

Crowd Wisdom and Deliberative Voting: Does Deliberation Improve Estimation Accuracy?

Jan Lorenz¹, Heiko Rauhut², Bernhard Kittel¹

¹ Center for Social Science Methodology, Carl von Ossietzky Universität Oldenburg

² Chair of Sociology, in particular Modeling and Simulation, ETH Zurich

ECPR General Conference, Reykjavík

August 25, 2011

Problem: Use of expert committees in politics

- ▶ Expert committees (e.g. on economic growth or tax estimation) are semi-political:
 - ▶ They shall find the best possible predictions (scientific deliberation, driven by truth-finding)
 - ▶ They shall or want to deliver a common opinion (political deliberation or bargaining)
- ▶ Example: **Intergovernmental Panel on Climate Change**
 - ▶ Shall predict sea level 2050 and 2100. Correctness is important:
 - ▶ Build dikes without need would be a waste of money...
 - ▶ High sea level and no dikes cause costly disasters...
 - ▶ Prevention of CO₂-emissions would be a waste of money if sea level rises anyway...
 - ▶ Described as “bazaar” where future sea level is “negotiated”.
 - ▶ Discusses its internal report and decision structures.

Research question on: The wisdom of the Crowd and Deliberation

- ▶ In 1907 Galton found that the average estimate of a weight judging competition outperformed every individual estimate (called the wisdom-of-crowd effect = large diversity of opinions but surprisingly close to correct average opinion)
- ▶ So, expert committees might use the effect!
- ▶ The effect is statistical and builds on independent estimates.
- ▶ Social influence can undermine the effect¹
- ▶ Can expert committees use it if they deliberate?
- ▶ Or can they even improve the outcome when they deliberate?
- ▶ If yes under what decision rules?

¹Lorenz, Rauhut, Schweitzer, Helbing; PNAS 2011

Experimental Design

Operationalize expert committee's task in laboratory experiment:

- ▶ Prediction → **Factual Questions**. Example:
 - ▶ What percentage of Saudi Arabian oil is used in Saudi Arabia?
- ▶ Optimum at true value → **incentivize** close to **correct** answers

Deviation	Payment
$\leq 1.5\%$	4.00 EUR
$\leq 3\%$	2.00 EUR
$\leq 6\%$	1.00 EUR
$\leq 12\%$	0.50 EUR
larger	0.00 EUR

- ▶ Strong incentive to come closer to the truth!

Experimental Design

Operationalize expert committee's task in laboratory experiment:

- ▶ Prediction → **Factual Questions**. Example:
 - ▶ What percentage of Saudi Arabian oil is used in Saudi Arabia?
 - ▶ **19%**
- ▶ Optimum at true value → **incentivize** close to **correct** answers

Deviation	Payment
$\leq 1.5\%$	4.00 EUR
$\leq 3\%$	2.00 EUR
$\leq 6\%$	1.00 EUR
$\leq 12\%$	0.50 EUR
larger	0.00 EUR

- ▶ Strong incentive to come closer to the truth!

Experimental Design

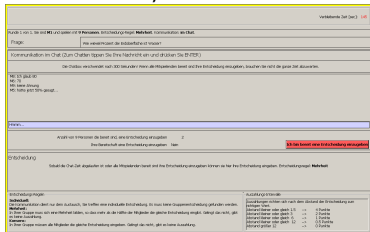
- ▶ Invited for 4 **sessions** of 24 participants each
- ▶ In each session 12 **rounds** with 12 questions (from 18)
- ▶ **Decision rules**
 1. **individual rule** ← no group decision, all estimates in report
 2. **Majority rule** ← common estimate desired, even disagreed
 3. **Consensus rule** ← agreement desired/necessary
 - ▶ No group decision ← committee fails to deliver report

Experimental Design

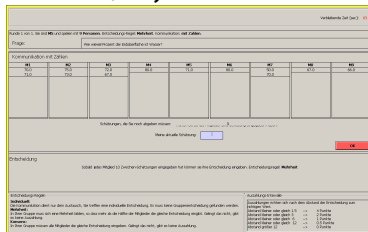
► All 12 rounds had 4 stages

1. Subjects enter private individual **initial estimate**, and **certainty**: 1 (very uncertain) to 6 (very certain)
2. Subjects randomly **grouped** to 3, 3, 9, and 9. Rule info.
3. Communication, deliberation and decision. 2 interfaces:

all-to-all **chat**, free text



numbers, "My current estimate"



