LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

TWEAKING DEMOCRACY: INNOVATIONS IN DEMOCRATIC DECISION MAKING

THE WISDOM OF CROWDS

THE GROUP CAN BE SMART, SO YOU DON'T HAVE TO

Adrian Haret
a.haret@lmu.de

December 13, 2023

Let's warm up with a little pop quiz.

This bridge connects Manhattan to what other New York borough?

- ☐ Brooklyn
- ☐ Queens

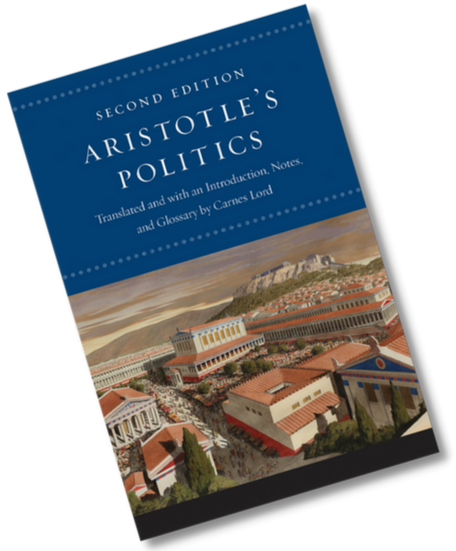
Before we find out the answer, let's get comfortable with the idea behind the wisdom of crowds.

THE WISDOM OF THE CROWDS IN THE WILD

PARTY-BOY ARISTOTLE

The many, who are not as individuals excellent men, nevertheless can, when they have come together, be better than the few best people, not individually but collectively…

…just as feasts to which many contribute are better than feasts provided at one person's expense.



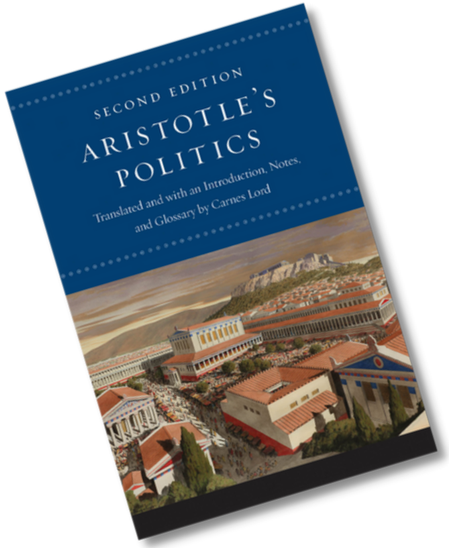
Aristotle (1984). *The Politics*. Translated and with an introduction by Carnes Lord. Chicago: University of Chicago Press,

PARTY-BOY ARISTOTLE

The many, who are not as individuals excellent men, nevertheless can, when they have come together, be better than the few best people, not individually but collectively…

…just as feasts to which many contribute are better than feasts provided at one person's expense.

In plain words: two (or more) heads are better than one.

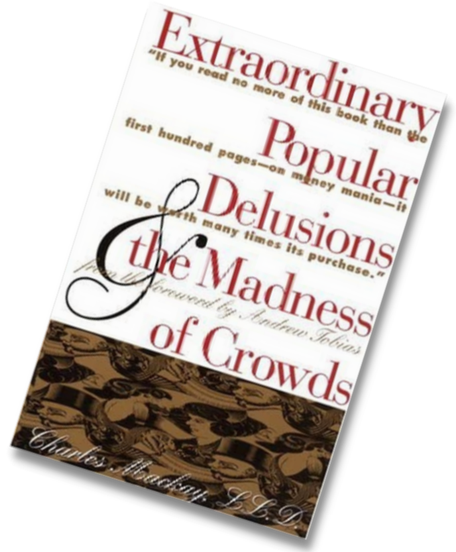


CHARLES MACKAY

No they're not.

So many examples of collective folly: economic bubbles, crusades, witch manias.

Brexit…



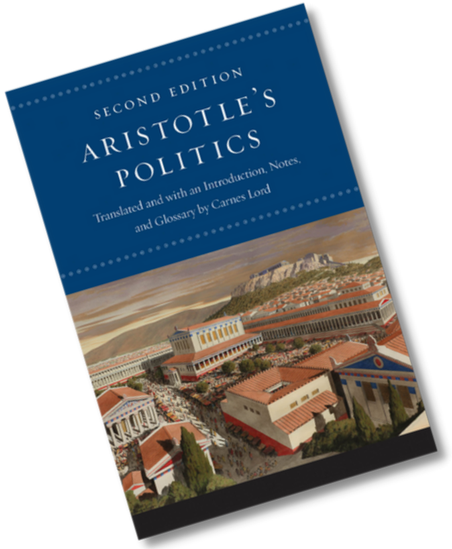
Aristotle (1984). *The Politics*. Translated and with an introduction by Carnes Lord. Chicago: University of Chicago Press,
MacKay, C. (1841). *Memoirs of Extraordinary Popular Delusions and the Madness of Crowds (Volumes I-III)*. London: Richard Bentley.

PARTY-BOY ARISTOTLE

The many, who are not as individuals excellent men, nevertheless can, when they have come together, be better than the few best people, not individually but collectively…

…just as feasts to which many contribute are better than feasts provided at one person's expense.

In plain words: two (or more) heads are better than one.



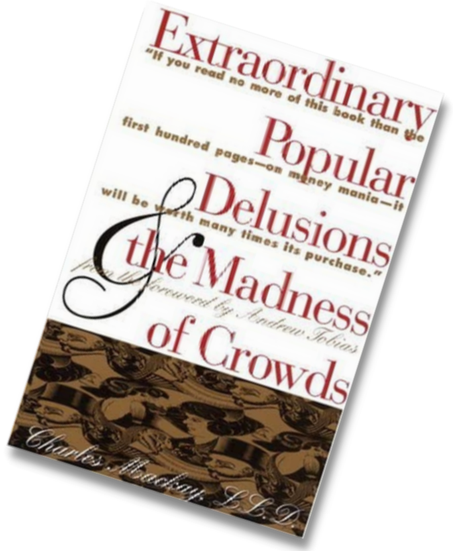
CHARLES MACKAY

No they're not.

So many examples of collective folly: economic bubbles, crusades, witch manias.

Brexit…

The success of Marvel movies…



Aristotle (1984). *The Politics*. Translated and with an introduction by Carnes Lord. Chicago: University of Chicago Press,
MacKay, C. (1841). *Memoirs of Extraordinary Popular Delusions and the Madness of Crowds (Volumes I-III)*. London: Richard Bentley.

FRANCIS GALTON

If I may be allowed!

I can personally attest to the accuracy of groups.

I am talking, of course, about

FRANCIS GALTON

If I may be allowed!

I can personally attest to the accuracy of groups.

I am talking, of course, about

the ox!

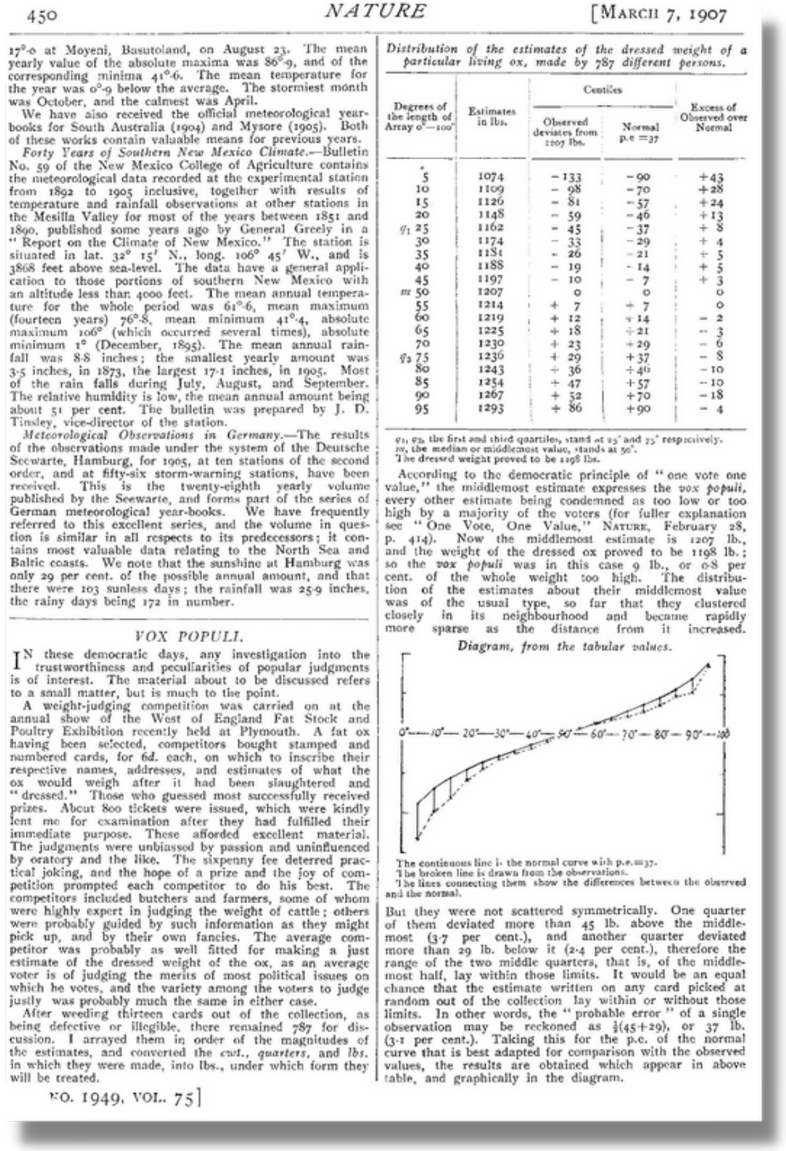
FRANCIS GALTON

A weight-judging competition was carried on at the annual show of the West of England Fat Stock and Poultry Exhibition recently held at Plymouth.

A fat ox having been selected, competitors bought stamped and numbered cards, for 6d. each, on which to inscribe their respective names, addresses, and estimates of what the ox would weigh after it had been slaughtered and "dressed".

Those who guessed most successfully received prizes.

About 800 tickets were issued, which were kindly lent me for examination after they had fulfilled their immediate purpose. .. [of which] there remained 787 for discussion.



450 NATURE [MARCH 7, 1907

17°·0 at Moyeni, Basutoland, on August 23. The mean yearly value of the absolute maxima was 86°·9, and of the corresponding minima 41°·6. The mean temperature for the year was 0°·9 below the average. The stormiest month was October, and the calmest was April.

We have also received the official meteorological year-books for South Australia (1904) and Mysore (1905). Both of these works contain valuable means for previous years.

Forty Years of Southern New Mexico Climate.—Bulletin No. 59 of the New Mexico College of Agriculture contains the meteorological data recorded at the experimental station from 1892 to 1905 inclusive, together with results of temperature and rainfall observations at other stations in the Mesilla Valley for most of the years between 1851 and 1890, published some years ago by General Greely in a "Report on the Climate of New Mexico." The station is situated in lat. 32° 15′ N., long. 106° 45′ W., and is 3868 feet above sea-level. The data have a general application to those portions of southern New Mexico with an altitude less than 4000 feet. The mean annual temperature for the whole period was 61°·6, mean maximum (fourteen years) 76°·8, mean minimum 41°·4, absolute maximum 106° (which occurred several times), absolute minimum 1° (December, 1895). The mean annual rainfall was 8·8 inches; the smallest yearly amount was 3·5 inches, in 1873, the largest 17·1 inches, in 1905. Most of the rain falls during July, August, and September. The relative humidity is low, the mean annual amount being about 51 per cent. The bulletin was prepared by J. D. Tinsley, vice-director of the station.

Meteorological Observations in Germany.—The results of the observations made under the system of the Deutsche Seewarte, Hamburg, for 1905, at ten stations of the second order, and at fifty-six storm-warning stations, have been received. This is the twenty-eighth yearly volume published by the Seewarte, and forms part of the series of German meteorological year-books. We have frequently referred to this excellent series, and the volume in question is similar in all respects to its predecessors; it contains most valuable data relating to the North Sea and Baltic coasts. We note that the sunshine at Hamburg was only 29 per cent. of the possible annual amount, and that there were 103 sunless days; the rainfall was 25·9 inches, the rainy days being 172 in number.

VOX POPULI.

IN these democratic days, any investigation into the trustworthiness and peculiarities of popular judgments is of interest. The material about to be discussed refers to a small matter, but is much to the point.

A weight-judging competition was carried on at the annual show of the West of England Fat Stock and Poultry Exhibition recently held at Plymouth. A fat ox having been selected, competitors bought stamped and numbered cards, for 6d. each, on which to inscribe their respective names, addresses, and estimates of what the ox would weigh after it had been slaughtered and "dressed." Those who guessed most successfully received prizes. About 800 tickets were issued, which were kindly lent me for examination after they had fulfilled their immediate purpose. These afforded excellent material. The judgments were unbiassed by passion and uninfluenced by oratory and the like. The sixpenny fee deterred practical joking, and the hope of a prize and the joy of competition prompted each competitor to do his best. The competitors included butchers and farmers, some of whom were highly expert in judging the weight of cattle; others were probably guided by such information as they might pick up, and by their own fancies. The average competitor was probably as well fitted for making a just estimate of the dressed weight of the ox, as an average voter is of judging the merits of most political issues on which he votes, and the variety among the voters to judge justly was probably much the same in either case.

After weeding thirteen cards out of the collection, as being defective or illegible, there remained 787 for discussion. I arrayed them in order of the magnitudes of the estimates, and converted the *cwt., quarters,* and *lbs.* in which they were made, into lbs., under which form they will be treated.

