TRUST

DAVID MILLARD

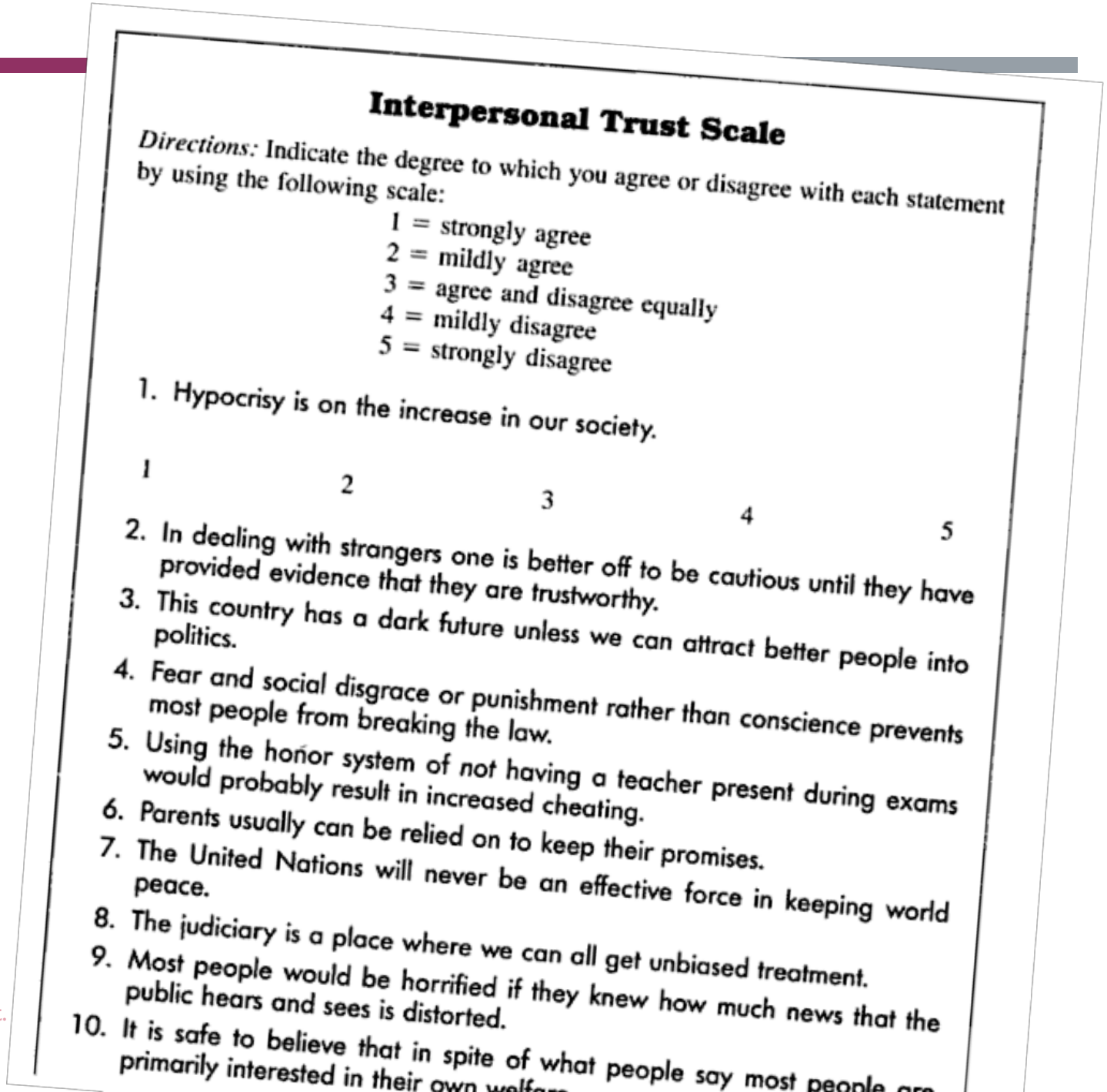
DEM@SOTON.AC.UK

THIS COUNTRY HAS A DARK FUTURE UNLESS WE CAN ATTRACT BETTER PEOPLE INTO POLITICS

IT IS SAFE TO BELIEVE THAT IN SPITE OF WHAT PEOPLE SAY MOST PEOPLE ARE PRIMARILY INTERESTED IN THEIR OWN WELFARE

ROTTER'S INTERPERSONAL TRUST SCALE

- The ITS measures an individual's general tendency to trust different groups of people
  - E.g. teachers, parents, politicians, physicians, classmates, and friends
- The survey measures the individual's "general optimism regarding the society"



Interpersonal Trust Scale

Directions: Indicate the degree to which you agree or disagree with each statement by using the following scale:

1 = strongly agree
2 = mildly agree
3 = agree and disagree equally
4 = mildly disagree
5 = strongly disagree

1. Hypocrisy is on the increase in our society.

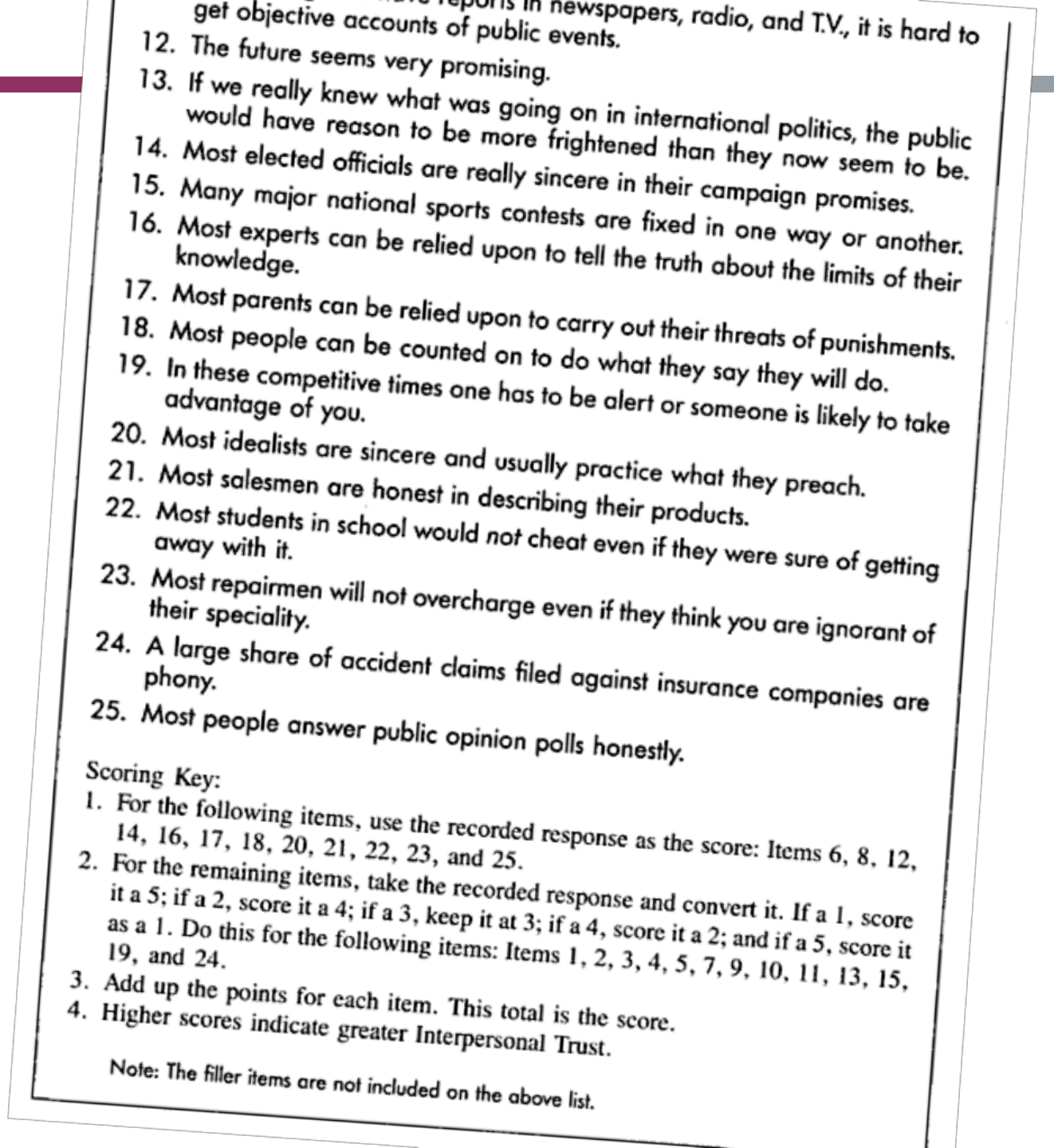
1 2 3 4 5

2. In dealing with strangers one is better off to be cautious until they have provided evidence that they are trustworthy.
3. This country has a dark future unless we can attract better people into politics.
4. Fear and social disgrace or punishment rather than conscience prevents most people from breaking the law.
5. Using the honor system of *not* having a teacher present during exams would probably result in increased cheating.
6. Parents usually can be relied on to keep their promises.
7. The United Nations will never be an effective force in keeping world peace.
8. The judiciary is a place where we can all get unbiased treatment.
9. Most people would be horrified if they knew how much news that the public hears and sees is distorted.
10. It is safe to believe that in spite of what people say most people are primarily interested in their own welfare.