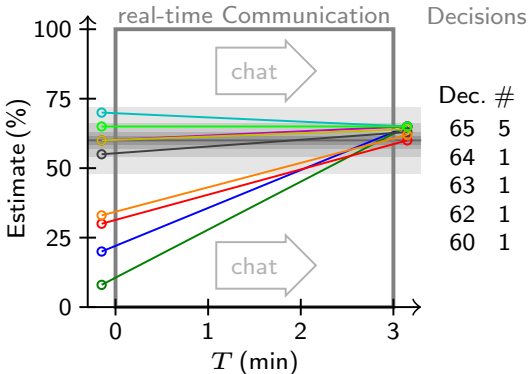
4. Info if group decision reached. True value only at very end!

Examples – Individual

ID 2 Session 1

Share of people in Asia (Q1a)

Decision Rule: Individual

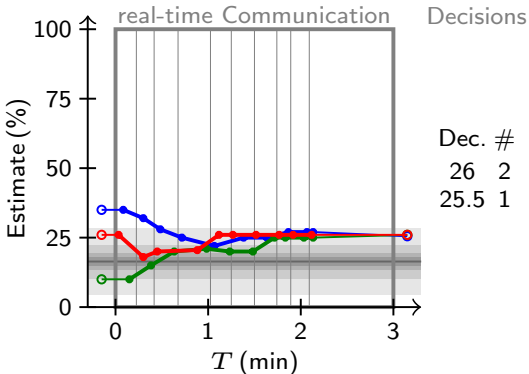


Examples – Individual

ID 167 Session 4

Votes USA at World Bank (Q6)

Decision Rule: Individual

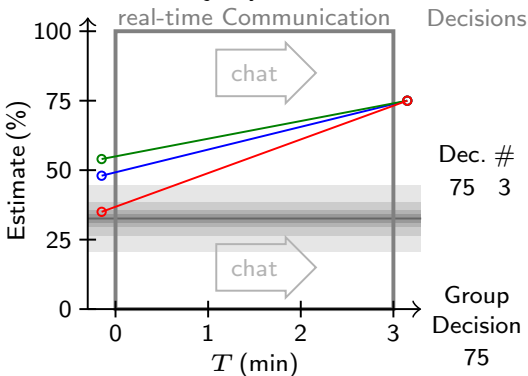


Examples – Majority

ID 68 Session 2

Turnout EU parliament Great Britain 1984 (Q5a)

Decision Rule: Majority

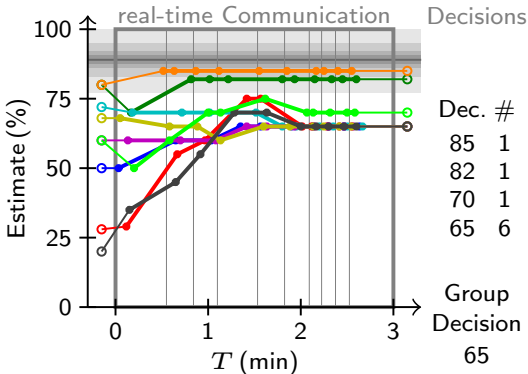


Examples – Majority

ID 122 Session 3

Water content onion (Q7b)

Decision Rule: Majority

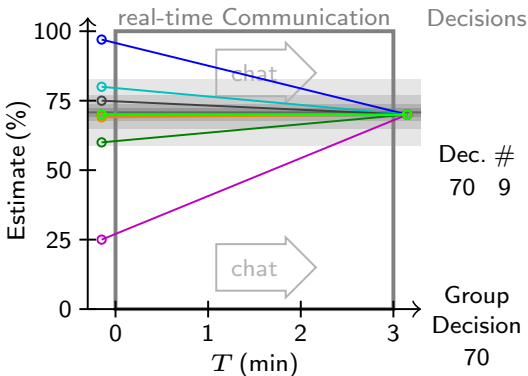


Examples – Consensus

ID 150 Session 4

Water on earth surface (Q2)