Distribution of the estimates of the dressed weight of a particular living ox, made by 787 different persons.

Degrees of the length of Array 0°—100°	Estimates in lbs.	Centiles: Observed deviates from 1207 lbs.	Centiles: Normal p.e. =37	Excess of Observed over Normal
*				
5	1074	−133	−90	+43
10	1109	− 98	−70	+28
15	1126	− 81	−57	+24
20	1148	− 59	−46	+13
q_1 25	1162	− 45	−37	+ 8
30	1174	− 33	−29	+ 4
35	1181	− 26	−21	+ 5
40	1188	− 19	−14	+ 5
45	1197	− 10	− 7	+ 3
m 50	1207	0	0	0
55	1214	+ 7	+ 7	0
60	1219	+ 12	+14	− 2
65	1225	+ 18	+21	− 3
70	1230	+ 23	+29	− 6
q_3 75	1236	+ 29	+37	− 8
80	1243	+ 36	+46	−10
85	1254	+ 47	+57	−10
90	1267	+ 52	+70	−18
95	1293	+ 86	+90	− 4

q_1, q_3, the first and third quartiles, stand at 25° and 75° respectively.
m, the median or middlemost value, stands at 50°.
The dressed weight proved to be 1198 lbs.

According to the democratic principle of "one vote one value," the middlemost estimate expresses the *vox populi*, every other estimate being condemned as too low or too high by a majority of the voters (for fuller explanation see "One Vote, One Value," NATURE, February 28, p. 414). Now the middlemost estimate is 1207 lb., and the weight of the dressed ox proved to be 1198 lb.; so the *vox populi* was in this case 9 lb., or 0·8 per cent. of the whole weight too high. The distribution of the estimates about their middlemost value was of the usual type, so far that they clustered closely in its neighbourhood and became rapidly more sparse as the distance from it increased.

Diagram, from the tabular values.

The continuous line is the normal curve with p.e.=37.
The broken line is drawn from the observations.
The lines connecting them show the differences between the observed and the normal.

But they were not scattered symmetrically. One quarter of them deviated more than 45 lb. above the middlemost (3·7 per cent.), and another quarter deviated more than 29 lb. below it (2·4 per cent.), therefore the range of the two middle quarters, that is, of the middlemost half, lay within those limits. It would be an equal chance that the estimate written on any card picked at random out of the collection lay within or without those limits. In other words, the "probable error" of a single observation may be reckoned as ¼(45+29), or 37 lb. (3·1 per cent.). Taking this for the p.c. of the normal curve that is best adapted for comparison with the observed values, the results are obtained which appear in above table, and graphically in the diagram.

Galton, F. (1907). Vox populi. *Nature*, 75(7), 450–451.

FRANCIS GALTON

A weight-judging competition was carried on at the annual show of the West of England Fat Stock and Poultry Exhibition recently held at Plymouth.

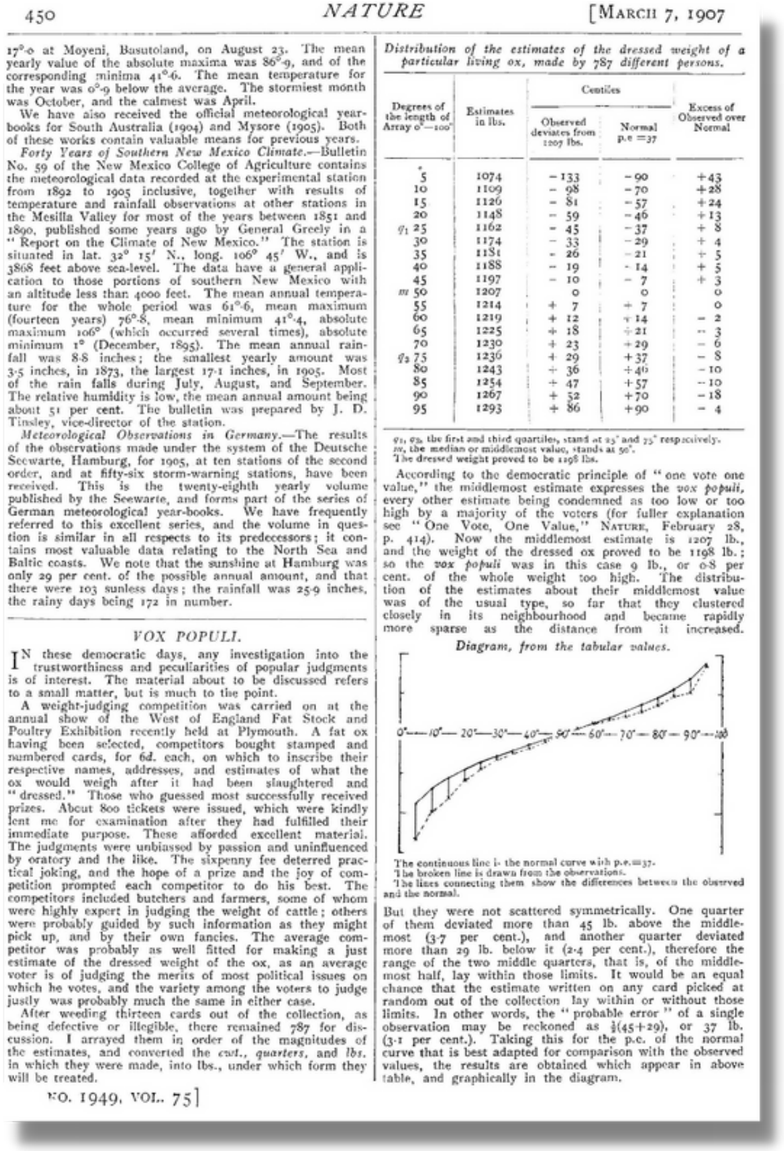
A fat ox having been selected, competitors bought stamped and numbered cards, for 6d. each, on which to inscribe their respective names, addresses, and estimates of what the ox would weigh after it had been slaughtered and "dressed".

Those who guessed most successfully received prizes.

About 800 tickets were issued, which were kindly lent me for examination after they had fulfilled their immediate purpose. .. [of which] there remained 787 for discussion.

Now the middlemost estimate is 1207 lb., and the weight of the dressed ox proved to be 1198 lb.

...so the *vox populi* was in this case 9 lb., or 0.8 per cent. of the whole weight too high.



450 NATURE [MARCH 7, 1907

17°·0 at Moyeni, Basutoland, on August 23. The mean yearly value of the absolute maxima was 86°·9, and of the corresponding minima 41°·6. The mean temperature for the year was 0°·9 below the average. The stormiest month was October, and the calmest was April.

We have also received the official meteorological year-books for South Australia (1904) and Mysore (1905). Both of these works contain valuable means for previous years.

Forty Years of Southern New Mexico Climate.—Bulletin No. 59 of the New Mexico College of Agriculture contains the meteorological data recorded at the experimental station from 1892 to 1905 inclusive, together with results of temperature and rainfall observations at other stations in the Mesilla Valley for most of the years between 1851 and 1890, published some years ago by General Greely in a "Report on the Climate of New Mexico." The station is situated in lat. 32° 15′ N., long. 106° 45′ W., and is 3868 feet above sea-level. The data have a general application to those portions of southern New Mexico with an altitude less than 4000 feet. The mean annual temperature for the whole period was 61°·6, mean maximum (fourteen years) 76°·8, mean minimum 41°·4, absolute maximum 106° (which occurred several times), absolute minimum 1° (December, 1895). The mean annual rainfall was 8·8 inches; the smallest yearly amount was 3·5 inches, in 1873, the largest 17·1 inches, in 1905. Most of the rain falls during July, August, and September. The relative humidity is low, the mean annual amount being about 51 per cent. The bulletin was prepared by J. D. Tinsley, vice-director of the station.

Meteorological Observations in Germany.—The results of the observations made under the system of the Deutsche Seewarte, Hamburg, for 1905, at ten stations of the second order, and at fifty-six storm-warning stations, have been received. This is the twenty-eighth yearly volume published by the Seewarte, and forms part of the series of German meteorological year-books. We have frequently referred to this excellent series, and the volume in question is similar in all respects to its predecessors; it contains most valuable data relating to the North Sea and Baltic coasts. We note that the sunshine at Hamburg was only 29 per cent. of the possible annual amount, and that there were 103 sunless days; the rainfall was 25·9 inches, the rainy days being 172 in number.

VOX POPULI.

IN these democratic days, any investigation into the trustworthiness and peculiarities of popular judgments is of interest. The material about to be discussed refers to a small matter, but is much to the point.

A weight-judging competition was carried on at the annual show of the West of England Fat Stock and Poultry Exhibition recently held at Plymouth. A fat ox having been selected, competitors bought stamped and numbered cards, for 6d. each, on which to inscribe their respective names, addresses, and estimates of what the ox would weigh after it had been slaughtered and "dressed." Those who guessed most successfully received prizes. About 800 tickets were issued, which were kindly lent me for examination after they had fulfilled their immediate purpose. These afforded excellent material. The judgments were unbiassed by passion and uninfluenced by oratory and the like. The sixpenny fee deterred practical joking, and the hope of a prize and the joy of competition prompted each competitor to do his best. The competitors included butchers and farmers, some of whom were highly expert in judging the weight of cattle; others were probably guided by such information as they might pick up, and by their own fancies. The average competitor was probably as well fitted for making a just estimate of the dressed weight of the ox, as an average voter is of judging the merits of most political issues on which he votes, and the variety among the voters to judge justly was probably much the same in either case.

After weeding thirteen cards out of the collection, as being defective or illegible, there remained 787 for discussion. I arrayed them in order of the magnitudes of the estimates, and converted the *cwt.*, *quarters*, and *lbs.* in which they were made, into lbs., under which form they will be treated.

Distribution of the estimates of the dressed weight of a particular living ox, made by 787 different persons.

Degrees of the length of Array 0°—100°	Estimates in lbs.	Centiles: Observed deviates from 1207 lbs.	Centiles: Normal p.e. =37	Excess of Observed over Normal
*				
5	1074	−133	−90	+43
10	1109	− 98	−70	+28
15	1126	− 81	−57	+24
20	1148	− 59	−46	+13
q_1 25	1162	− 45	−37	+ 8
30	1174	− 33	−29	+ 4
35	1181	− 26	−21	+ 5
40	1188	− 19	−14	+ 5
45	1197	− 10	− 7	+ 3
m 50	1207	0	0	0
55	1214	+ 7	+ 7	0
60	1219	+ 12	+ 14	− 2
65	1225	+ 18	+ 21	− 3
70	1230	+ 23	+ 29	− 6
q_3 75	1236	+ 29	+37	− 8
80	1243	+ 36	+46	−10
85	1254	+ 47	+57	−10
90	1267	+ 52	+70	−18
95	1293	+ 86	+90	− 4

q_1, q_3, the first and third quartiles, stand at 25° and 75° respectively.
m, the median or middlemost value, stands at 50°.
The dressed weight proved to be 1198 lbs.

According to the democratic principle of "one vote one value," the middlemost estimate expresses the *vox populi*, every other estimate being condemned as too low or too high by a majority of the voters (for fuller explanation see "One Vote, One Value," NATURE, February 28, p. 414). Now the middlemost estimate is 1207 lb., and the weight of the dressed ox proved to be 1198 lb.; so the *vox populi* was in this case 9 lb., or 0·8 per cent. of the whole weight too high. The distribution of the estimates about their middlemost value was of the usual type, so far that they clustered closely in its neighbourhood and became rapidly more sparse as the distance from it increased.

Diagram, from the tabular values.

The continuous line is the normal curve with p.e.=37.
The broken line is drawn from the observations.
The lines connecting them show the differences between the observed and the normal.

But they were not scattered symmetrically. One quarter of them deviated more than 45 lb. above the middlemost (3·7 per cent.), and another quarter deviated more than 29 lb. below it (2·4 per cent.), therefore the range of the two middle quarters, that is, of the middlemost half, lay within those limits. It would be an equal chance that the estimate written on any card picked at random out of the collection lay within or without those limits. In other words, the "probable error" of a single observation may be reckoned as ½(45+29), or 37 lb. (3·1 per cent.). Taking this for the p.c. of the normal curve that is best adapted for comparison with the observed values, the results are obtained which appear in above table, and graphically in the diagram.

NO. 1949, VOL. 75]

Galton, F. (1907). Vox populi. *Nature*, 75(7), 450—451.

FRANCIS GALTON

A weight-judging competition was carried on at the annual show of the West of England Fat Stock and Poultry Exhibition recently held at Plymouth.