Rotter, J. B. 1967. A new scale for the measurement of inter- personal trust. Journal of Personality. 35; 615-665

ROTTER'S INTERPERSONAL TRUST SCALE

- Rotter summarised his 13 years of findings in 1980
- Subjects that are high in interpersonal trust tend to be:
  - more trustworthy
  - better liked
  - happier
- Subjects low in interpersonal trust tend to exhibit untrustworthy behavior (self-reported cheating and lying)



reports in newspapers, radio, and T.V., it is hard to get objective accounts of public events.

12. The future seems very promising.
13. If we really knew what was going on in international politics, the public would have reason to be more frightened than they now seem to be.
14. Most elected officials are really sincere in their campaign promises.
15. Many major national sports contests are fixed in one way or another.
16. Most experts can be relied upon to tell the truth about the limits of their knowledge.
17. Most parents can be relied upon to carry out their threats of punishments.
18. Most people can be counted on to do what they say they will do.
19. In these competitive times one has to be alert or someone is likely to take advantage of you.
20. Most idealists are sincere and usually practice what they preach.
21. Most salesmen are honest in describing their products.
22. Most students in school would *not* cheat even if they were sure of getting away with it.
23. Most repairmen will not overcharge even if they think you are ignorant of their speciality.
24. A large share of accident claims filed against insurance companies are phony.
25. Most people answer public opinion polls honestly.

Scoring Key:

1. For the following items, use the recorded response as the score: Items 6, 8, 12, 14, 16, 17, 18, 20, 21, 22, 23, and 25.
2. For the remaining items, take the recorded response and convert it. If a 1, score it a 5; if a 2, score it a 4; if a 3, keep it at 3; if a 4, score it a 2; and if a 5, score it as a 1. Do this for the following items: Items 1, 2, 3, 4, 5, 7, 9, 10, 11, 13, 15, 19, and 24.
3. Add up the points for each item. This total is the score.
4. Higher scores indicate greater Interpersonal Trust.

Note: The filler items are not included on the above list.

Rotter, J. B. 1967. A new scale for the measurement of inter- personal trust. Journal of Personality. 35; 615-665

Trust

"Trust is a psychological state comprising **the intention to accept vulnerability** based upon positive expectations of the intentions or behavior of another."

- Roussau et al. (1998)

Rousseau et al. Not so different after all: a cross-discipline view of trust. Academy of Management Review (1998) vol. 23 pp. 393-404

Trust

"Trust is a psychological state comprising **the intention to accept vulnerability** based upon positive expectations of the intentions or behavior of another."

- Roussau et al. (1998)

Risk

"Risk is the perceived probability of loss, as interpreted by a decision maker "

Interdependence

"where the interests of one party cannot be achieved without reliance upon another."

Rousseau et al. Not so different after all: a cross-discipline view of trust. Academy of Management Review (1998) vol. 23 pp. 393-404

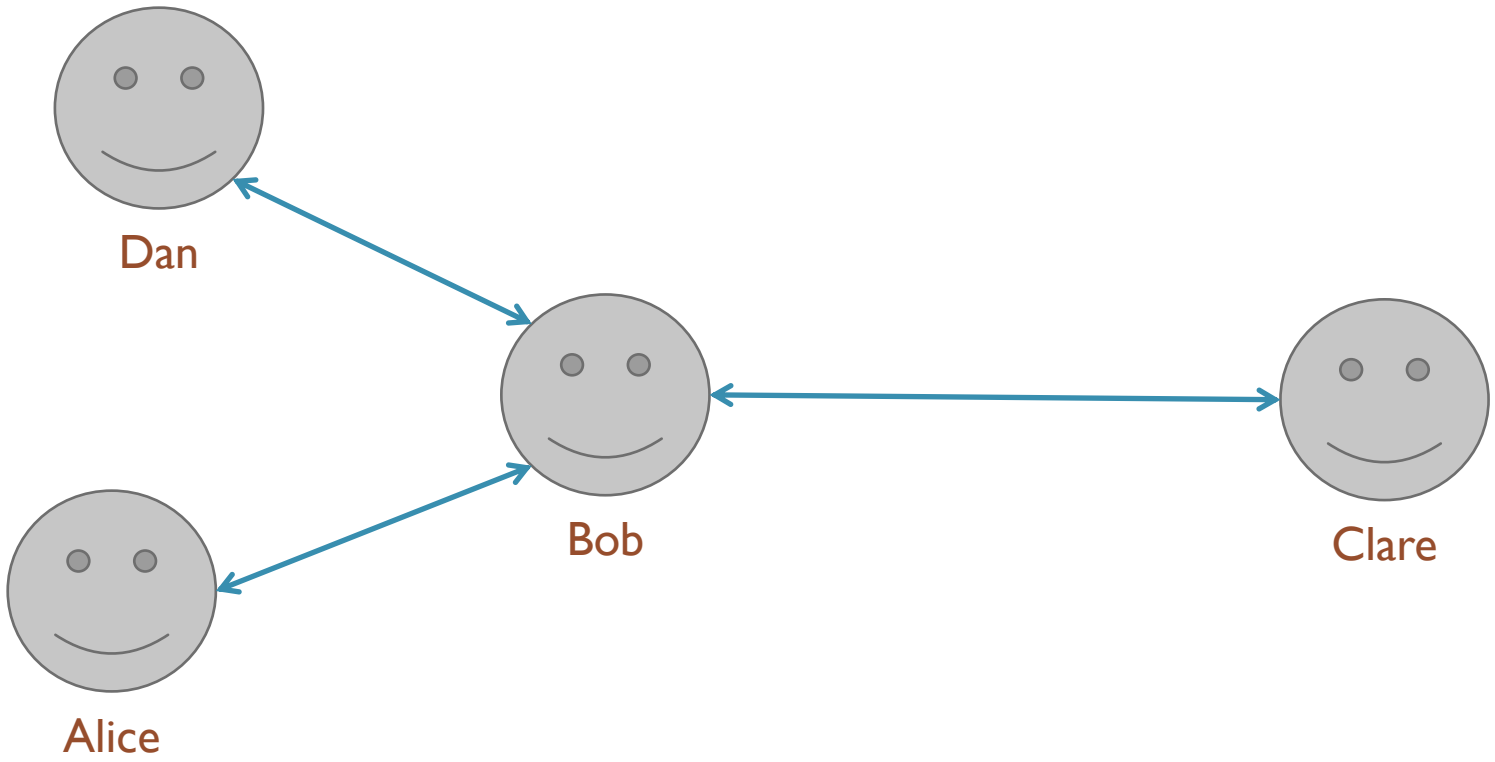
PART I

CALCULATING INTERPERSONAL TRUST ONLINE

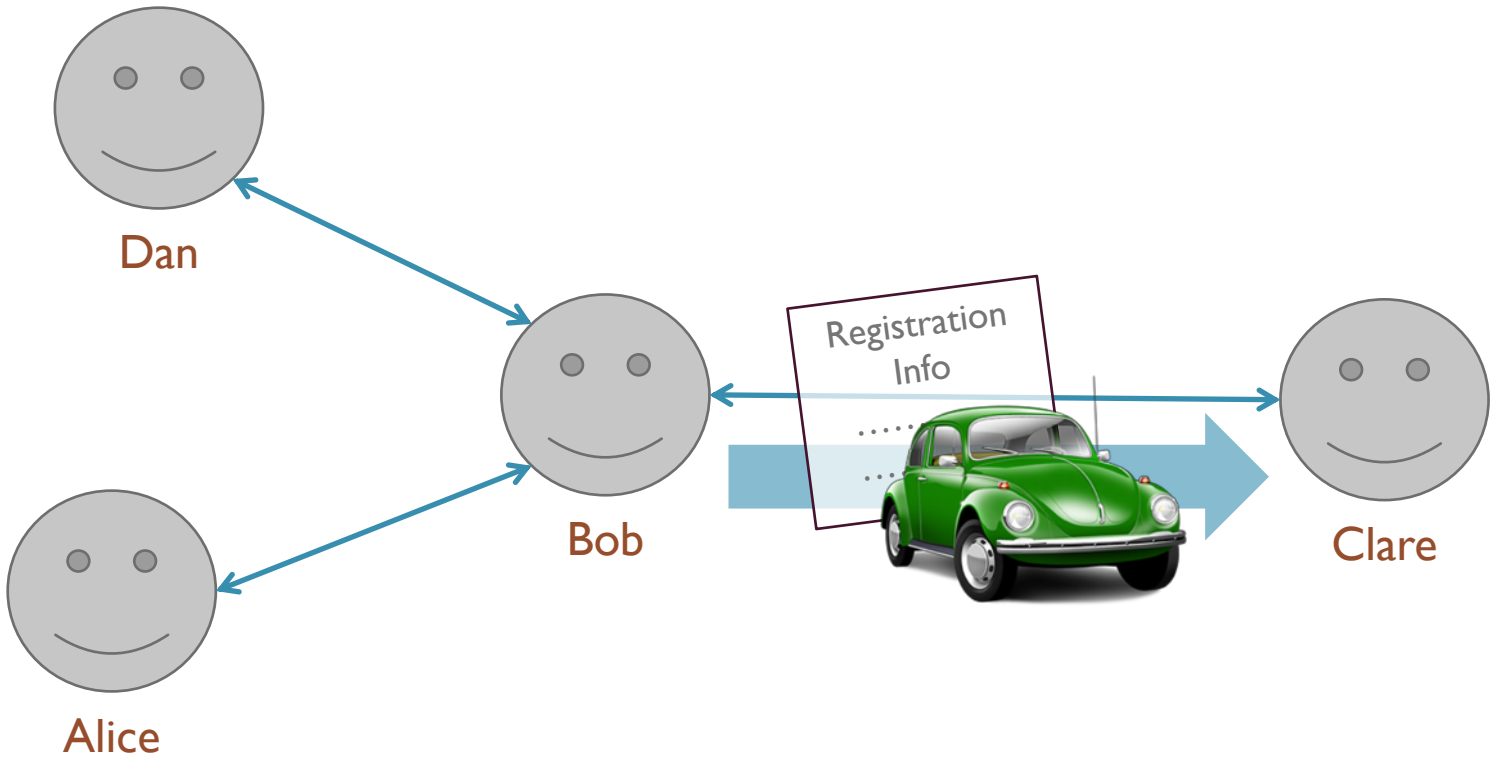
HOW MIGHT WE CALCULATE TRUST?

