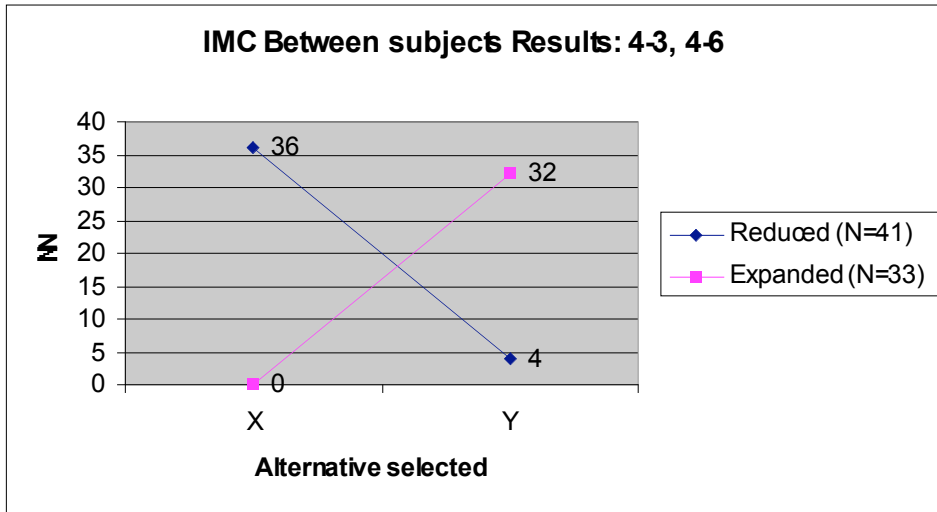


Figure 2 - Ordering

	<b>51</b>	<b>49</b>
	<b>%</b>	<b>%</b>
<b>Ranked Alternatives</b>	L	P
	O	L
	M	O
	P	N
	N	Q
	Q	M

Out of 59 subjects, 55 ranked P ahead of M, while only 4 ranked M ahead of P. Assuming that they would have ranked M ahead of P if only shown the voting blocs' preferences over these two alternatives, we can say that the subject group, as a whole, is violating the IMC criterion.

Figure 3. Within-subjects tests for 5 voters x 3 alternatives pairings, rank format presentation

<b>Test</b>	<b>Violators</b>	<b>Nonviolators</b>
IMC	84%	16%
IIA	92%	8%
IUA	64%	36%

Figure 4 – Profiles (Pairwise presentation)

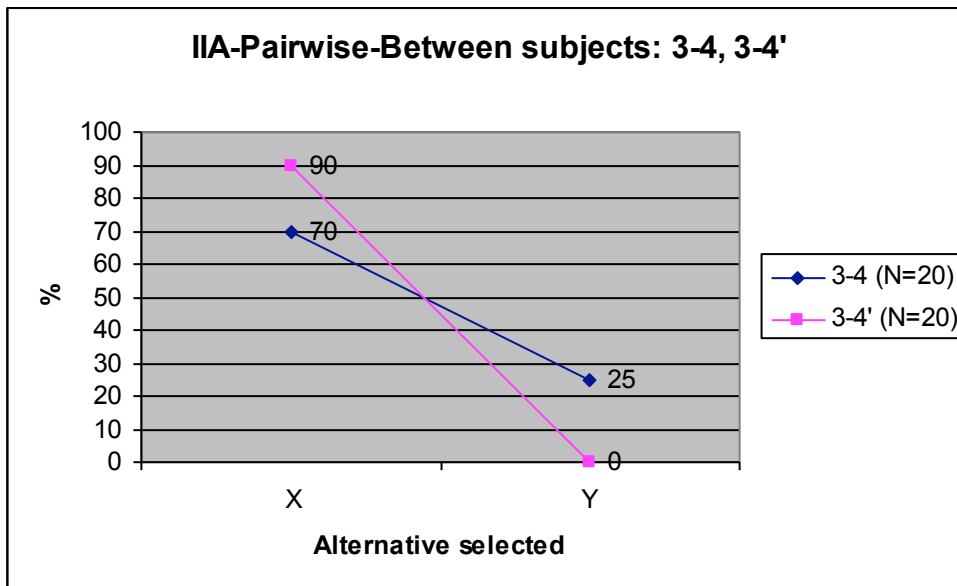
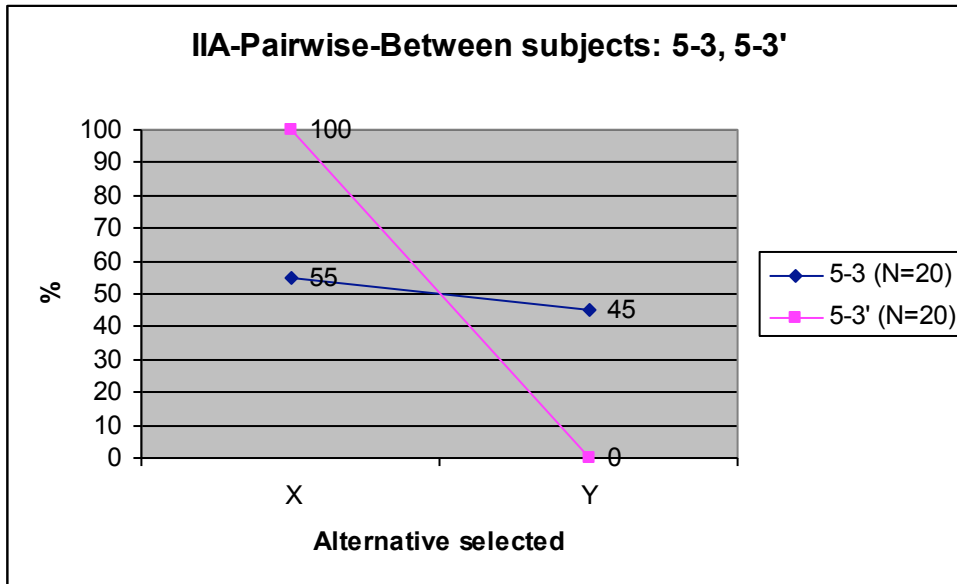
5-3 (Pairwise)