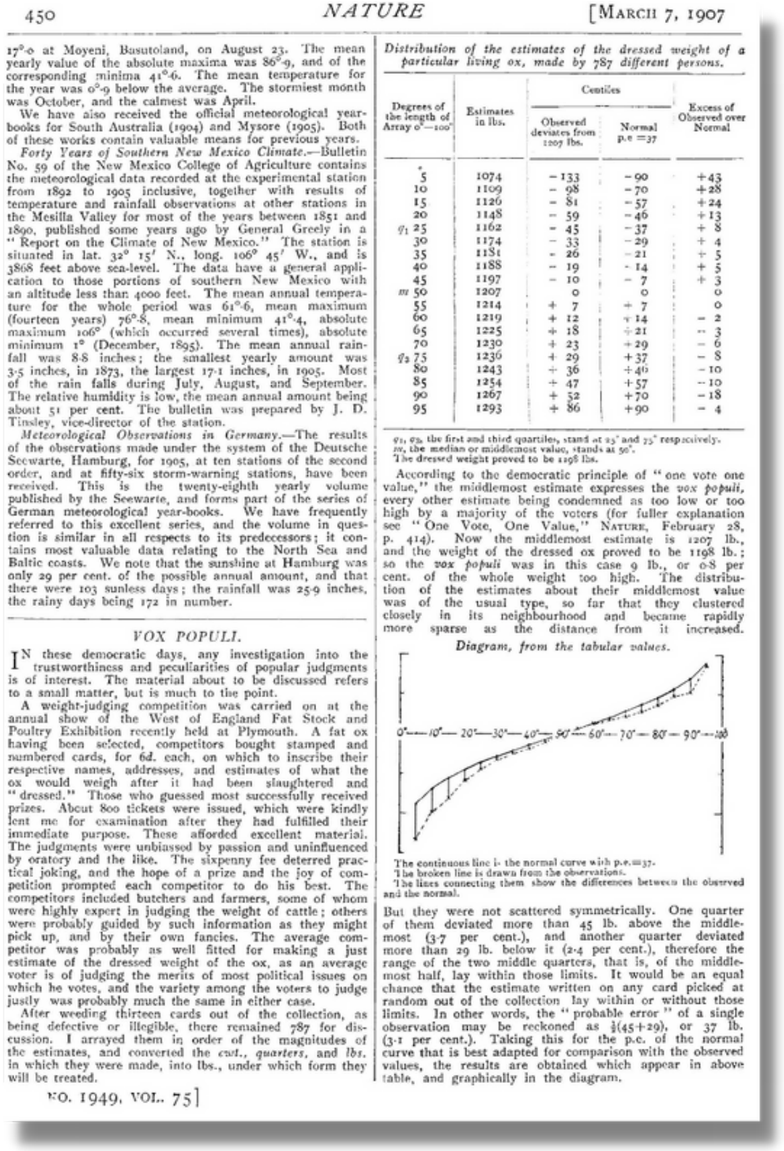
A fat ox having been selected, competitors bought stamped and numbered cards, for 6d. each, on which to inscribe their respective names, addresses, and estimates of what the ox would weigh after it had been slaughtered and "dressed".

Those who guessed most successfully received prizes.

About 800 tickets were issued, which were kindly lent me for examination after they had fulfilled their immediate purpose. .. [of which] there remained 787 for discussion.

Now the middlemost estimate is 1207 lb., and the weight of the dressed ox proved to be 1198 lb.

...so the *vox populi* was in this case 9 lb., or 0.8 per cent. of the whole weight too high.



450 NATURE [MARCH 7, 1907

17°·0 at Moyeni, Basutoland, on August 23. The mean yearly value of the absolute maxima was 86°·9, and of the corresponding minima 41°·6. The mean temperature for the year was 0°·9 below the average. The stormiest month was October, and the calmest was April.

We have also received the official meteorological year-books for South Australia (1904) and Mysore (1905). Both of these works contain valuable means for previous years.

Forty Years of Southern New Mexico Climate.—Bulletin No. 59 of the New Mexico College of Agriculture contains the meteorological data recorded at the experimental station from 1892 to 1905 inclusive, together with results of temperature and rainfall observations at other stations in the Mesilla Valley for most of the years between 1851 and 1890, published some years ago by General Greely in a "Report on the Climate of New Mexico." The station is situated in lat. 32° 15′ N., long. 106° 45′ W., and is 3868 feet above sea-level. The data have a general application to those portions of southern New Mexico with an altitude less than 4000 feet. The mean annual temperature for the whole period was 61°·6, mean maximum (fourteen years) 76°·8, mean minimum 41°·4, absolute maximum 106° (which occurred several times), absolute minimum 1° (December, 1895). The mean annual rainfall was 8·8 inches; the smallest yearly amount was 3·5 inches, in 1873, the largest 17·1 inches, in 1905. Most of the rain falls during July, August, and September. The relative humidity is low, the mean annual amount being about 51 per cent. The bulletin was prepared by J. D. Tinsley, vice-director of the station.

Meteorological Observations in Germany.—The results of the observations made under the system of the Deutsche Seewarte, Hamburg, for 1905, at ten stations of the second order, and at fifty-six storm-warning stations, have been received. This is the twenty-eighth yearly volume published by the Seewarte, and forms part of the series of German meteorological year-books. We have frequently referred to this excellent series, and the volume in question is similar in all respects to its predecessors; it contains most valuable data relating to the North Sea and Baltic coasts. We note that the sunshine at Hamburg was only 29 per cent. of the possible annual amount, and that there were 103 sunless days; the rainfall was 25·9 inches, the rainy days being 172 in number.

VOX POPULI.

IN these democratic days, any investigation into the trustworthiness and peculiarities of popular judgments is of interest. The material about to be discussed refers to a small matter, but is much to the point.

A weight-judging competition was carried on at the annual show of the West of England Fat Stock and Poultry Exhibition recently held at Plymouth. A fat ox having been selected, competitors bought stamped and numbered cards, for 6d. each, on which to inscribe their respective names, addresses, and estimates of what the ox would weigh after it had been slaughtered and "dressed." Those who guessed most successfully received prizes. About 800 tickets were issued, which were kindly lent me for examination after they had fulfilled their immediate purpose. These afforded excellent material. The judgments were unbiassed by passion and uninfluenced by oratory and the like. The sixpenny fee deterred practical joking, and the hope of a prize and the joy of competition prompted each competitor to do his best. The competitors included butchers and farmers, some of whom were highly expert in judging the weight of cattle; others were probably guided by such information as they might pick up, and by their own fancies. The average competitor was probably as well fitted for making a just estimate of the dressed weight of the ox, as an average voter is of judging the merits of most political issues on which he votes, and the variety among the voters to judge justly was probably much the same in either case.

After weeding thirteen cards out of the collection, as being defective or illegible, there remained 787 for discussion. I arrayed them in order of the magnitudes of the estimates, and converted the *cwt.*, *quarters*, and *lbs.* in which they were made, into lbs., under which form they will be treated.

Distribution of the estimates of the dressed weight of a particular living ox, made by 787 different persons.

Degrees of the length of Array 0°—100°	Estimates in lbs.	Centiles: Observed deviates from 1207 lbs.	Centiles: Normal p.e. =37	Excess of Observed over Normal
*				
5	1074	−133	−90	+43
10	1109	− 98	−70	+28
15	1126	− 81	−57	+24
20	1148	− 59	−46	+13
q_1 25	1162	− 45	−37	+ 8
30	1174	− 33	−29	+ 4
35	1181	− 26	−21	+ 5
40	1188	− 19	−14	+ 5
45	1197	− 10	− 7	+ 3
m 50	1207	0	0	0
55	1214	+ 7	+ 7	0
60	1219	+ 12	+ 14	− 2
65	1225	+ 18	+ 21	− 3
70	1230	+ 23	+ 29	− 6
q_3 75	1236	+ 29	+37	− 8
80	1243	+ 36	+46	−10
85	1254	+ 47	+57	−10
90	1267	+ 52	+70	−18
95	1293	+ 86	+90	− 4

q_1, q_3, the first and third quartiles, stand at 25° and 75° respectively.
m, the median or middlemost value, stands at 50°.
The dressed weight proved to be 1198 lbs.

According to the democratic principle of "one vote one value," the middlemost estimate expresses the *vox populi*, every other estimate being condemned as too low or too high by a majority of the voters (for fuller explanation see "One Vote, One Value," NATURE, February 28, p. 414). Now the middlemost estimate is 1207 lb., and the weight of the dressed ox proved to be 1198 lb.; so the *vox populi* was in this case 9 lb., or 0·8 per cent. of the whole weight too high. The distribution of the estimates about their middlemost value was of the usual type, so far that they clustered closely in its neighbourhood and became rapidly more sparse as the distance from it increased.

Diagram, from the tabular values.

The continuous line is the normal curve with p.e.=37.
The broken line is drawn from the observations.
The lines connecting them show the differences between the observed and the normal.

But they were not scattered symmetrically. One quarter of them deviated more than 45 lb. above the middlemost (3·7 per cent.), and another quarter deviated more than 29 lb. below it (2·4 per cent.), therefore the range of the two middle quarters, that is, of the middlemost half, lay within those limits. It would be an equal chance that the estimate written on any card picked at random out of the collection lay within or without those limits. In other words, the "probable error" of a single observation may be reckoned as ½(45+29), or 37 lb. (3·1 per cent.). Taking this for the p.e. of the normal curve that is best adapted for comparison with the observed values, the results are obtained which appear in above table, and graphically in the diagram.

NO. 1949, VOL. 75]

Galton, F. (1907). Vox populi. *Nature*, 75(7), 450–451.

FRANCIS GALTON

By middlemost I mean what you might call today the median.

People have since pointed out that the mean was even more accurate: 1197 lbs.

17°·0 at Moyeni, Basutoland, on August 23. The mean yearly value of the absolute maxima was 86°·9, and of the corresponding minima 41°·6. The mean temperature for the year was 0°·9 below the average. The stormiest month was October, and the calmest was April.

We have also received the official meteorological year-books for South Australia (1904) and Mysore (1905). Both of these works contain valuable means for previous years.

Forty Years of Southern New Mexico Climate.—Bulletin No. 59 of the New Mexico College of Agriculture contains the meteorological data recorded at the experimental station from 1892 to 1905 inclusive, together with results of temperature and rainfall observations at other stations in the Mesilla Valley for most of the years between 1851 and 1890, published some years ago by General Greely in a "Report on the Climate of New Mexico." The station is situated in lat. 32° 15′ N., long. 106° 45′ W., and is 3868 feet above sea-level. The data have a general application to those portions of southern New Mexico with an altitude less than 4000 feet. The mean annual temperature for the whole period was 61°·6, mean maximum (fourteen years) 76°·8, mean minimum 41°·4, absolute maximum 106° (which occurred several times), absolute minimum 1° (December, 1895). The mean annual rainfall was 8·8 inches; the smallest yearly amount was 3·5 inches, in 1873, the largest 17·1 inches, in 1905. Most of the rain falls during July, August, and September. The relative humidity is low, the mean annual amount being about 51 per cent. The bulletin was prepared by J. D. Tinsley, vice-director of the station.

Meteorological Observations in Germany.—The results of the observations made under the system of the Deutsche Seewarte, Hamburg, for 1905, at ten stations of the second order, and at fifty-six storm-warning stations, have been received. This is the twenty-eighth yearly volume published by the Seewarte, and forms part of the series of German meteorological year-books. We have frequently referred to this excellent series, and the volume in question is similar in all respects to its predecessors; it contains most valuable data relating to the North Sea and Baltic coasts. We note that the sunshine at Hamburg was only 29 per cent. of the possible annual amount, and that there were 103 sunless days; the rainfall was 25·9 inches, the rainy days being 172 in number.

VOX POPULI.

IN these democratic days, any investigation into the trustworthiness and peculiarities of popular judgments is of interest. The material about to be discussed refers to a small matter, but is much to the point.

A weight-judging competition was carried on at the annual show of the West of England Fat Stock and Poultry Exhibition recently held at Plymouth. A fat ox having been selected, competitors bought stamped and numbered cards, for 6*d.* each, on which to inscribe their respective names, addresses, and estimates of what the ox would weigh after it had been slaughtered and "dressed." Those who guessed most successfully received prizes. About 800 tickets were issued, which were kindly lent me for examination after they had fulfilled their immediate purpose. These afforded excellent material. The judgments were unbiassed by passion and uninfluenced by oratory and the like. The sixpenny fee deterred practical joking, and the hope of a prize and the joy of competition prompted each competitor to do his best. The competitors included butchers and farmers, some of whom were highly expert in judging the weight of cattle; others were probably guided by such information as they might pick up, and by their own fancies. The average competitor was probably as well fitted for making a just estimate of the dressed weight of the ox, as an average voter is of judging the merits of most political issues on which he votes, and the variety among the voters to judge justly was probably much the same in either case.

After weeding thirteen cards out of the collection, as being defective or illegible, there remained 787 for discussion. I arrayed them in order of the magnitudes of the estimates, and converted the *cwt.*, *quarters*, and *lbs.* in which they were made, into lbs., under which form they will be treated.

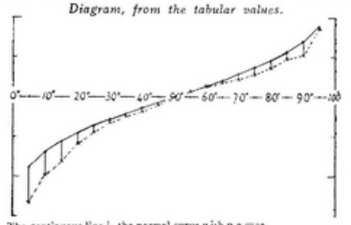
Distribution of the estimates of the dressed weight of a particular living ox, made by 787 different persons.

Degrees of the length of Array 0°—100°	Estimates in lbs.	Centiles: Observed deviates from 1207 lbs.	Centiles: Normal p.e. =37	Excess of Observed over Normal
°				
5	1074	−133	−90	+43
10	1109	− 98	−70	+28
15	1126	− 81	−57	+24
20	1148	− 59	−46	+13
q_1 25	1162	− 45	−37	+ 8
30	1174	− 33	−29	+ 4
35	1181	− 26	−21	+ 5
40	1188	− 19	−14	+ 5
45	1197	− 10	− 7	+ 3
m 50	1207	0	0	0
55	1214	+ 7	+ 7	0
60	1219	+ 12	+14	− 2
65	1225	+ 18	+21	− 3
70	1230	+ 23	+29	− 6
q_3 75	1236	+ 29	+37	− 8
80	1243	+ 36	+46	−10
85	1254	+ 47	+57	−10
90	1267	+ 52	+70	−18
95	1293	+ 86	+90	− 4

q_1, q_3, the first and third quartiles, stand at 25° and 75° respectively.
m, the median or middlemost value, stands at 50°.
The dressed weight proved to be 1198 lbs.