1. Policy-Based Trust
2. Provenance-Based Trust
3. Reputation-Based Trust

HOW MIGHT WE CALCULATE TRUST?



HOW MIGHT WE CALCULATE TRUST?



Bob is selling Clare a car that used to belong to Alice

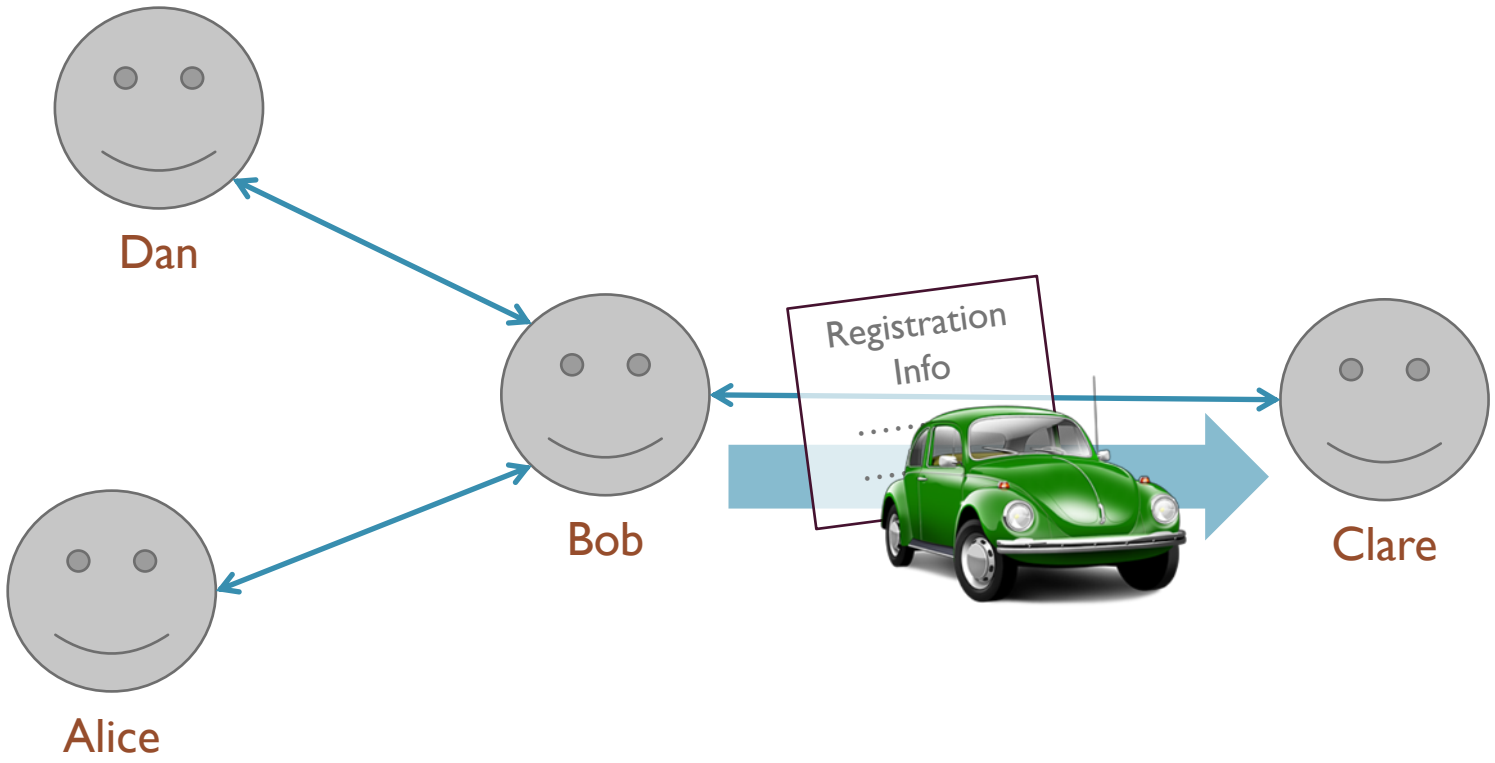
Dan has bought stuff from Bob before

Can Clare trust the info Bob gives her about the car?

HOW MIGHT WE CALCULATE TRUST?

1. Policy-Based Trust
2. Provenance-Based Trust
3. Reputation-Based Trust

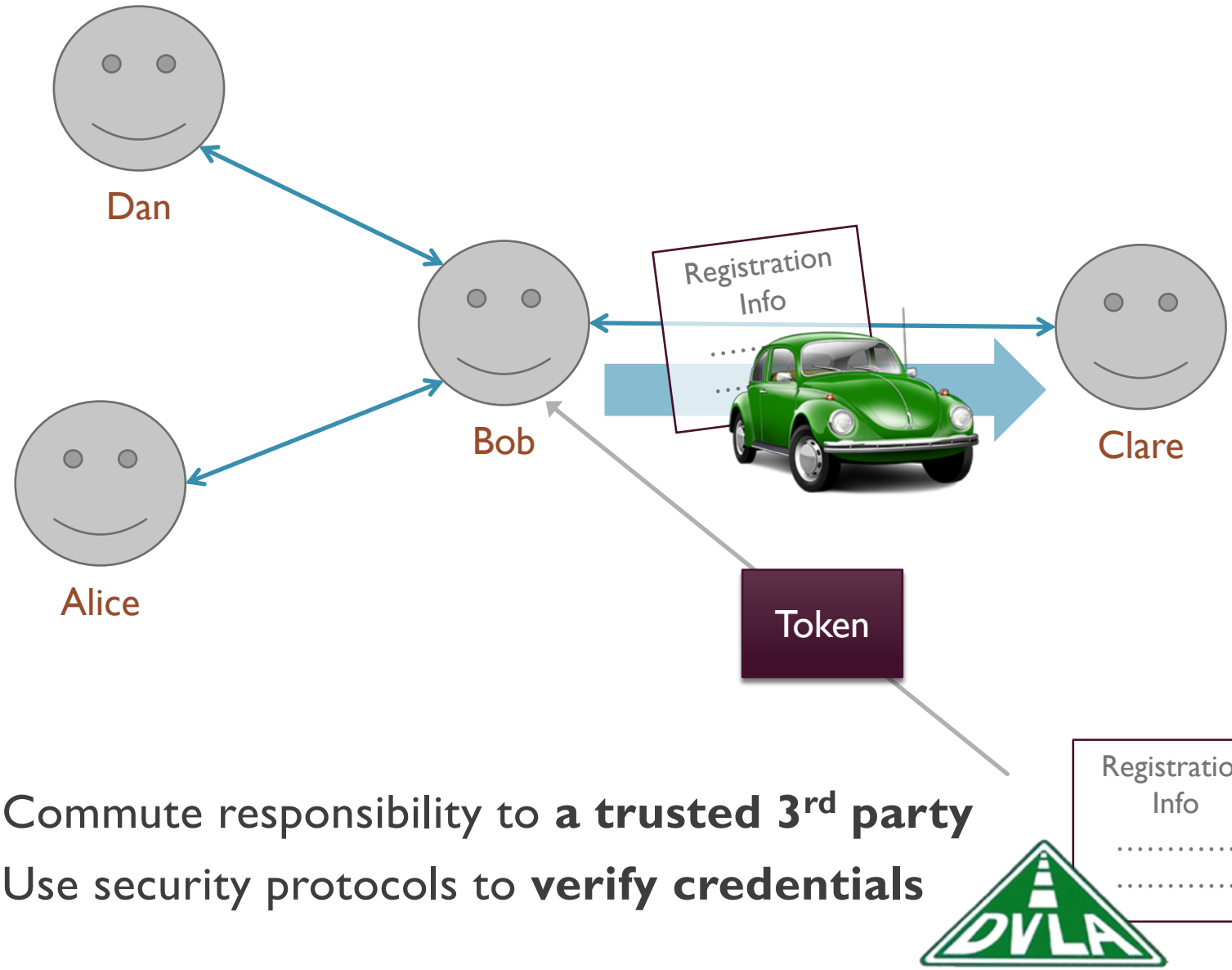
HOW MIGHT WE CALCULATE TRUST?