	<b>X vs. Y</b>	<b>Y vs. Z</b>	<b>Z vs. X</b>
<b>Voter A</b>	X	Z	Z
<b>Voter B</b>	X	Y	X
<b>Voter C</b>	X	Y	X
<b>Voter D</b>	Y	Y	Z
<b>Voter E</b>	Y	Y	Z

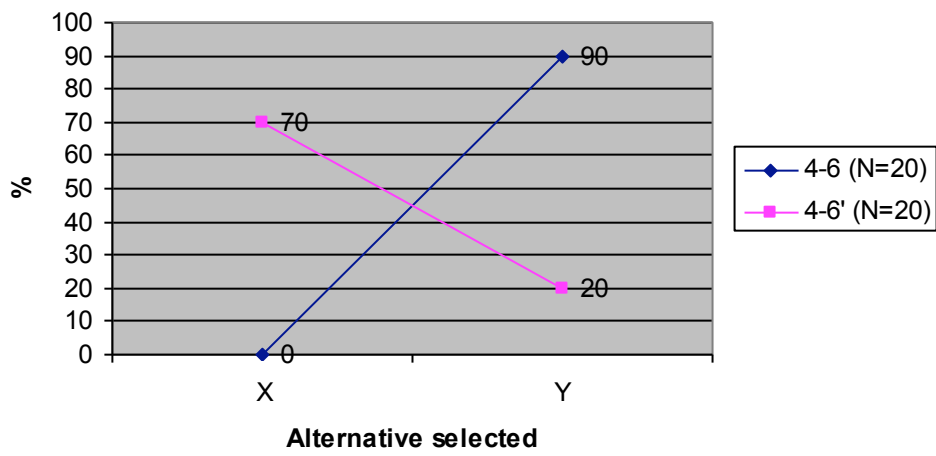
5-3' (Pairwise)

	<b>X vs. Y</b>	<b>Y vs. Z</b>	<b>Z vs. X</b>
<b>Voter F</b>	X	Z	X
<b>Voter G</b>	X	Z	X
<b>Voter H</b>	X	Y	X
<b>Voter I</b>	Y	Z	Z
<b>Voter J</b>	Y	Y	X