According to the democratic principle of "one vote one value," the middlemost estimate expresses the *vox populi*, every other estimate being condemned as too low or too high by a majority of the voters (for fuller explanation see "One Vote, One Value," NATURE, February 28, p. 414). Now the middlemost estimate is 1207 lb., and the weight of the dressed ox proved to be 1198 lb.; so the *vox populi* was in this case 9 lb., or 0·8 per cent. of the whole weight too high. The distribution of the estimates about their middlemost value was of the usual type, so far that they clustered closely in its neighbourhood and became rapidly more sparse as the distance from it increased.

Diagram, from the tabular values.



The continuous line is the normal curve with p.e.=37.
The broken line is drawn from the observations.
The lines connecting them show the differences between the observed and the normal.

But they were not scattered symmetrically. One quarter of them deviated more than 45 lb. above the middlemost (3·7 per cent.), and another quarter deviated more than 29 lb. below it (2·4 per cent.), therefore the range of the two middle quarters, that is, of the middlemost half, lay within those limits. It would be an equal chance that the estimate written on any card picked at random out of the collection lay within or without those limits. In other words, the "probable error" of a single observation may be reckoned as ½(45+29), or 37 lb. (3·1 per cent.). Taking this for the p.e. of the normal curve that is best adapted for comparison with the observed values, the results are obtained which appear in above table, and graphically in the diagram.

Galton, F. (1907). Vox populi. *Nature*, 75(7), 450–451.

FRANCIS GALTON

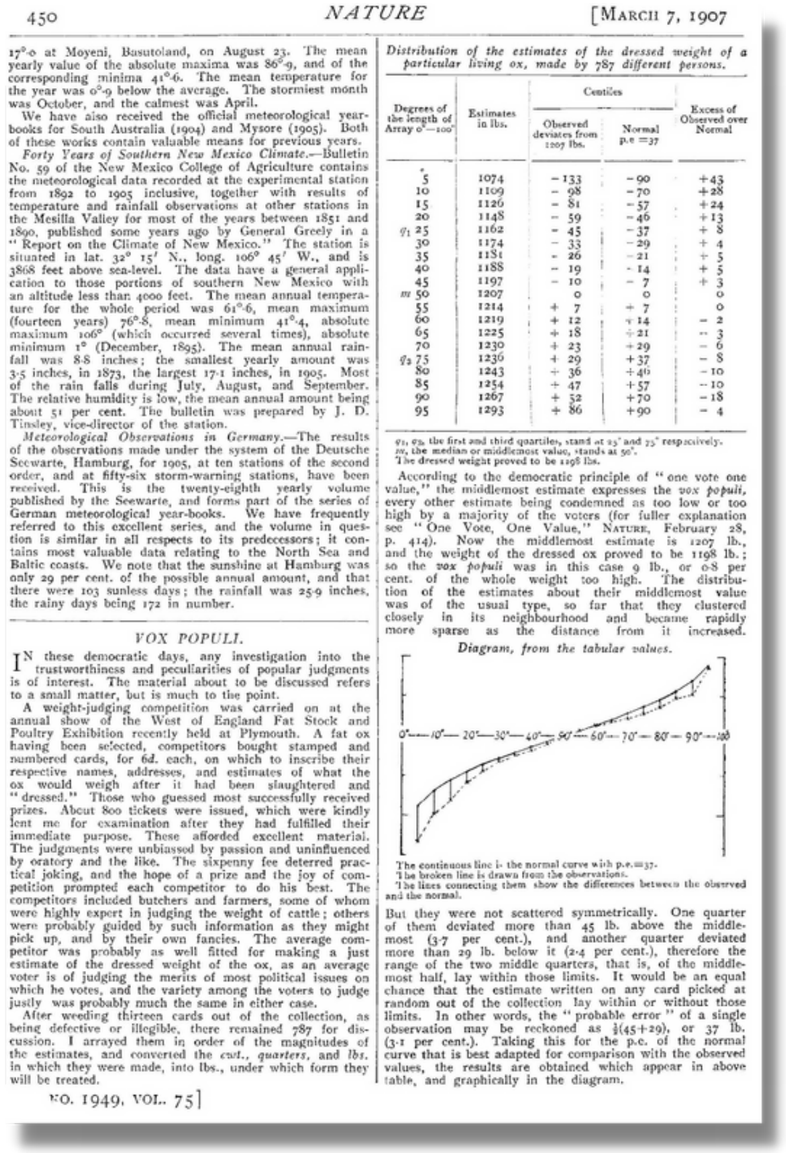
By middlemost I mean what you might call today the median.

People have since pointed out that the mean was even more accurate: 1197 lbs.

!!!

THE OX

The crowd was, on average, within 1 lb of the true weight!



450 NATURE [MARCH 7, 1907

17°·0 at Moyeni, Basutoland, on August 23. The mean yearly value of the absolute maxima was 86°·9, and of the corresponding minima 41°·6. The mean temperature for the year was 0°·9 below the average. The stormiest month was October, and the calmest was April.

We have also received the official meteorological year-books for South Australia (1904) and Mysore (1905). Both of these works contain valuable means for previous years.

Forty Years of Southern New Mexico Climate.—Bulletin No. 59 of the New Mexico College of Agriculture contains the meteorological data recorded at the experimental station from 1892 to 1905 inclusive, together with results of temperature and rainfall observations at other stations in the Mesilla Valley for most of the years between 1851 and 1890, published some years ago by General Greely in a "Report on the Climate of New Mexico." The station is situated in lat. 32° 15′ N., long. 106° 45′ W., and is 3868 feet above sea-level. The data have a general application to those portions of southern New Mexico with an altitude less than 4000 feet. The mean annual temperature for the whole period was 61°·6, mean maximum (fourteen years) 76°·8, mean minimum 41°·4, absolute maximum 106° (which occurred several times), absolute minimum 1° (December, 1895). The mean annual rainfall was 8·8 inches; the smallest yearly amount was 3·5 inches, in 1873, the largest 17·1 inches, in 1905. Most of the rain falls during July, August, and September. The relative humidity is low, the mean annual amount being about 51 per cent. The bulletin was prepared by J. D. Tinsley, vice-director of the station.

Meteorological Observations in Germany.—The results of the observations made under the system of the Deutsche Seewarte, Hamburg, for 1905, at ten stations of the second order, and at fifty-six storm-warning stations, have been received. This is the twenty-eighth yearly volume published by the Seewarte, and forms part of the series of German meteorological year-books. We have frequently referred to this excellent series, and the volume in question is similar in all respects to its predecessors; it contains most valuable data relating to the North Sea and Baltic coasts. We note that the sunshine at Hamburg was only 29 per cent. of the possible annual amount, and that there were 103 sunless days; the rainfall was 25·9 inches, the rainy days being 172 in number.

VOX POPULI.

IN these democratic days, any investigation into the trustworthiness and peculiarities of popular judgments is of interest. The material about to be discussed refers to a small matter, but is much to the point.

A weight-judging competition was carried on at the annual show of the West of England Fat Stock and Poultry Exhibition recently held at Plymouth. A fat ox having been selected, competitors bought stamped and numbered cards, for 6*d.* each, on which to inscribe their respective names, addresses, and estimates of what the ox would weigh after it had been slaughtered and "dressed." Those who guessed most successfully received prizes. About 800 tickets were issued, which were kindly lent me for examination after they had fulfilled their immediate purpose. These afforded excellent material. The judgments were unbiassed by passion and uninfluenced by oratory and the like. The sixpenny fee deterred practical joking, and the hope of a prize and the joy of competition prompted each competitor to do his best. The competitors included butchers and farmers, some of whom were highly expert in judging the weight of cattle; others were probably guided by such information as they might pick up, and by their own fancies. The average competitor was probably as well fitted for making a just estimate of the dressed weight of the ox, as an average voter is of judging the merits of most political issues on which he votes, and the variety among the voters to judge justly was probably much the same in either case.

After weeding thirteen cards out of the collection, as being defective or illegible, there remained 787 for discussion. I arrayed them in order of the magnitudes of the estimates, and converted the *cwt.*, *quarters*, and *lbs.* in which they were made, into lbs., under which form they will be treated.

Distribution of the estimates of the dressed weight of a particular living ox, made by 787 different persons.

Degrees of the length of Array 0°—100°	Estimates in lbs.	Centiles: Observed deviates from 1207 lbs.	Centiles: Normal p.e. =37	Excess of Observed over Normal
*				
5	1074	−133	−90	+43
10	1109	− 98	−70	+28
15	1126	− 81	−57	+24
20	1148	− 59	−46	+13
q_1 25	1162	− 45	−37	+ 8
30	1174	− 33	−29	+ 4
35	1181	− 26	−21	+ 5
40	1188	− 19	−14	+ 5
45	1197	− 10	− 7	+ 3
m 50	1207	0	0	0
55	1214	+ 7	+ 7	0
60	1219	+ 12	+14	− 2
65	1225	+ 18	+21	− 3
70	1230	+ 23	+29	− 6
q_3 75	1236	+ 29	+37	− 8
80	1243	+ 36	+46	−10
85	1254	+ 47	+57	−10
90	1267	+ 52	+70	−18
95	1293	+ 86	+90	− 4

q_1, q_3, the first and third quartiles, stand at 25° and 75° respectively.
m, the median or middlemost value, stands at 50°.
The dressed weight proved to be 1198 lbs.

According to the democratic principle of "one vote one value," the middlemost estimate expresses the *vox populi*, every other estimate being condemned as too low or too high by a majority of the voters (for fuller explanation see "One Vote, One Value," NATURE, February 28, p. 414). Now the middlemost estimate is 1207 lb., and the weight of the dressed ox proved to be 1198 lb.; so the *vox populi* was in this case 9 lb., or 0·8 per cent. of the whole weight too high. The distribution of the estimates about their middlemost value was of the usual type, so far that they clustered closely in its neighbourhood and became rapidly more sparse as the distance from it increased.

Diagram, from the tabular values.

The continuous line is the normal curve with p.e.=37.
The broken line is drawn from the observations.
The lines connecting them show the differences between the observed and the normal.

But they were not scattered symmetrically. One quarter of them deviated more than 45 lb. above the middlemost (3·7 per cent.), and another quarter deviated more than 29 lb. below it (2·4 per cent.), therefore the range of the two middle quarters, that is, of the middlemost half, lay within those limits. It would be an equal chance that the estimate written on any card picked at random out of the collection lay within or without those limits. In other words, the "probable error" of a single observation may be reckoned as ½(45+29), or 37 lb. (3·1 per cent.). Taking this for the p.e. of the normal curve that is best adapted for comparison with the observed values, the results are obtained which appear in above table, and graphically in the diagram.

Galton, F. (1907). Vox populi. *Nature*, 75(7), 450–451.

FRANCIS GALTON

By middlemost I mean what you might call today the median.

People have since pointed out that the mean was even more accurate: 1197 lbs.

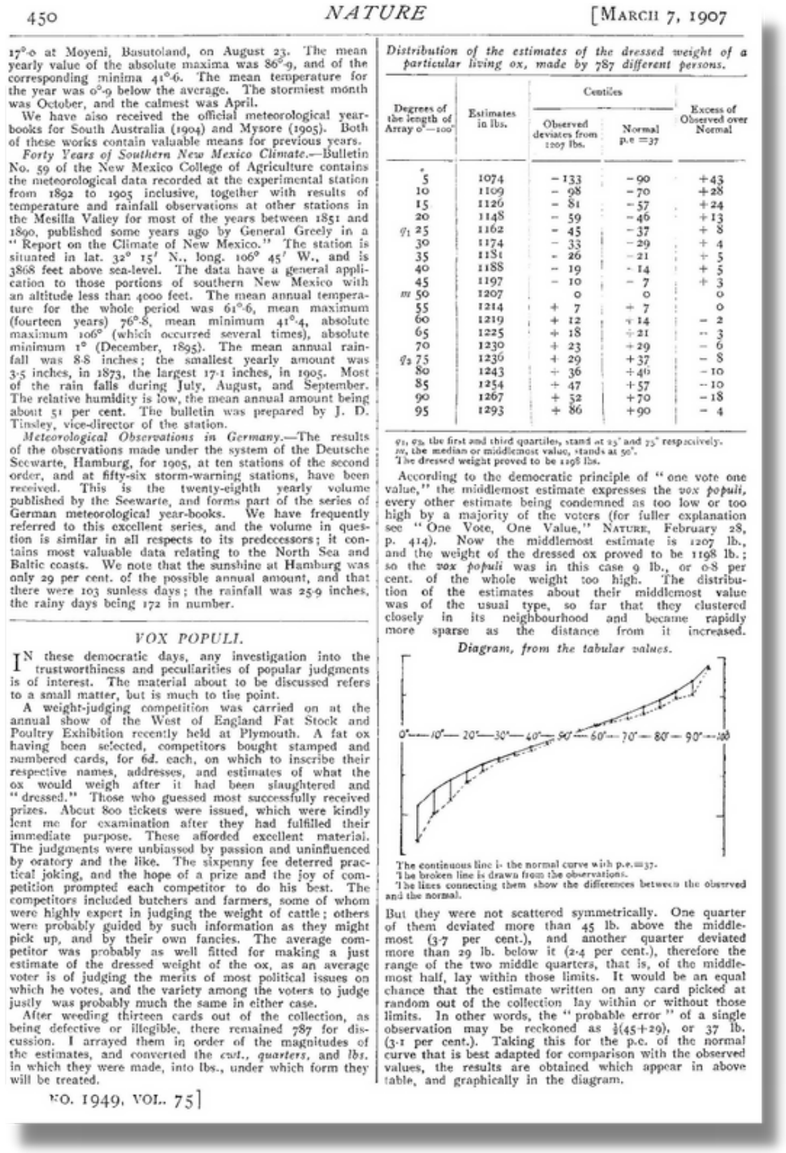
!!!