Commute responsibility to **a trusted 3rd party**

Use security protocols to **verify credentials**

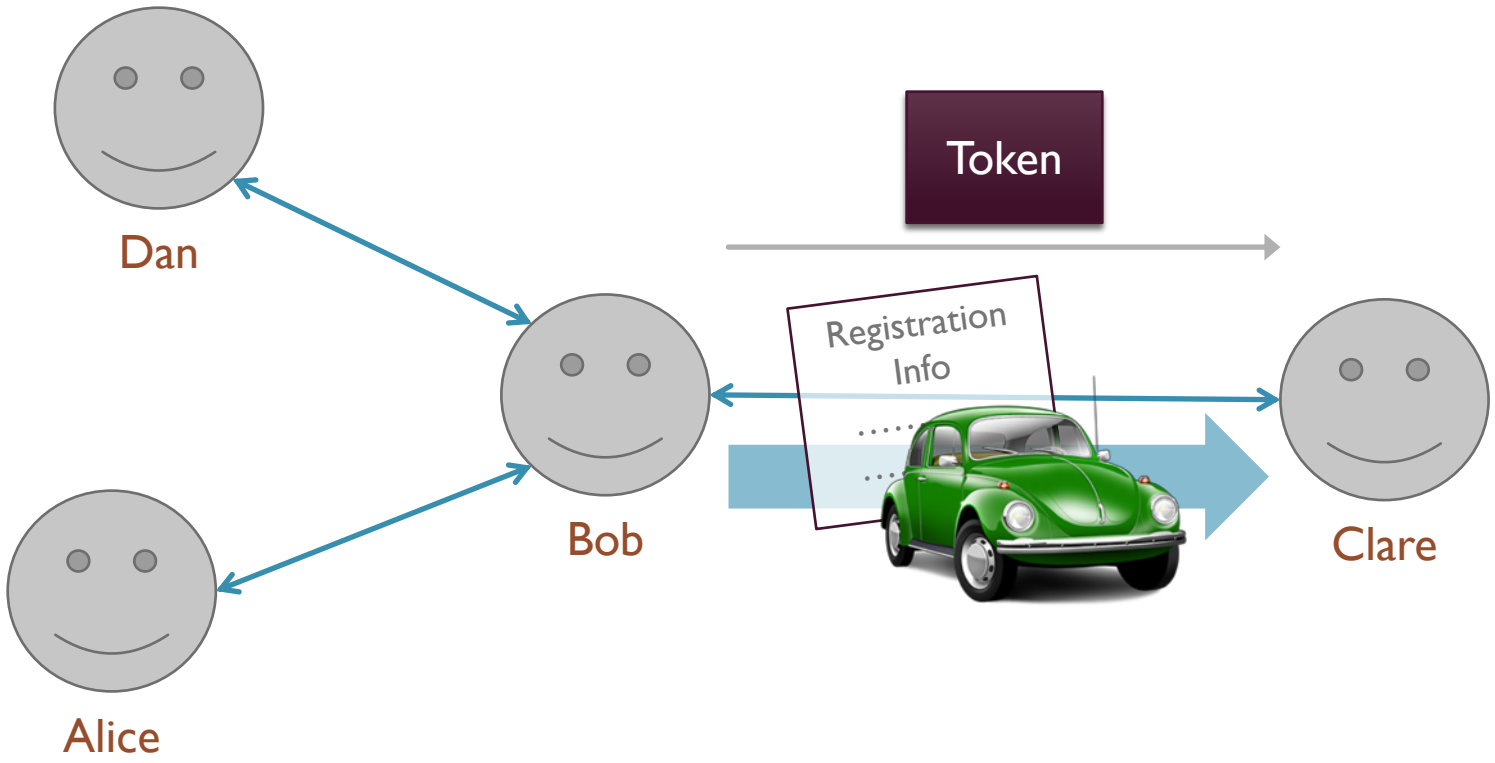
HOW MIGHT WE CALCULATE TRUST?



Commute responsibility to **a trusted 3rd party**

Use security protocols to **verify credentials**

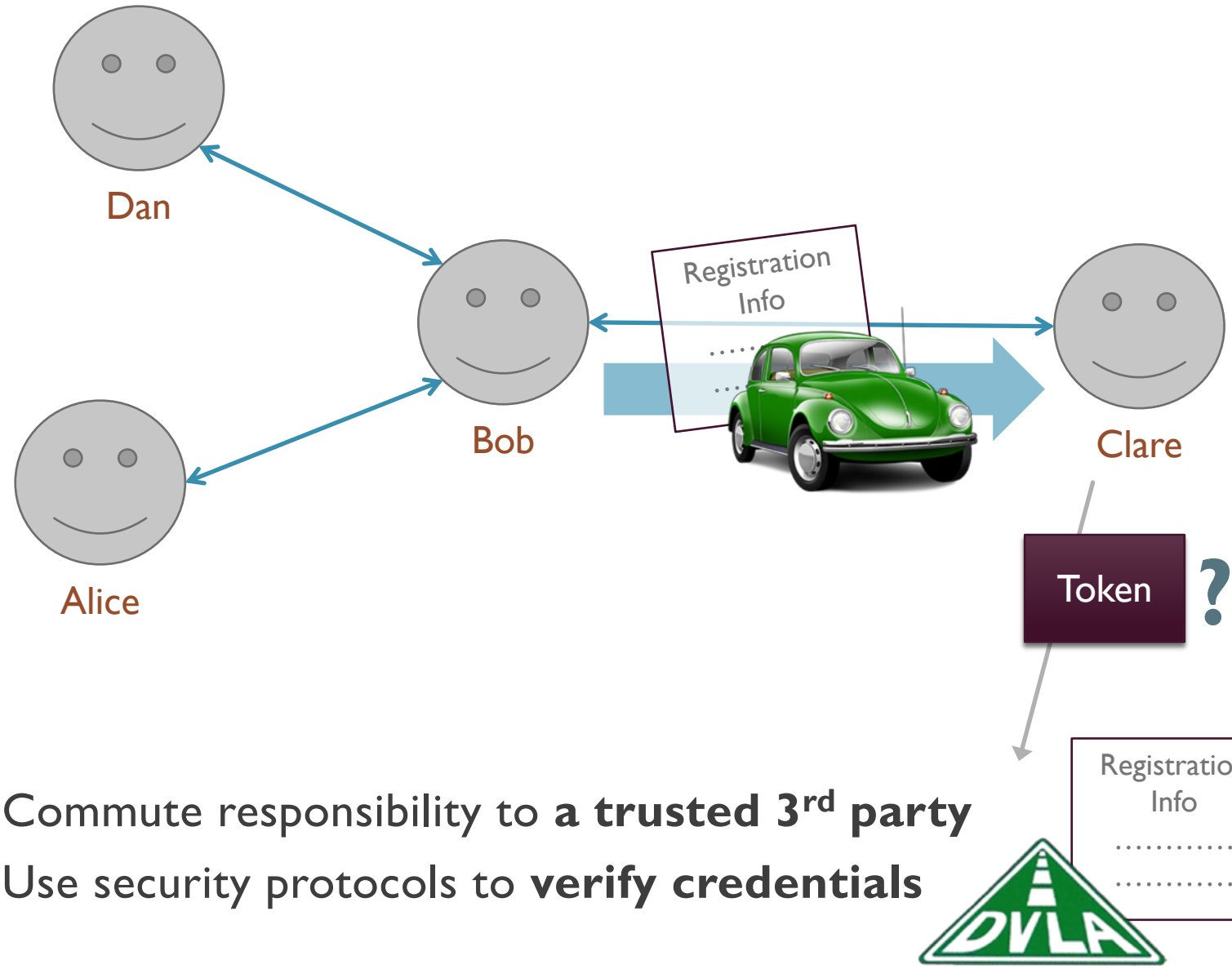
HOW MIGHT WE CALCULATE TRUST?



Commute responsibility to **a trusted 3rd party**

Use security protocols to **verify credentials**

HOW MIGHT WE CALCULATE TRUST?