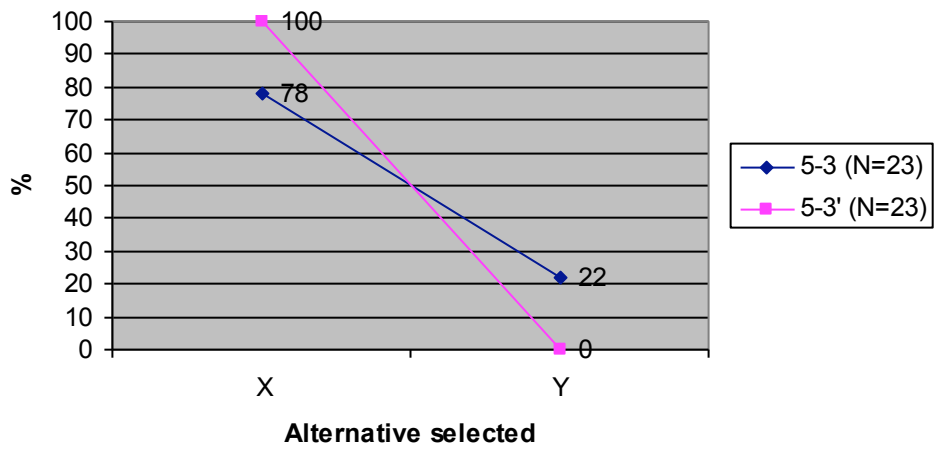
Figures 5a-5f. Results for pairwise presentation, between subjects:



**IIA-Pairwise-Between subjects: 4-6, 4-6'**



**IUA-Pairwise-Between subjects: 5-3, 5-3'**



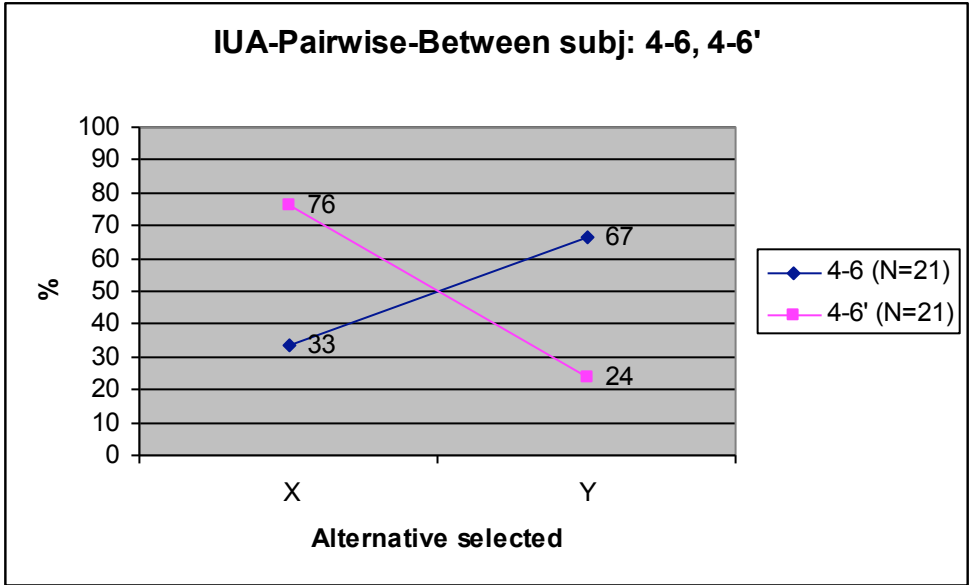
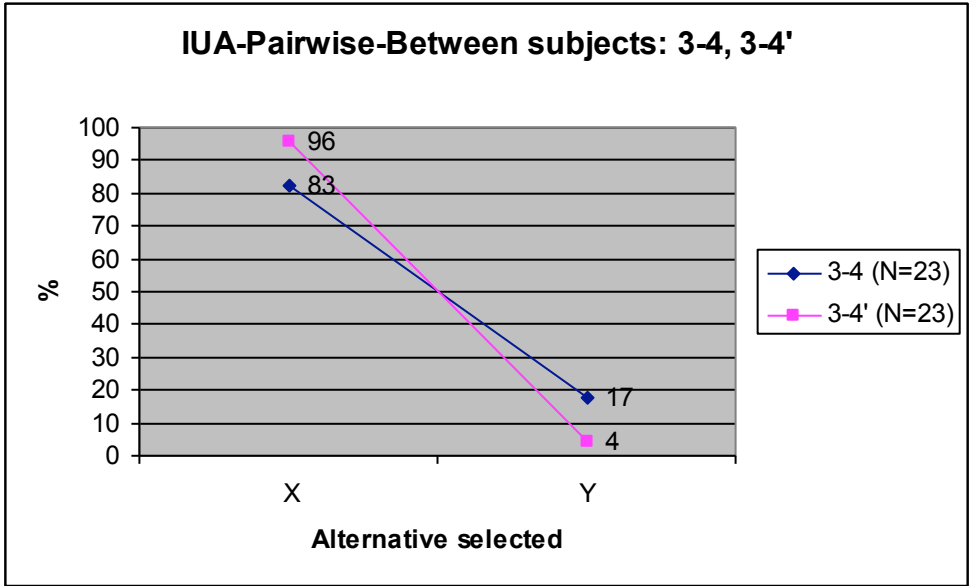