THE OX

The crowd was, on average, within 1 lb of the true weight!

FRANCIS GALTON

This result is, I think, more creditable to the trustworthiness of a democratic judgment than might have been expected.



450 NATURE [MARCH 7, 1907

17°·0 at Moyeni, Basutoland, on August 23. The mean yearly value of the absolute maxima was 86°·9, and of the corresponding minima 41°·6. The mean temperature for the year was 0°·9 below the average. The stormiest month was October, and the calmest was April.

We have also received the official meteorological year-books for South Australia (1904) and Mysore (1905). Both of these works contain valuable means for previous years.

Forty Years of Southern New Mexico Climate.—Bulletin No. 59 of the New Mexico College of Agriculture contains the meteorological data recorded at the experimental station from 1892 to 1905 inclusive, together with results of temperature and rainfall observations at other stations in the Mesilla Valley for most of the years between 1851 and 1890, published some years ago by General Greely in a "Report on the Climate of New Mexico." The station is situated in lat. 32° 15′ N., long. 106° 45′ W., and is 3868 feet above sea-level. The data have a general application to those portions of southern New Mexico with an altitude less than 4000 feet. The mean annual temperature for the whole period was 61°·6, mean maximum (fourteen years) 76°·8, mean minimum 41°·4, absolute maximum 106° (which occurred several times), absolute minimum 1° (December, 1895). The mean annual rainfall was 8·8 inches; the smallest yearly amount was 3·5 inches, in 1873, the largest 17·1 inches, in 1905. Most of the rain falls during July, August, and September. The relative humidity is low, the mean annual amount being about 51 per cent. The bulletin was prepared by J. D. Tinsley, vice-director of the station.

Meteorological Observations in Germany.—The results of the observations made under the system of the Deutsche Seewarte, Hamburg, for 1905, at ten stations of the second order, and at fifty-six storm-warning stations, have been received. This is the twenty-eighth yearly volume published by the Seewarte, and forms part of the series of German meteorological year-books. We have frequently referred to this excellent series, and the volume in question is similar in all respects to its predecessors; it contains most valuable data relating to the North Sea and Baltic coasts. We note that the sunshine at Hamburg was only 29 per cent. of the possible annual amount, and that there were 103 sunless days; the rainfall was 25·9 inches, the rainy days being 172 in number.

VOX POPULI.

IN these democratic days, any investigation into the trustworthiness and peculiarities of popular judgments is of interest. The material about to be discussed refers to a small matter, but is much to the point.

A weight-judging competition was carried on at the annual show of the West of England Fat Stock and Poultry Exhibition recently held at Plymouth. A fat ox having been selected, competitors bought stamped and numbered cards, for 6*d.* each, on which to inscribe their respective names, addresses, and estimates of what the ox would weigh after it had been slaughtered and "dressed." Those who guessed most successfully received prizes. About 800 tickets were issued, which were kindly lent me for examination after they had fulfilled their immediate purpose. These afforded excellent material. The judgments were unbiassed by passion and uninfluenced by oratory and the like. The sixpenny fee deterred practical joking, and the hope of a prize and the joy of competition prompted each competitor to do his best. The competitors included butchers and farmers, some of whom were highly expert in judging the weight of cattle; others were probably guided by such information as they might pick up, and by their own fancies. The average competitor was probably as well fitted for making a just estimate of the dressed weight of the ox, as an average voter is of judging the merits of most political issues on which he votes, and the variety among the voters to judge justly was probably much the same in either case.

After weeding thirteen cards out of the collection, as being defective or illegible, there remained 787 for discussion. I arrayed them in order of the magnitudes of the estimates, and converted the *cwt.*, *quarters*, and *lbs.* in which they were made, into lbs., under which form they will be treated.

Distribution of the estimates of the dressed weight of a particular living ox, made by 787 different persons.

Degrees of the length of Array 0°—100°	Estimates in lbs.	Centiles: Observed deviates from 1207 lbs.	Centiles: Normal p.e. =37	Excess of Observed over Normal
*				
5	1074	−133	−90	+43
10	1109	− 98	−70	+28
15	1126	− 81	−57	+24
20	1148	− 59	−46	+13
q_1 25	1162	− 45	−37	+ 8
30	1174	− 33	−29	+ 4
35	1181	− 26	−21	+ 5
40	1188	− 19	−14	+ 5
45	1197	− 10	− 7	+ 3
m 50	1207	0	0	0
55	1214	+ 7	+ 7	0
60	1219	+ 12	+ 14	− 2
65	1225	+ 18	+ 21	− 3
70	1230	+ 23	+ 29	− 6
q_3 75	1236	+ 29	+37	− 8
80	1243	+ 36	+46	−10
85	1254	+ 47	+57	−10
90	1267	+ 52	+70	−18
95	1293	+ 86	+90	− 4

q_1, q_3, the first and third quartiles, stand at 25° and 75° respectively.
m, the median or middlemost value, stands at 50°.
The dressed weight proved to be 1198 lbs.

According to the democratic principle of "one vote one value," the middlemost estimate expresses the *vox populi*, every other estimate being condemned as too low or too high by a majority of the voters (for fuller explanation see "One Vote, One Value," NATURE, February 28, p. 414). Now the middlemost estimate is 1207 lb., and the weight of the dressed ox proved to be 1198 lb.; so the *vox populi* was in this case 9 lb., or 0·8 per cent. of the whole weight too high. The distribution of the estimates about their middlemost value was of the usual type, so far that they clustered closely in its neighbourhood and became rapidly more sparse as the distance from it increased.

Diagram, from the tabular values.

The continuous line is the normal curve with p.e.=37.
The broken line is drawn from the observations.
The lines connecting them show the differences between the observed and the normal.

But they were not scattered symmetrically. One quarter of them deviated more than 45 lb. above the middlemost (3·7 per cent.), and another quarter deviated more than 29 lb. below it (2·4 per cent.), therefore the range of the two middle quarters, that is, of the middlemost half, lay within those limits. It would be an equal chance that the estimate written on any card picked at random out of the collection lay within or without those limits. In other words, the "probable error" of a single observation may be reckoned as ½(45+29), or 37 lb. (3·1 per cent.). Taking this for the p.e. of the normal curve that is best adapted for comparison with the observed values, the results are obtained which appear in above table, and graphically in the diagram.

NO. 1949, VOL. 75]

Galton, F. (1907). Vox populi. *Nature*, 75(7), 450–451.

So groups can be used to estimate the weights of oxen…

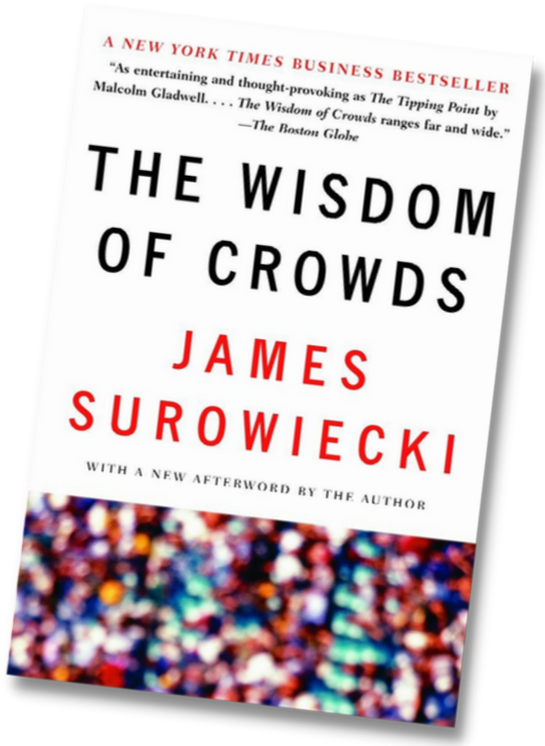
What's the big deal?

JAMES SUROWIECKI

There are many more examples of the wisdom of crowds at work.

Like the market response to the *Challenger* disaster.

Or the finding of the *Scorpion* submarine.



Surowiecki, J. (2005). *The Wisdom of Crowds*. Anchor

JAMES SUROWIECKI

There are many more examples of the wisdom of crowds at work.

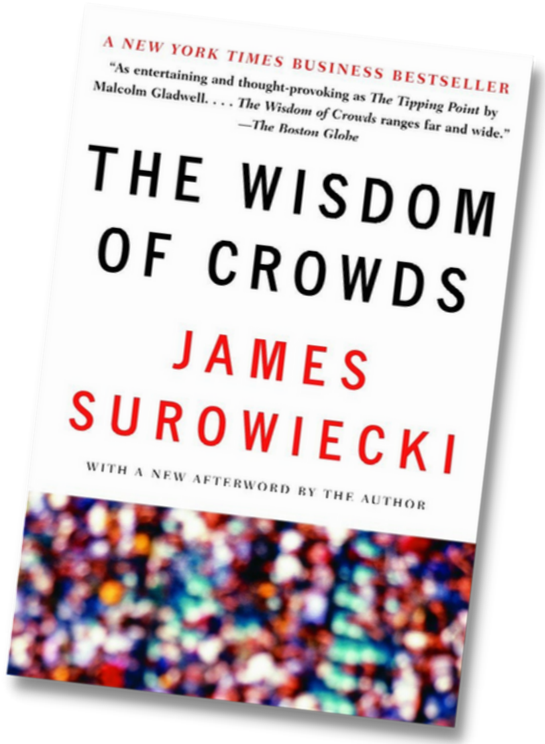
Like the market response to the *Challenger* disaster.

Or the finding of the *Scorpion* submarine.

IAIN COUZIN

Or golden shiners, as a group, finding patches of shade.

Even though individuals are bad at it.



Surowiecki, J. (2005). *The Wisdom of Crowds*. Anchor

Berdahl, A., Torney, C. J., Ioannou, C. C., Faria, J. J., & Couzin, I. D. (2013). Emergent sensing of complex environments by mobile animal groups. *Science*, 339(6119), 574-576.

Yong, E. (2013). The Real Wisdom of Crowds. *National Geographic*.

Wise crowds also give us a way of looking at democratic institutions: not just fair, but also good at finding solutions.

PARTY-POOPER PLATO

When you give people democratic choice they end up doing something stupid.

Good statecraft is like flying a plane.

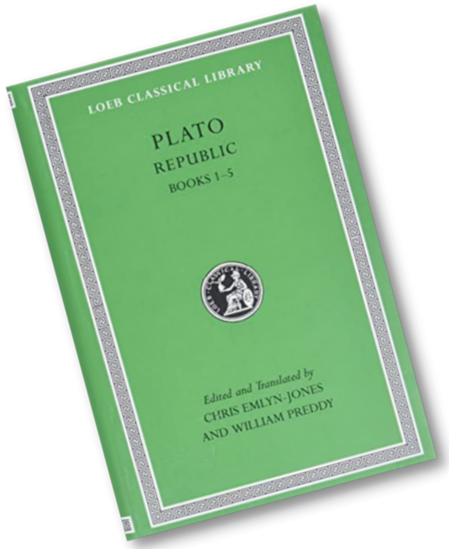
And you need a good pilot for that.

EDWIN HUTCHINS

Actually… flying a plane (or running a ship) requires a lot of coordination and teamwork.

HÉLÈNE LANDEMORE

Similarly, modern societies need the input of as many and as diverse parties as possible to work well.



Plato. *Republic*. Translated by Paul Shorey, 2 volumes. Loeb.
Hutchins, E. (1995). *Cognition in the Wild*. MIT Press
Landemore, H. (2012). *Democratic Reason*. Princeton University Press

Applications?

JUSTIN WOLFERS

Prediction markets!

Simple markets can be used to aggregate disparate information into efficient forecasts of uncertain future events.

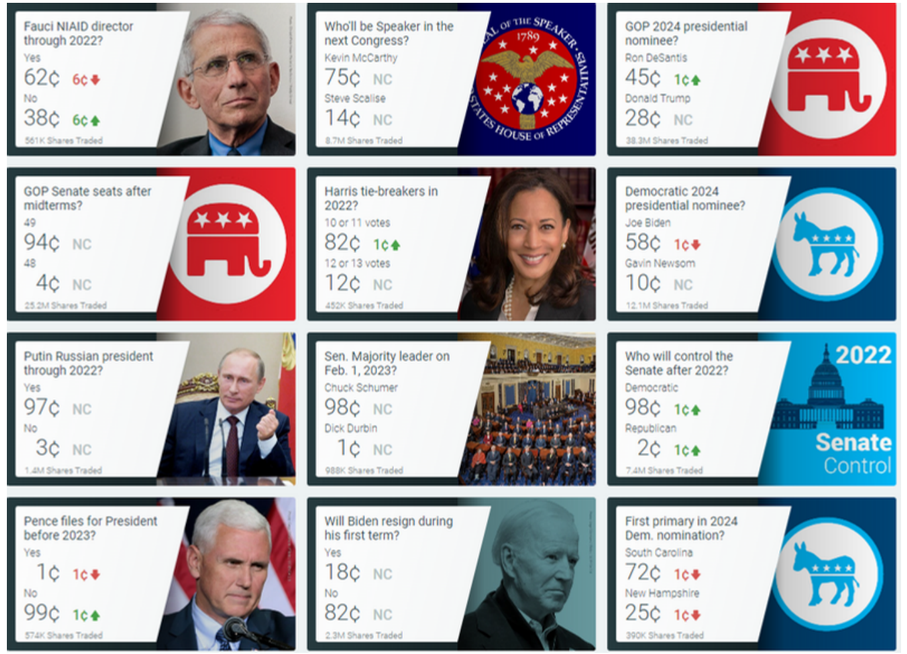
ERIC ZITZEWITZ

People buy and sell shares in future events (by a double auction).