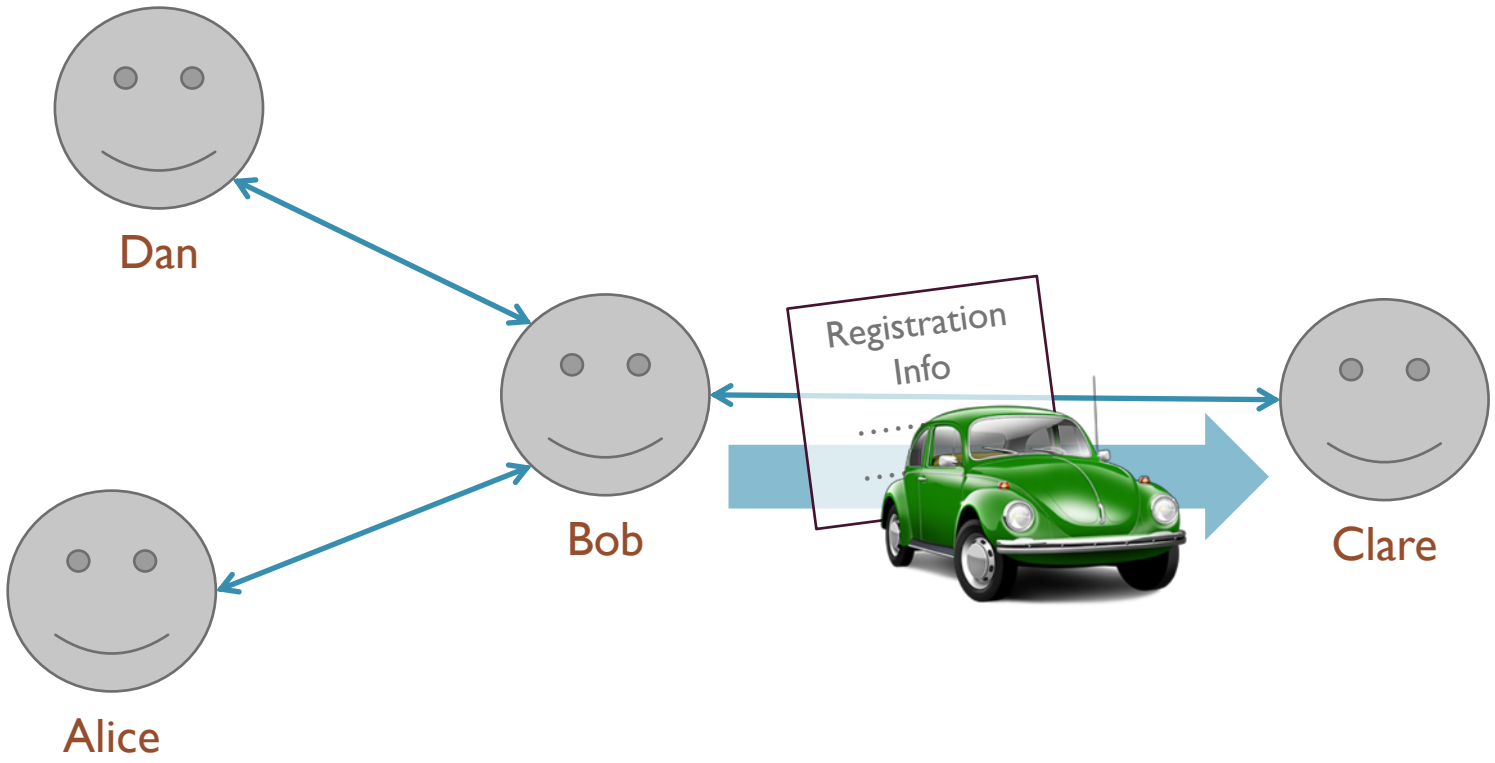
Commute responsibility to **a trusted 3rd party**

Use security protocols to **verify credentials**

HOW MIGHT WE CALCULATE TRUST?

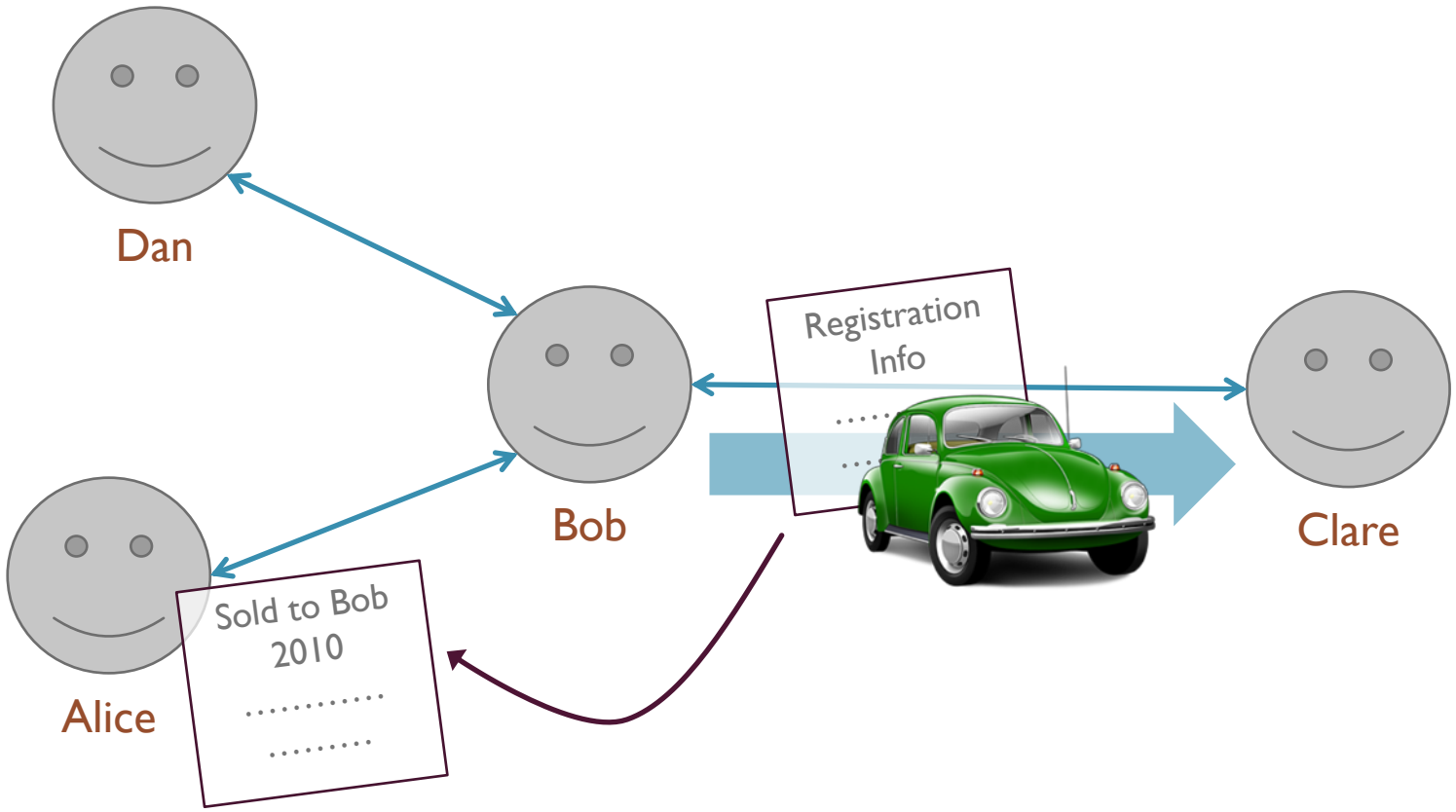
1. Policy-Based Trust
2. Provenance-Based Trust
3. Reputation-Based Trust

HOW MIGHT WE CALCULATE TRUST?



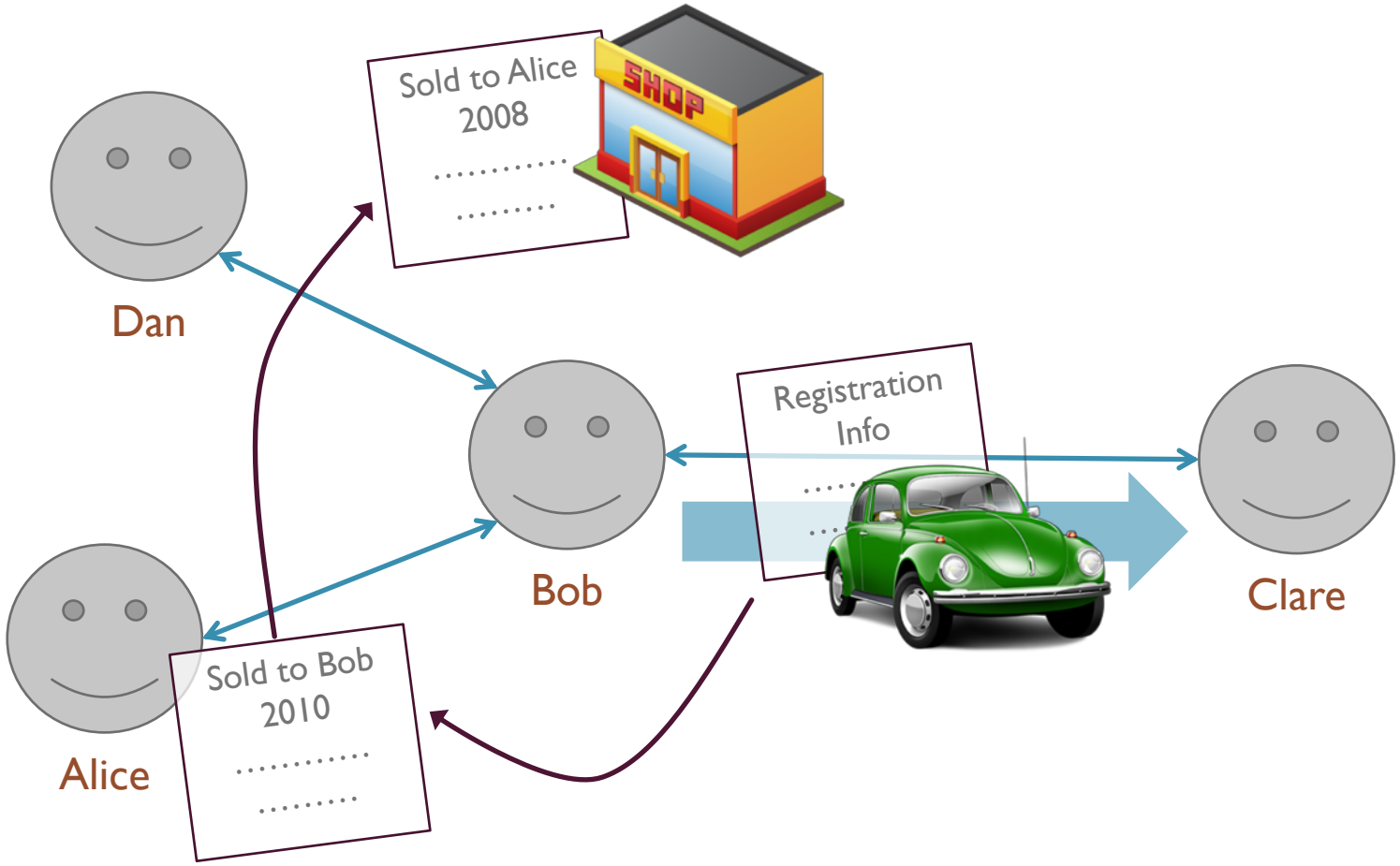
The **provenance** of a piece of data is the process that led to that piece of data.