Decision Rule: Consensus

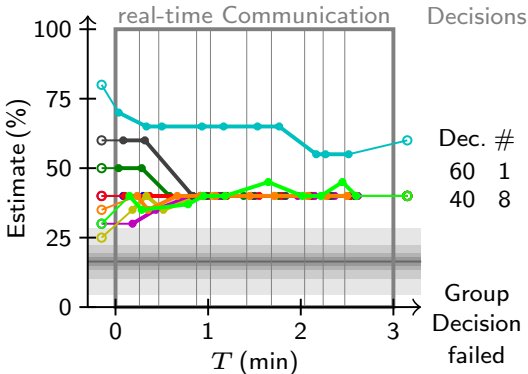


Examples – Consensus

ID 21 Session 1

Votes USA at World Bank (Q6)

Decision Rule: Consensus



Results – Is there a wisdom-of-crowds effect?

	question	truth	mean	median	IQR	med. pay	pay med.
1a	Share of people in Asia	60.00	41.31	40.00	25.00	0.00	0.00
1b	Share of people in Americas	13.50	23.65	20.00	12.00	0.50	0.50
2	Water on earth surface	70.80	67.89	70.00	10.00	1.00	4.00
3a	Mass iron in earth crust	4.70	20.30	11.50	20.00	0.75	0.50
3b	Mass aluminium in earth crust	8.10	14.12	10.00	13.25	1.00	2.00
4	Workers in agriculture	40.00	34.52	31.00	26.25	0.00	0.50
5a	Turnout EU parliament UK 1984	32.60	41.51	35.00	32.00	0.00	2.00
5b	Turnout EU parliament FR 1984	56.80	41.58	40.00	29.00	0.00	0.00
6	Votes USA at World Bank	16.40	33.16	30.00	18.50	0.00	0.00
7a	Water content tomato	93.00	77.11	80.00	15.00	0.00	0.00
7b	Water content onion	89.00	53.61	50.00	45.25	0.00	0.00
8	Use of own oil Saudi Arabia	19.00	13.29	10.00	15.00	0.50	0.50
9a	Turnout Bundestag DE 2009	70.90	49.74	53.50	29.52	0.00	0.00
9b	Turnout Bundestag DE 1998	82.20	59.94	61.50	17.50	0.00	0.00
10	Land area with no agriculture	82.00	52.53	50.00	26.25	0.00	0.00
11a	Fat content reindeer's milk	16.90	26.81	23.50	32.00	0.25	0.50
11b	Fat content mother's milk	4.00	29.96	22.50	39.25	0.00	0.00
12	US presidents from Republicans	40.90	53.08	60.00	20.75	0.00	0.00

- ▶ Aggregation of wisdom of crowd: **Median** better than mean
- ▶ WoC effect: Payment of median better than median payment

Results – Payment improvement after deliberation?

	All	Numbers	Chat	3	9
All	0.3**	0.15	0.46**	0.18	0.43**
Individual	0.54**	0.55.	0.53*	0.33.	0.75*
Majority	-0.05	-0.41.	0.3	-0.08	-0.03
Consensus	0.43*	0.31	0.55**	0.3	0.56*

- ▶ There is a **significant average improvement** in (potential) payment from the median initial estimate in groups to the median/group decision.
- ▶ This is not explained by the WoC-effect.
- ▶ Majority rule does not improve on average (especially with number communication).
- ▶ Consensus and individual rule benefit a lot from larger groups.

Results – Certainty improvement after deliberation?

	All	Numbers	Chat	3	9
All	0.3***	0.3***	0.31**	0.36***	0.24**
Individual	0.39***	0.39***	0.4***	0.4***	0.39***
Majority	0.51***	0.51***	0.5**	0.6***	0.41**
Consensus	0.01	-0.01	0.03	0.09	-0.07

- ▶ Average confidence in groups rises under individual and majority rule from initial estimate to (group) decision.
- ▶ It stays constant with consensus rule.
- ▶ Increase in confidence comes with increase in group accuracy under individual rule.
- ▶ Increase in confidence comes without increase in group accuracy under majority rule.

Conclusion

- ▶ It is interesting that groups can improve the group accuracy beyond the statistical wisdom-of-crowds effect.
- ▶ Free text communication seems to better improve group accuracy than communication by numbers.
- ▶ Groups of 9 perform much better on average than groups of 3.
- ▶ Majority rule as a mechanism in expert committees seems to be a bad choice.

Warning: This work is preliminary, all conditions are not fully factorized over all question!