The price indicates the collective estimate of the probability of the event.

See PredictIt.

PREDICTIT



JUSTIN WOLFERS

And other prediction platforms, like Metaculus or Good Judgment Open.

Wolfers, J., & Zitzewitz, E. (2004). Prediction Markets. *The Journal of Economic Perspectives: A Journal of the American Economic Association*, 18(2), 107–126.

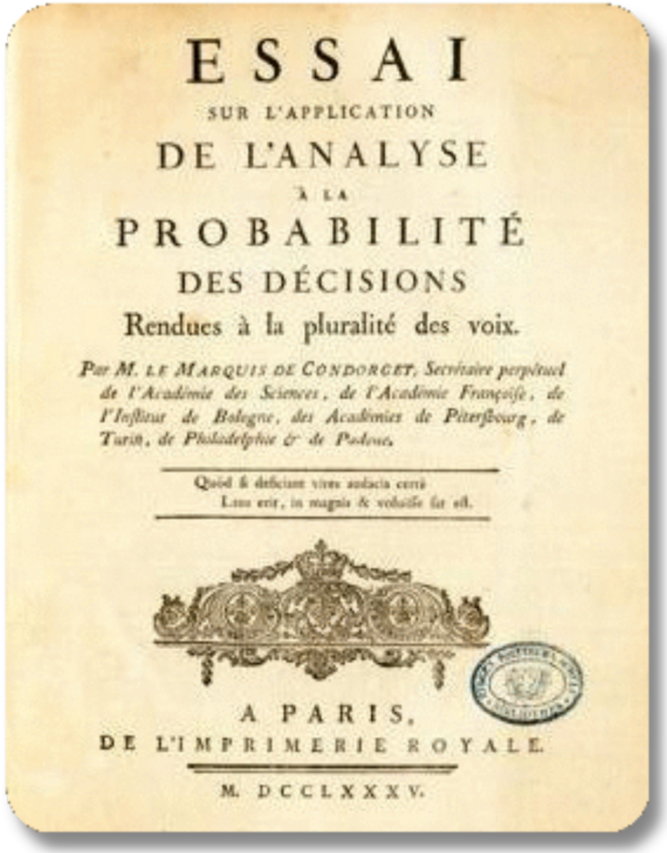
Where did it all start?

CONDORCET

The role of the government is to implement measures that are in the best interest of society.

But how to decide on what outcomes are good?

Democratic procedures can work well.



ESSAI

SUR L'APPLICATION

DE L'ANALYSE

À LA

PROBABILITÉ

DES DÉCISIONS

Rendues à la pluralité des voix.

Par M. LE MARQUIS DE CONDORCET, Secrétaire perpétuel de l'Académie des Sciences, de l'Académie Françoise, de l'Institut de Bologne, des Académies de Pétersbourg, de Turin, de Philadelphie & de Padoue.

Quòd si deficiunt vires audacia certè
Laus erit, in magnis & voluisse sat est.

A PARIS,

DE L'IMPRIMERIE ROYALE.

M. DCCLXXXV.

Marie Jean Antoine Nicolas de Caritat, Marquis of Condorcet (1785). *Essai sur l'application de l'analyse à la probabilité des decisions rendues à la pluralité des voix.*

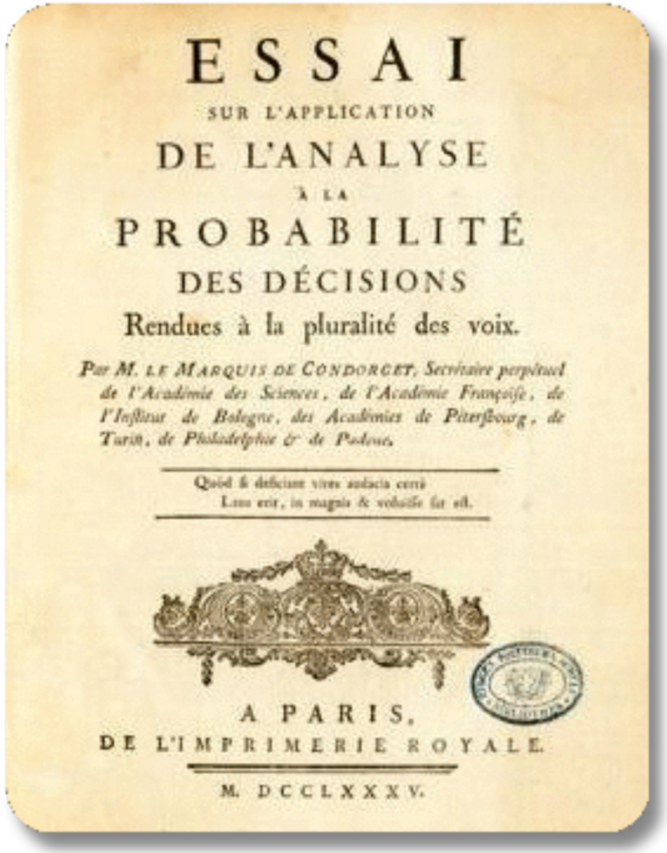
CONDORCET

The role of the government is to implement measures that are in the best interest of society.

But how to decide on what outcomes are good?

Democratic procedures can work well.

And I can show it using this newfangled theory of probabilities.



ESSAI

SUR L'APPLICATION

DE L'ANALYSE

À LA

PROBABILITÉ

DES DÉCISIONS

Rendues à la pluralité des voix.

Par M. LE MARQUIS DE CONDORCET, Secrétaire perpétuel de l'Académie des Sciences, de l'Académie Françoise, de l'Institut de Bologne, des Académies de Pétersbourg, de Turin, de Philadelphie & de Padoue.

Quòd si deficiunt vires audacia certè
Laus erit, in magnis & voluisse sat est.

A PARIS,

DE L'IMPRIMERIE ROYALE.

M. DCCLXXXV.

Marie Jean Antoine Nicolas de Caritat, Marquis of Condorcet (1785). *Essai sur l'application de l'analyse à la probabilité des decisions rendues à la pluralité des voix.*

THE CONDORCET JURY THEOREM

We work in a setting where an odd number of *agents* vote on two issues, one of which is correct.

Each agent has a specific *competence*, which is the probability of voting for the correct alternative.

agents	$N = \{1, \ldots, n\}$
alternatives	$A = \{a, b\}$
correct alternative	a
voter i's vote	v_i
profile of votes	$\boldsymbol{v} = (v_1, \ldots, v_n)$
voter i's competence	$p_i = \mathbb{P}[v_i = a]$
majority opinion	$F_{MAJ}(\boldsymbol{v}) = x$, such that $v_i = x$ for a strict majority of voters

assumed odd

i's guess of the right answer

probability that *i* gets the right answer

we write profiles as words:
(*a*, *a*, *b*, *a*, ...) → *aaba*...

CONDORCET

I want to make some assumptions.

ASSUMPTIONS

Competence

Agents are *competent*, i.e., better than random at being correct:

$$p_i > \frac{1}{2}, \text{ for every agent } i \in N.$$

ASSUMPTIONS

Competence

Agents are *competent*, i.e., better than random at being correct:

$$p_i > \frac{1}{2}, \text{ for every agent } i \in N.$$

Equal Competence

All agents have *the same competence*:

$$p_i = p_j = p, \text{ for any two agents } i, j \in N.$$

ASSUMPTIONS

Competence

Agents are *competent*, i.e., better than random at being correct:

$$p_i > \frac{1}{2}, \text{ for every agent } i \in N.$$

Equal Competence

All agents have *the same competence*:

$$p_i = p_j = p, \text{ for any two agents } i, j \in N.$$

Independence

Agents vote *independently* of each other:

$$\mathbb{P}[v_i = x, v_j = y] = \mathbb{P}[v_i = x] \cdot \mathbb{P}[v_j = y], \text{ for any two agents } i, j \in N.$$

CONDORCET

I claim that under these conditions, the majority tends to get it right!

We want to understand the probability that the majority opinion is correct, that is:

$$\mathbb{P}[F_{MAJ}(v_1, \ldots, v_n) = a].$$

Computing the probability of a correct majority becomes more and more involved as the number of agents grows.

But let's start simple.

ONE VOTER

$$\boldsymbol{v} = (v_1)$$

$$\begin{aligned}\mathbb{P}[F_{\mathsf{MAJ}}(v_1) = a] &= \mathbb{P}[v_1 = a] \\ &= p \\ &> 1/2.\end{aligned}$$

ONE VOTER

$$\boldsymbol{v} = (v_1)$$

$$\begin{aligned}\mathbb{P}[F_{\mathsf{MAJ}}(v_1) = a] &= \mathbb{P}[v_1 = a] \\ &= p \\ &> 1/2.\end{aligned}$$

by the Equal Competence assumption

by the Competence assumption

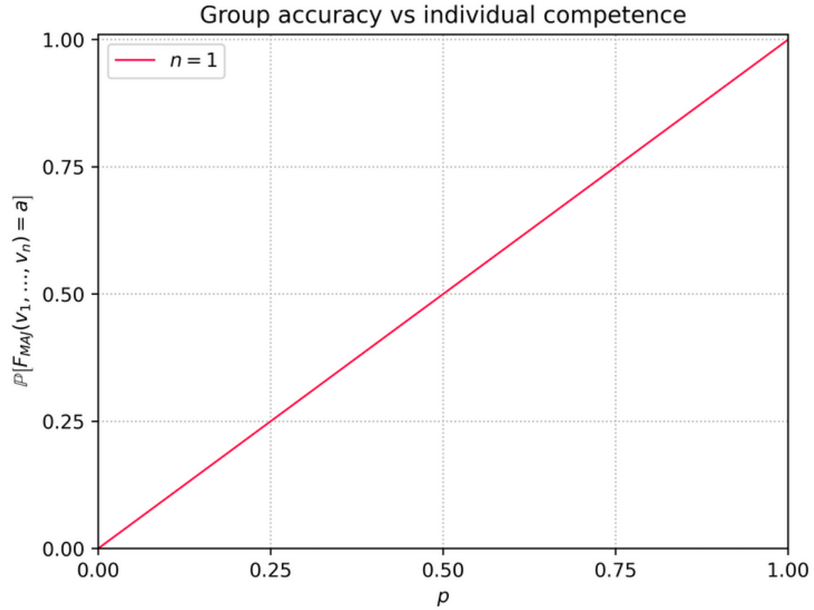
ONE VOTER

$$\boldsymbol{v} = (v_1)$$

$$\begin{aligned}\mathbb{P}[F_{MAJ}(v_1) = a] &= \mathbb{P}[v_1 = a] \\ &= p \\ &> 1/2.\end{aligned}$$

by the Equal Competence assumption

by the Competence assumption



Note

As p grows, so does group accuracy.

in this case, trivially

$$\boldsymbol{v} = (v_1, v_2)$$

TWO VOTERS

$$\boldsymbol{v} = (v_1, v_2)$$

TWO VOTERS

Oh yeah, we're not looking at this case.

THREE VOTERS

$$\boldsymbol{v} = (v_1, v_2, v_3)$$

$$\begin{aligned}
\mathbb{P}[F_{\mathsf{MAJ}}(\boldsymbol{v}) = a] &= \mathbb{P}[\boldsymbol{v} \in \{aab, aba, baa, aaa\}] \\
&= \mathbb{P}[\boldsymbol{v} = aab] + \mathbb{P}[\boldsymbol{v} = aba] + \mathbb{P}[\boldsymbol{v} = baa] + \mathbb{P}[\boldsymbol{v} = aaa] \\
&= \mathbb{P}[v_1 = a] \cdot \mathbb{P}[v_2 = a] \cdot \mathbb{P}[v_3 = b] + \\
&\quad \mathbb{P}[v_1 = a] \cdot \mathbb{P}[v_2 = b] \cdot \mathbb{P}[v_3 = a] + \\
&\quad \mathbb{P}[v_1 = b] \cdot \mathbb{P}[v_2 = a] \cdot \mathbb{P}[v_3 = a] + \\
&\quad \mathbb{P}[v_1 = a] \cdot \mathbb{P}[v_2 = a] \cdot \mathbb{P}[v_3 = a] + \\
&= p \cdot p \cdot (1-p) + p \cdot (1-p) \cdot p + (1-p) \cdot p \cdot p + p \cdot p \cdot p \\
&= 3p^2(1-p) + p^3 \\
&> p. \ (?)
\end{aligned}$$