HOW MIGHT WE CALCULATE TRUST?



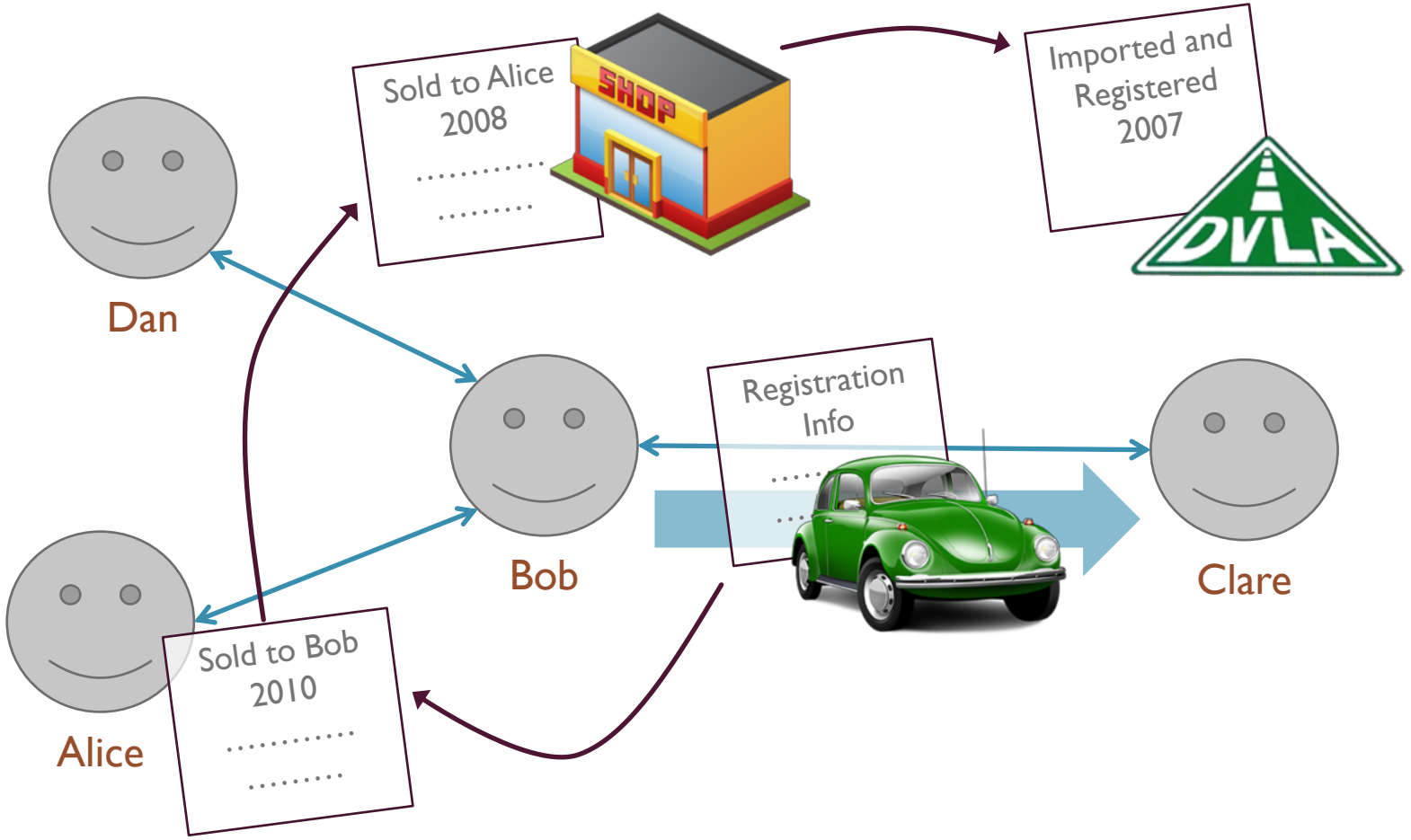
The **provenance** of a piece of data is the process that led to that piece of data.

HOW MIGHT WE CALCULATE TRUST?



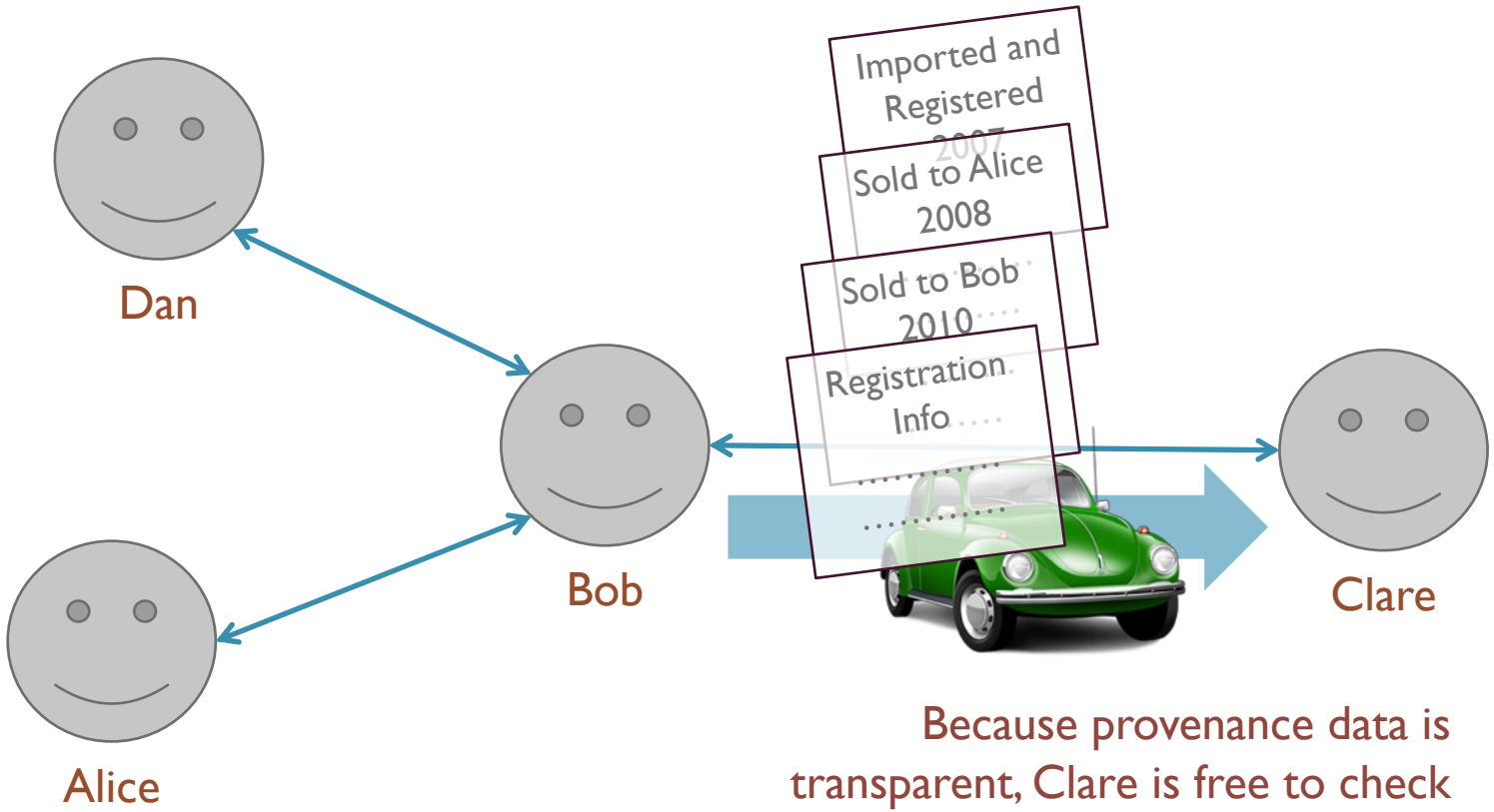
The **provenance** of a piece of data is the process that led to that piece of data.

HOW MIGHT WE CALCULATE TRUST?



The **provenance** of a piece of data is the process that led to that piece of data.

HOW MIGHT WE CALCULATE TRUST?