Figure 6. Pairwise Presentation Results, Within Subjects, 5 voters x 3 alternatives only

Test	Violators	Nonviolators
IIA	75%	25%
IUA	35%	65%



## Question 1 Summary

Question 1: Can we construct preference profiles for which the aggregation judgments of large majorities of neutral survey-takers robustly violate Arrow's criteria of IIA and collective rationality?

*Answer:* Yes. Large violations of IMC and IIA for all three profile sets in rank format. But pairwise format yields violation for 4x6 profile set only. Effect appears robust to presentation mode only for sufficiently large Borda reversals.

Figure 7. Size of Borda reversals for each profile set in IIA/IUA studies

<b>Profile set</b>	<b>Borda reversal</b>
5 voters, 3 alternatives	5:6 versus 7:4
3 voters, 4 alternatives	6:7 versus 7:3
4 voters, 6 alternatives	11:16 versus 17:10

Question 2: If such violations occur, are they the result of robust and widespread agreement among survey-takers about what preference aggregation procedure (or family thereof) should be applied to profiles generally?

*Answer:* Individual differences and presentation emerge as the size of the Borda reversal diminishes.

- Pairwise versus rank format
- Within versus between subjects
- Order effect within subjects (which profile presented first)

Question 3: What principles characterize neutral survey-takers' revealed social preferences?

*Some answers:*

- 1) Most subjects violate inter-menu consistency (IMC), pairwise independence (IIA), and independence of unavailable alternatives (IUA), regardless of presentation mode, when the Borda difference is large enough (i.e. there is a clear Borda winner in each profile).
- 2) When the Borda reversal is smaller,
  - a. Subjects adhere to IUA significantly more than IIA or IMC.
  - b. Pairwise presentation leads subjects toward aggregation that respects IIA and IUA (probably also IMC).
  - c. Within-subjects presentation leads subjects away from IMC and IIA (but not IUA)
  - d. Mode of presentation (rank/pairwise format, within/between subjects, and profile order) has a large biasing effect on survey takers.
  - e. Order of presentation in within-subjects comparisons biases survey-takers toward or away from IUA, depending on whether a format-cued procedure favors violation or consistency.

- 3) Survey-takers appear to apply a variety of principles to preference aggregation, none of which is fully determinative when they conflict:
- a.* PRINCIPLE 1: Independence of irrelevant voters (IIV). Social choices should be based only on the preferences of individuals affected by the choice.
  - b.* PRINCIPLE 2: Interpersonal comparisons of utility (ICU). Social choices should take into account, as much as possible, information about the relative strengths of preference that individuals in the choice group have for one alternative over another.
  - c.* PRINCIPLE 3: Inter-menu independence (IMI). Social choices should take into account only the alternatives that are available, and not utilize information about how individuals rank these available alternatives among other sets which include unavailable outcomes.
  - d.* PRINCIPLE 4: Independence of unavailable alternatives (IUA). Social choices should be the same whenever the set of available outcomes does not change, even if individuals change their preferences involving unavailable outcomes.
  - e.* OTHER PRINCIPLES: Pairwise majority (Condorcet defeat), plurality, elimination...

Question 4: What is the proper role of experiments in solving the practical problem of selecting a social choice rule for a given population?

*Two suggested answers:*

- 1) No principle of social choice should be regarded as indispensable if there are situations in which the fully-informed opinions of a large majority of intelligent people think that the best social decision is one that entails violating the principle.
- 2) Surveys that systematically assess neutral subjects' aggregation across a large number of abstract preference profiles could be the basis for inducing a weighted, multi-criterion hybrid social choice procedure that would optimally characterize people's intuitions about fair social choice.