by the Independence assumption

by the Equal Competence assumption

by the Competence assumption

THREE VOTERS

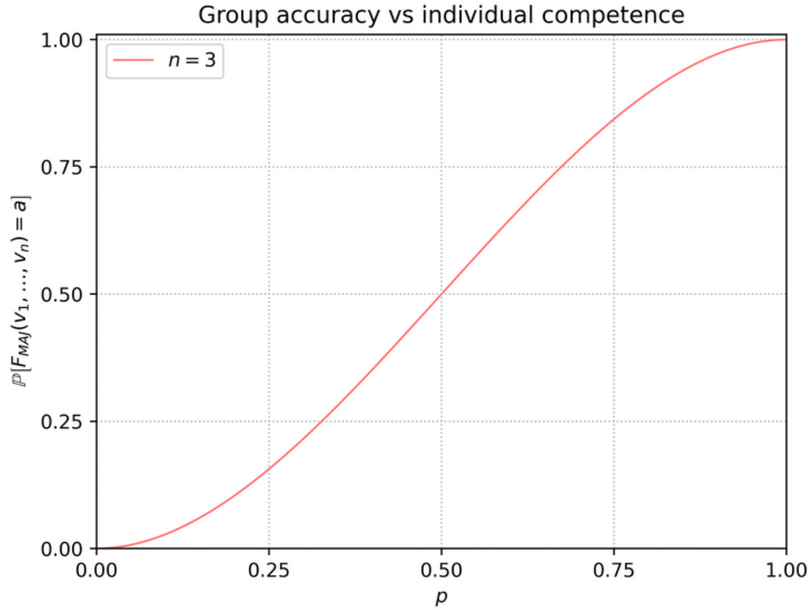
$$\boldsymbol{v} = (v_1, v_2, v_3)$$

$$\begin{aligned}
\mathbb{P}[F_{MAJ}(\boldsymbol{v}) = a] &= \mathbb{P}[\boldsymbol{v} \in \{aab, aba, baa, aaa\}] \\
&= \mathbb{P}[\boldsymbol{v} = aab] + \mathbb{P}[\boldsymbol{v} = aba] + \mathbb{P}[\boldsymbol{v} = baa] + \mathbb{P}[\boldsymbol{v} = aaa] \\
&= \mathbb{P}[v_1 = a] \cdot \mathbb{P}[v_2 = a] \cdot \mathbb{P}[v_3 = b] + \\
&\quad \mathbb{P}[v_1 = a] \cdot \mathbb{P}[v_2 = b] \cdot \mathbb{P}[v_3 = a] + \\
&\quad \mathbb{P}[v_1 = b] \cdot \mathbb{P}[v_2 = a] \cdot \mathbb{P}[v_3 = a] + \\
&\quad \mathbb{P}[v_1 = a] \cdot \mathbb{P}[v_2 = a] \cdot \mathbb{P}[v_3 = a] + \\
&= p \cdot p \cdot (1-p) + p \cdot (1-p) \cdot p + (1-p) \cdot p \cdot p + p \cdot p \cdot p \\
&= 3p^2(1-p) + p^3 \\
&> p. \ (?)
\end{aligned}$$

by the Independence assumption

by the Equal Competence assumption

by the Competence assumption



Note

As p grows, so does group accuracy.

THREE VOTERS

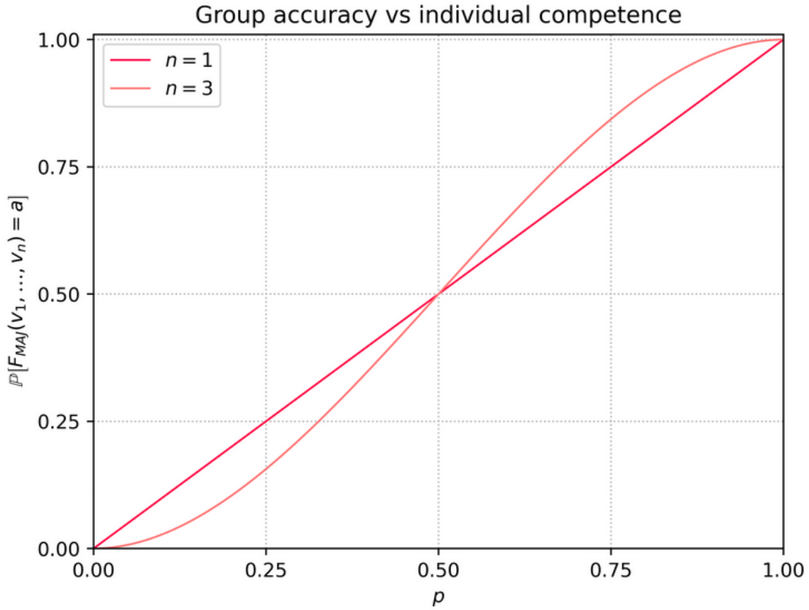
$$\boldsymbol{v} = (v_1, v_2, v_3)$$

$$\begin{aligned}
\mathbb{P}[F_{MAJ}(\boldsymbol{v}) = a] &= \mathbb{P}[\boldsymbol{v} \in \{aab, aba, baa, aaa\}] \\
&= \mathbb{P}[\boldsymbol{v} = aab] + \mathbb{P}[\boldsymbol{v} = aba] + \mathbb{P}[\boldsymbol{v} = baa] + \mathbb{P}[\boldsymbol{v} = aaa] \\
&= \mathbb{P}[v_1 = a] \cdot \mathbb{P}[v_2 = a] \cdot \mathbb{P}[v_3 = b] + \\
&\quad \mathbb{P}[v_1 = a] \cdot \mathbb{P}[v_2 = b] \cdot \mathbb{P}[v_3 = a] + \\
&\quad \mathbb{P}[v_1 = b] \cdot \mathbb{P}[v_2 = a] \cdot \mathbb{P}[v_3 = a] + \\
&\quad \mathbb{P}[v_1 = a] \cdot \mathbb{P}[v_2 = a] \cdot \mathbb{P}[v_3 = a] + \\
&= p \cdot p \cdot (1-p) + p \cdot (1-p) \cdot p + (1-p) \cdot p \cdot p + p \cdot p \cdot p \\
&= 3p^2(1-p) + p^3 \\
&> p. \ (?)
\end{aligned}$$

by the Independence assumption

by the Equal Competence assumption

by the Competence assumption



Note

As p grows, so does group accuracy.

A group of size 3 is more likely to be correct than a group of size 1.

$$\boldsymbol{v} = (v_1, v_2, v_3, v_4, v_5)$$

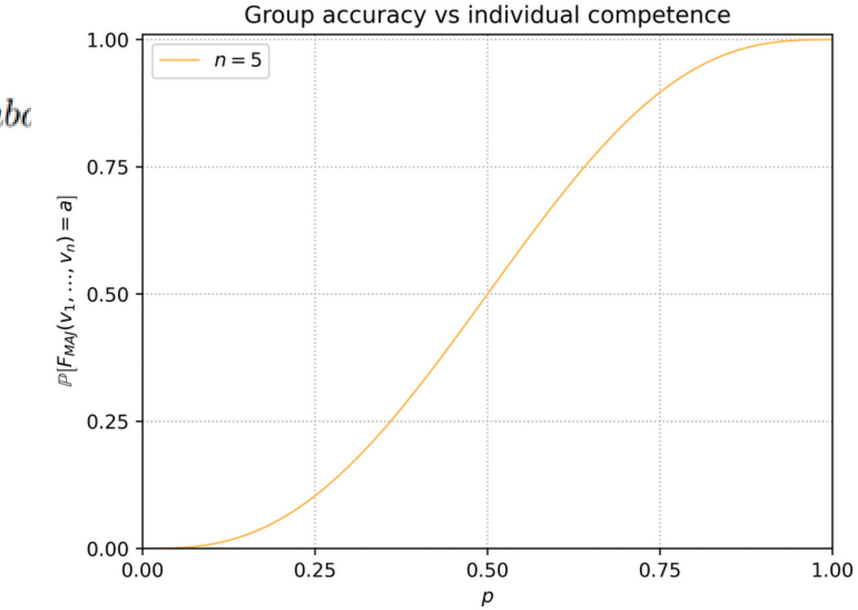
FIVE VOTERS

$$\begin{aligned}
\mathbb{P}[F_{\mathsf{MAJ}}(\boldsymbol{v}) = a] &= \mathbb{P}[\boldsymbol{v} \in \{aaabb, aabab, abaab, abbba, aabba, ababa, baaba, abbaa, babaa, bbaaa\}] + \\
&\quad \mathbb{P}[\boldsymbol{v} \in \{aaaab, aaaba, aabaa, abaaa, baaaa\}] + \\
&\quad \mathbb{P}[\boldsymbol{v} \in \{aaaaa\}] \\
&\quad \cdots \\
&= 10p^3(1-p)^2 + 5p^4(1-p) + p^5 \\
&= \binom{5}{3}p^3(1-p)^2 + \binom{5}{4}p^4(1-p)^1 + \binom{5}{5}p^5.
\end{aligned}$$

$$\boldsymbol{v} = (v_1, v_2, v_3, v_4, v_5)$$

FIVE VOTERS

$$\begin{aligned}
\mathbb{P}[F_{MAJ}(\boldsymbol{v}) = a] &= \mathbb{P}[\boldsymbol{v} \in \{aaabb, aabab, abaab, abbba, aabba, ababa, baabc \\
&\quad \mathbb{P}[\boldsymbol{v} \in \{aaaab, aaaba, aabaa, abaaa, baaaa\}]+ \\
&\quad \mathbb{P}[\boldsymbol{v} \in \{aaaaa\}] \\
&\quad \dots \\
&= 10p^3(1-p)^2 + 5p^4(1-p) + p^5 \\
&= \binom{5}{3}p^3(1-p)^2 + \binom{5}{4}p^4(1-p)^1 + \binom{5}{5}p^5.
\end{aligned}$$



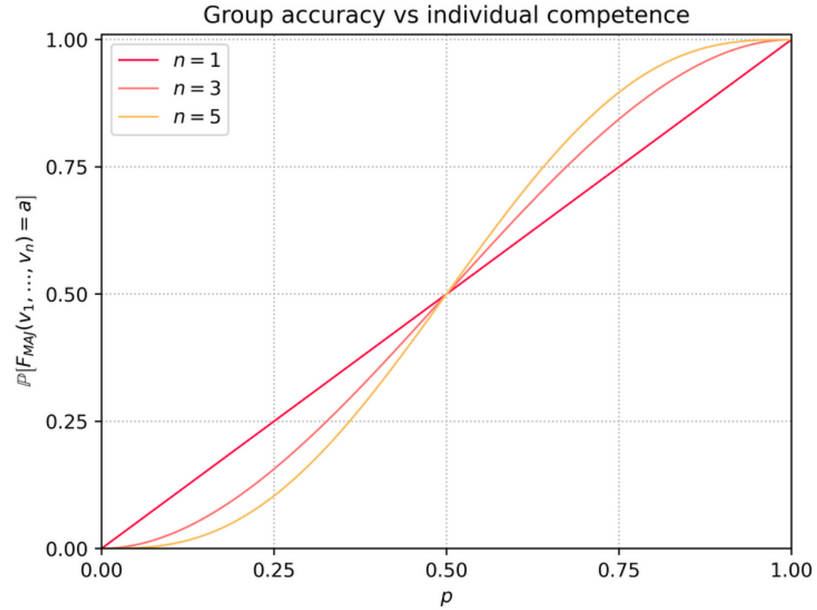
Note

Again: as *p* grows, so does group accuracy.

$$\boldsymbol{v} = (v_1, v_2, v_3, v_4, v_5)$$

FIVE VOTERS

$$\begin{aligned}
\mathbb{P}[F_{MAJ}(\boldsymbol{v}) = a] &= \mathbb{P}[\boldsymbol{v} \in \{aaabb, aabab, abaab, abbba, aabba, ababa, baabc \\
&\quad \mathbb{P}[\boldsymbol{v} \in \{aaaab, aaaba, aabaa, abaaa, baaaa\}]+ \\
&\quad \mathbb{P}[\boldsymbol{v} \in \{aaaaa\}] \\
&\quad \dots \\
&= 10p^3(1-p)^2 + 5p^4(1-p) + p^5 \\
&= \binom{5}{3}p^3(1-p)^2 + \binom{5}{4}p^4(1-p)^1 + \binom{5}{5}p^5.
\end{aligned}$$



Note

Again: as *p* grows, so does group accuracy.

A group of size 5 is more likely to be correct than a group of size 3.

ANY ODD NUMBER OF VOTERS

$$\boldsymbol{v} = (v_1, \ldots, v_n)$$

$$\begin{aligned}
\mathbb{P}[F_{MAJ}(\boldsymbol{v}) = a] &= \mathbb{P}\left[\boldsymbol{v} \text{ such that } > n/2 \text{ agents vote for } a\right] \\
&= \mathbb{P}\left[\boldsymbol{v} \text{ s.t. } \lfloor n/2 \rfloor + 1 \text{ agents vote for } a\right] + \cdots + \mathbb{P}\left[\boldsymbol{v} \text{ s.t. } n \text{ agents vote for } a\right] \\
&= \left(\mathbb{P}\left[\boldsymbol{v} = \underbrace{a \ldots a}_{\lfloor n/2 \rfloor + 1} b \ldots b \right] + \cdots + \mathbb{P}\left[\boldsymbol{v} = b \ldots b \underbrace{a \ldots a}_{\lfloor n/2 \rfloor + 1} \right] \right) + \ldots \\
&\qquad + \mathbb{P}[\boldsymbol{v} = \underbrace{a \ldots a}_{n}] \\
&= \binom{n}{\lfloor n/2 \rfloor + 1} p^{\lfloor n/2 \rfloor + 1} (1-p)^{n - (\lfloor n/2 \rfloor + 1)} + \cdots + \binom{n}{n-1} p^{n-1} (1-p)^1 + \binom{n}{n} p^n \\
&= \sum_{i = \lfloor n/2 \rfloor + 1}^{n} \binom{n}{i} p^i (1-p)^{n-i}.
\end{aligned}$$

CONDORCET

By the croissants of my ancestors: I claim that the larger the group, the more accurate it is!

And that in the limit, groups are infallible.

Provided there are no dumdums and people make their minds up independently.