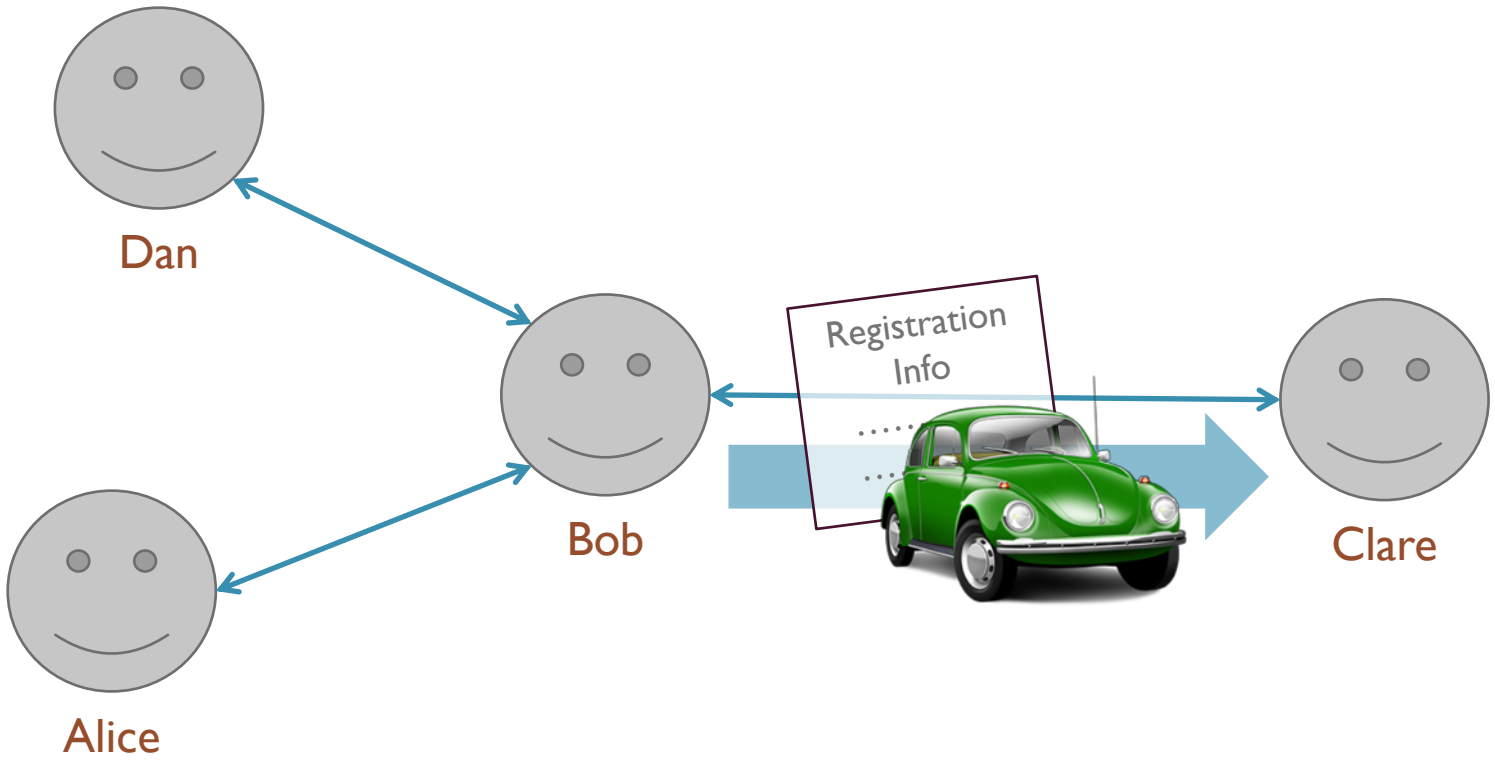
Because provenance data is transparent, Clare is free to check with anyone named in the process.

The **provenance** of a piece of data is the process that led to that piece of data.

HOW MIGHT WE CALCULATE TRUST?

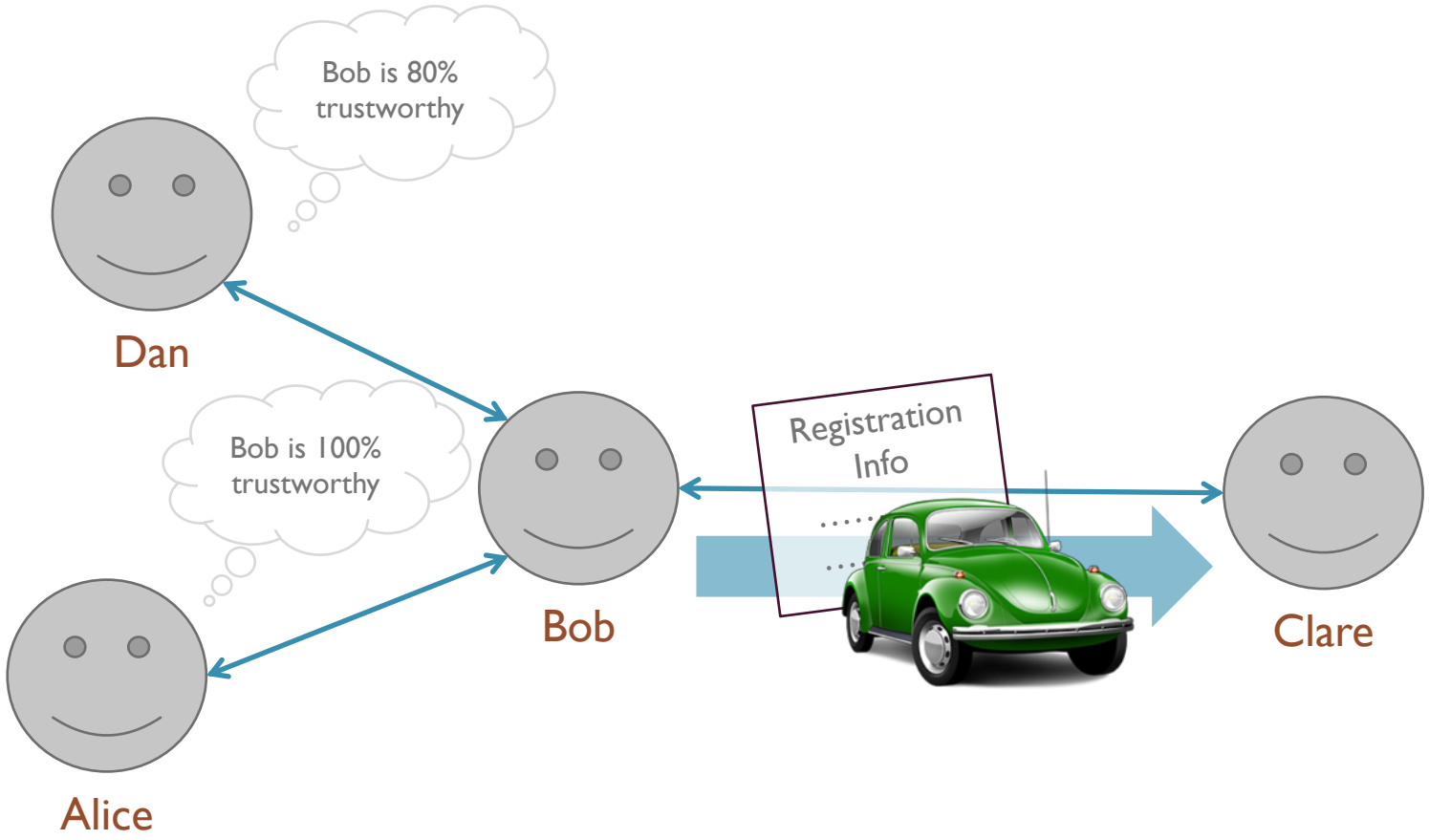
1. Policy-Based Trust
2. Provenance-Based Trust
3. Reputation-Based Trust

HOW MIGHT WE CALCULATE TRUST?



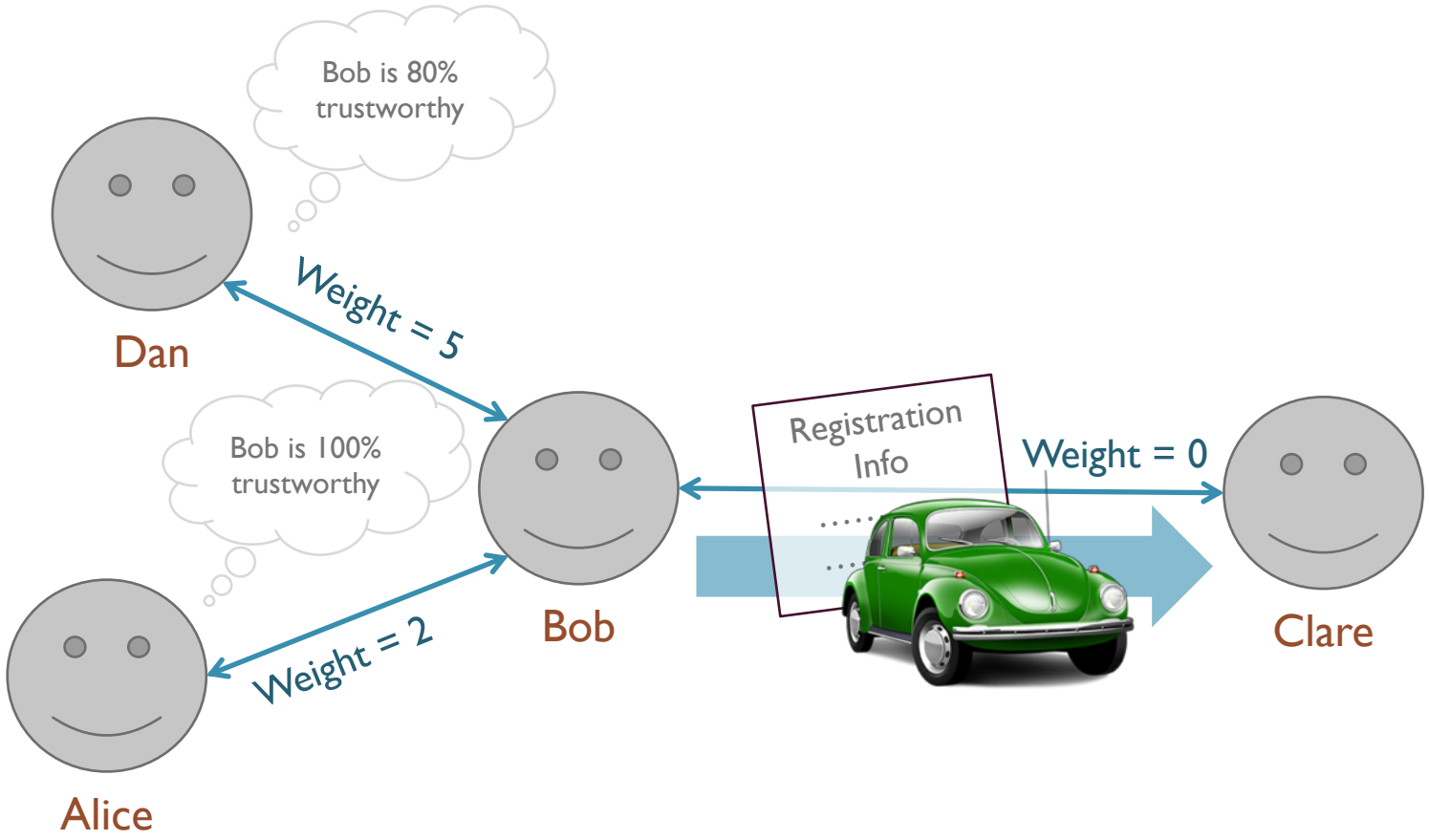
Trust based on personal experience or the experiences of others – their reputation