THEOREM (THE CONDORCET JURY THEOREM, OR CJT)

If all agents have the same, larger than ½, competence and vote independently of each other, then, for odd *n*, it holds that:

- the accuracy of the group improves as its size grows:

$$\mathbb{P}[F_{MAJ}(v_1, \ldots, v_{n+2}) = a] > \mathbb{P}[F_{MAJ}(v_1, \ldots, v_n) = a]$$

- the accuracy of the group is better than that of any of its members:

$$\mathbb{P}[F_{MAJ}(v_1, \ldots, v_n) = a] > \mathbb{P}[v_i = a], \text{ for } n \geq 3$$

- the accuracy of the group approaches 1 asymptotically:

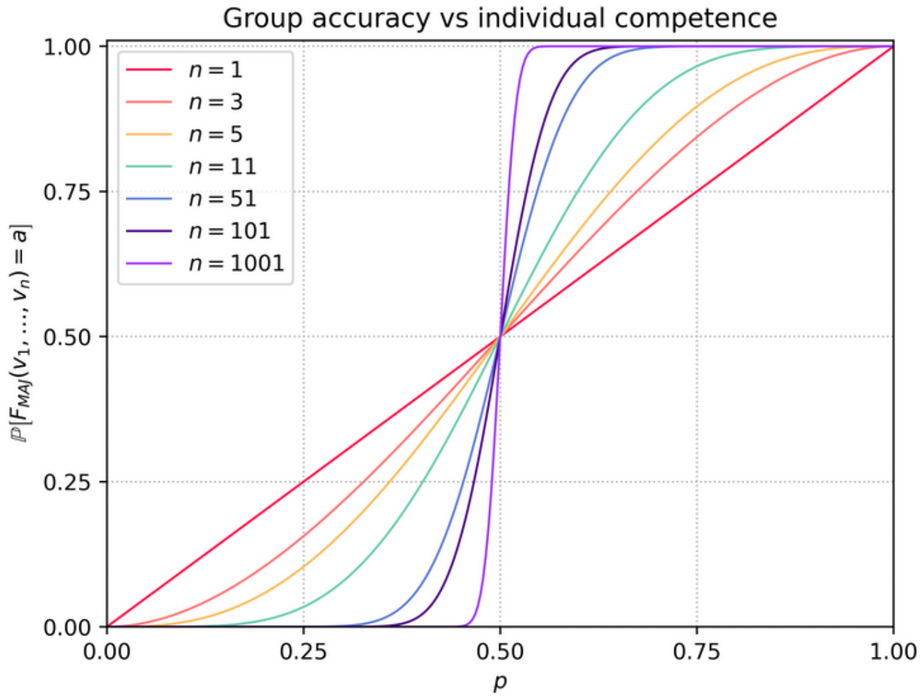
$$\lim_{n \to \infty} \mathbb{P}[F_{MAJ}(v_1, \ldots, v_n) = a] = 1$$

CONDORCET

Groups are better than their members.

The larger the group, the better.

In the limit, performance is perfect.



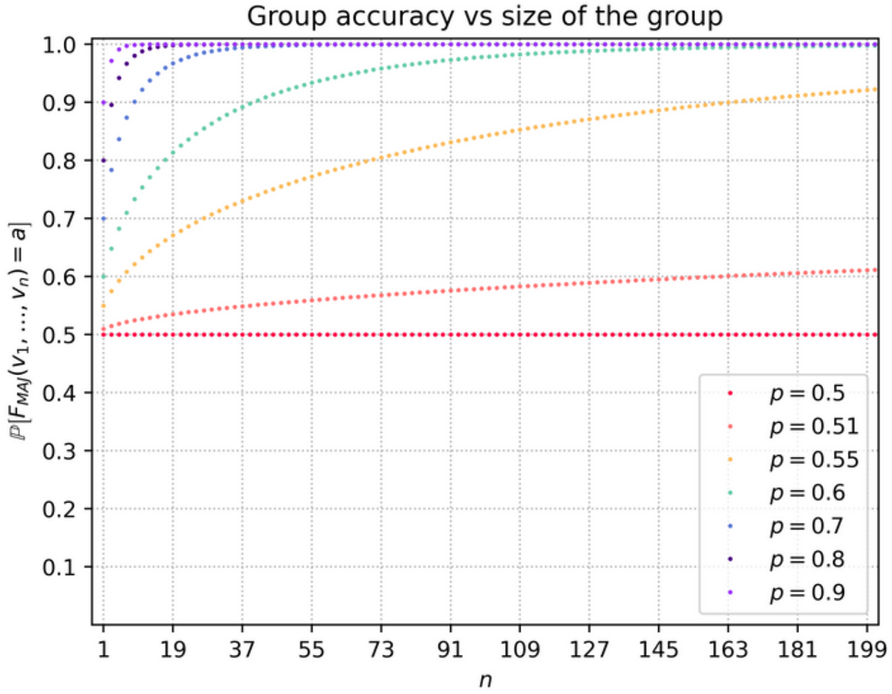
CONDORCET

Groups are better than their members.

The larger the group, the better.

In the limit, performance is perfect.

And performance grows fast with the size of the group.



CONDORCET

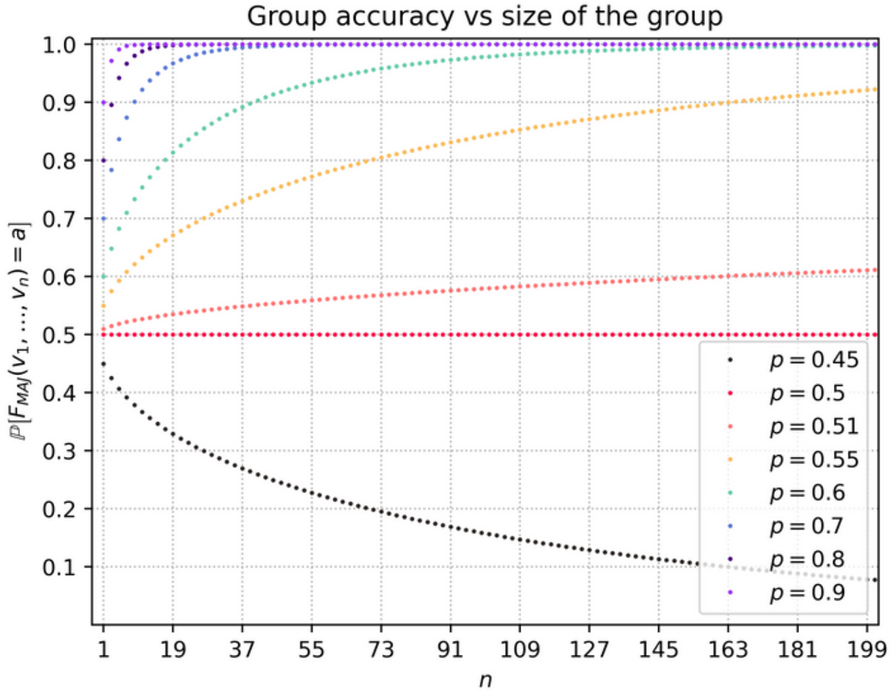
Groups are better than their members.

The larger the group, the better.

In the limit, performance is perfect.

And performance grows fast with the size of the group.

Provided $p > 0.5$.



RELAXING THE ASSUMPTIONS OF THE CONDORCET JURY THEOREM

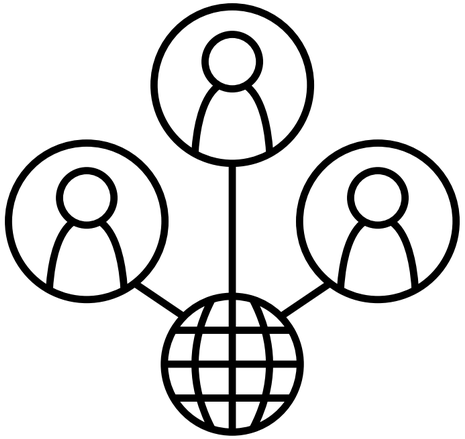
Relaxing independence.

CONDORCET

No point in denying it: the CJT has a major blindspot.

Independent voter beliefs.

Out there people interact and are exposed to common information sources, e.g., mass media.



KRISHNA K. LADHA

Introducing correlation between voters can make the optimistic results go away.

HÉLÈNE LANDEMORE

At the same time, there is more and more evidence that certain forms of communication, e.g., deliberation, are good for decision making.

Ladha, K. K. (1992). The Condorcet Jury Theorem, Free Speech, and Correlated Votes. *American Journal of Political Science*, 36(3), 617–634

Landemore, H. (2013). *Democratic Reason: Politics, Collective Intelligence, and the Rule of the Many*. Princeton University Press.

Landemore, H. (2020). *Open Democracy: Reinventing Popular Rule for the Twenty-First Century*. Princeton University Press.

Relaxing the competence assumption.

CONDORCET

What would be a reason for p to be below 0.5?

CONDORCET

What would be a reason for p to be below 0.5?

DANIEL KAHNEMAN

Biases!

You thought it was Brooklyn, didn't you?

This bridge connects Manhattan to what other New York borough?

- ☐ Brooklyn
- ☑ Queens

Kahneman, D. (2013). *Thinking, Fast and Slow*. Farrar, Straus and Giroux.

CONDORCET

What would be a reason for p to be below 0.5?

DANIEL KAHNEMAN

Biases!

You thought it was Brooklyn, didn't you?

BRYAN CAPLAN

Most people can't be relied on take good decisions.

JASON BRENNAN

Especially when it comes to political issues.

HÉLÈNE LANDEMORE

Let's not exaggerate.

Kahneman, D. (2013). *Thinking, Fast and Slow*. Farrar, Straus and Giroux.
Caplan, B. (2011). *The Myth of the Rational Voter: Why Democracies Choose Bad Policies*. Princeton University Press.
Brennan, J. (2017). *Against Democracy*. Princeton University Press.
Landemore, H. (2013). *Democratic Reason: Politics, Collective Intelligence, and the Rule of the Many*. Princeton University Press.

This bridge connects Manhattan to what other New York borough?

- ☐ Brooklyn
- ☑ Queens

And what does this p even mean, anyway?

CONDORCET

Can we rate people's accuracies, especially if predicting rare, or unique, events?

GLENN BRIEN

Sure!

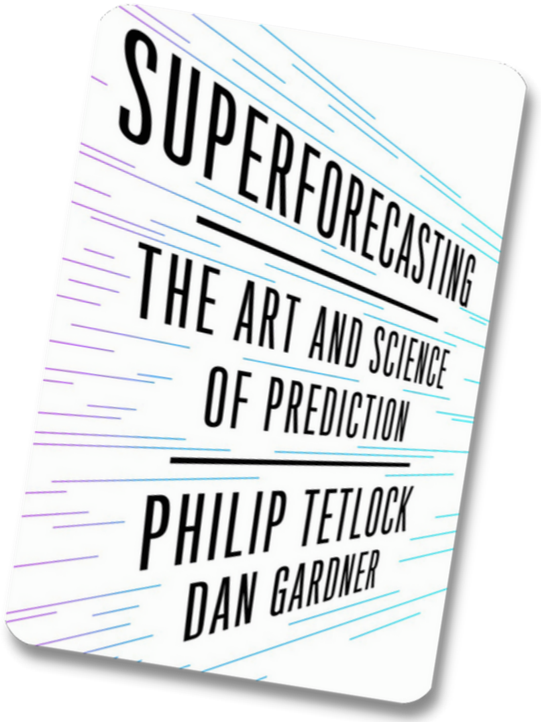
Check out the Brier score.

CONDORCET

Even so: is it realistic to assume that $p > 0.5$?

PHILIP E. TETLOCK

Some people seem to manage it: superforecasters.



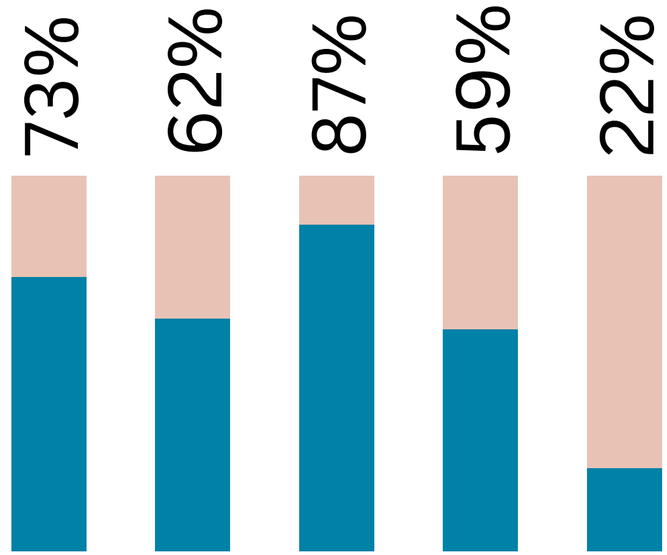
Tetlock, P. E., & Gardner, D. (2016). *Superforecasting: The Art and Science of Prediction*. Random House.

And what if agents don't all have the same competence p?

BERNARD GROFMAN

It's not so clear if the conclusions of the CJT still hold.

It gets kind of complicated...



Grofman, B., Owen, G., & Feld, S. L. (1983). Thirteen theorems in search of the truth. *Theory and Decision*, 15(3), 261–278.
Owen, G., Grofman, B., & Feld, S. L. (1989). Proving a distribution-free generalization of the Condorcet Jury Theorem. *Mathematical Social Sciences*, 17(1), 1–16.

Wrapping up...

CONDORCET

The Condorcet Jury Theorem is a cornerstone of the idea that groups can be wise.

But it also feels like a fragile result, based on unrealistic assumptions.

Can we find better results, for modern-day challenges?