HOW MIGHT WE CALCULATE TRUST?



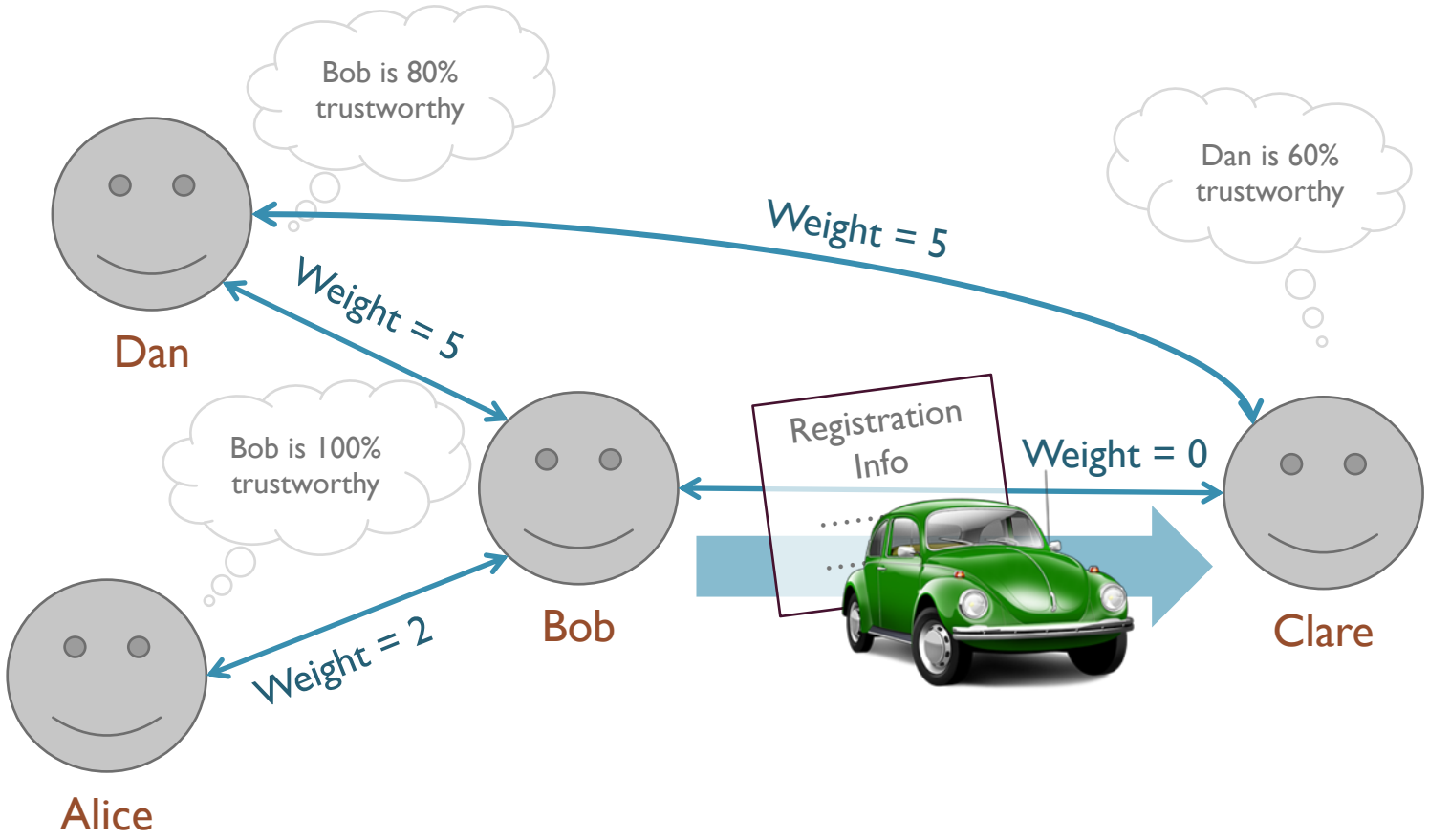
Trust based on personal experience or the experiences of others – their reputation

HOW MIGHT WE CALCULATE TRUST?



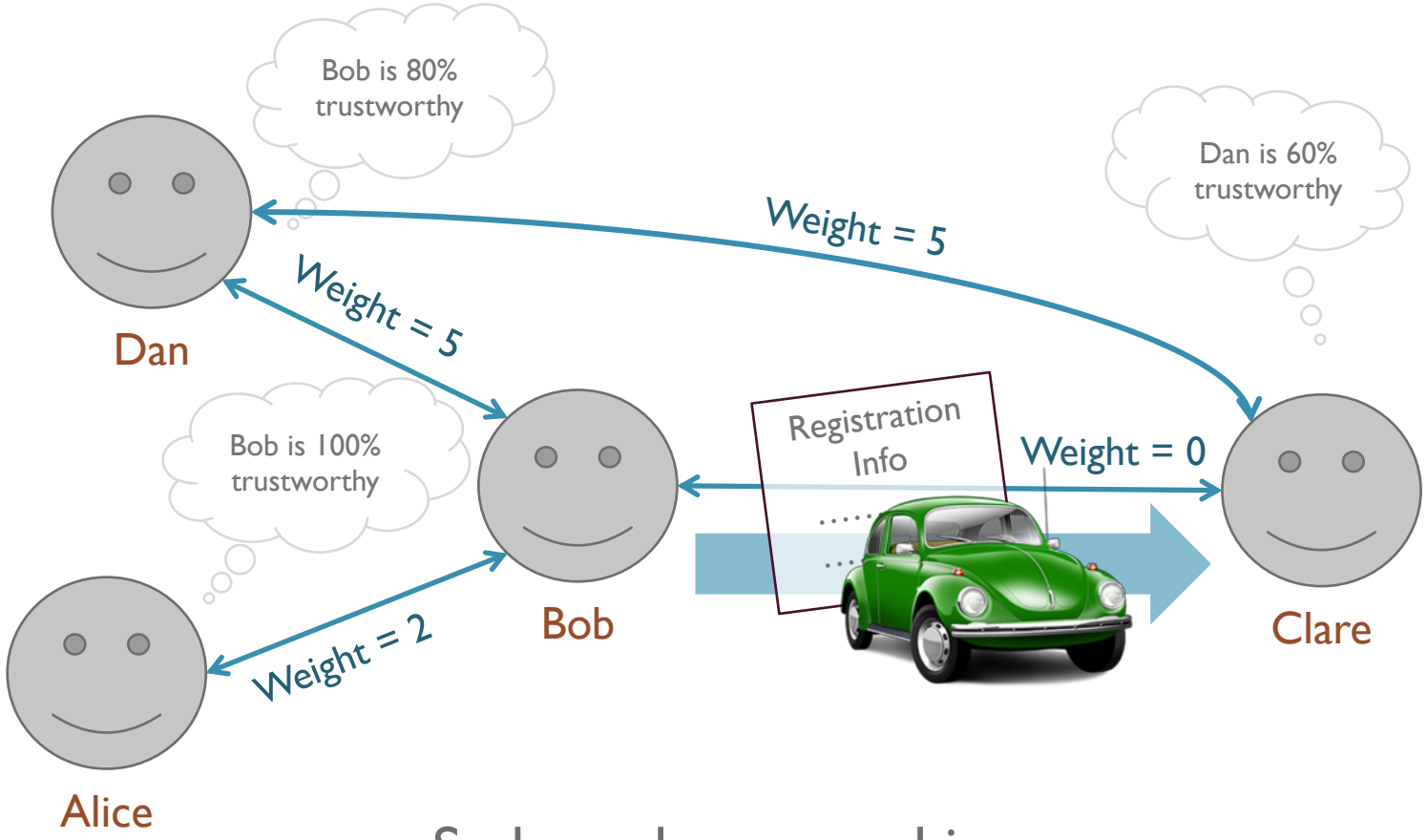
Trust based on personal experience or the experiences of others – their reputation

HOW MIGHT WE CALCULATE TRUST?



Trust based on personal experience or the experiences of others – their reputation

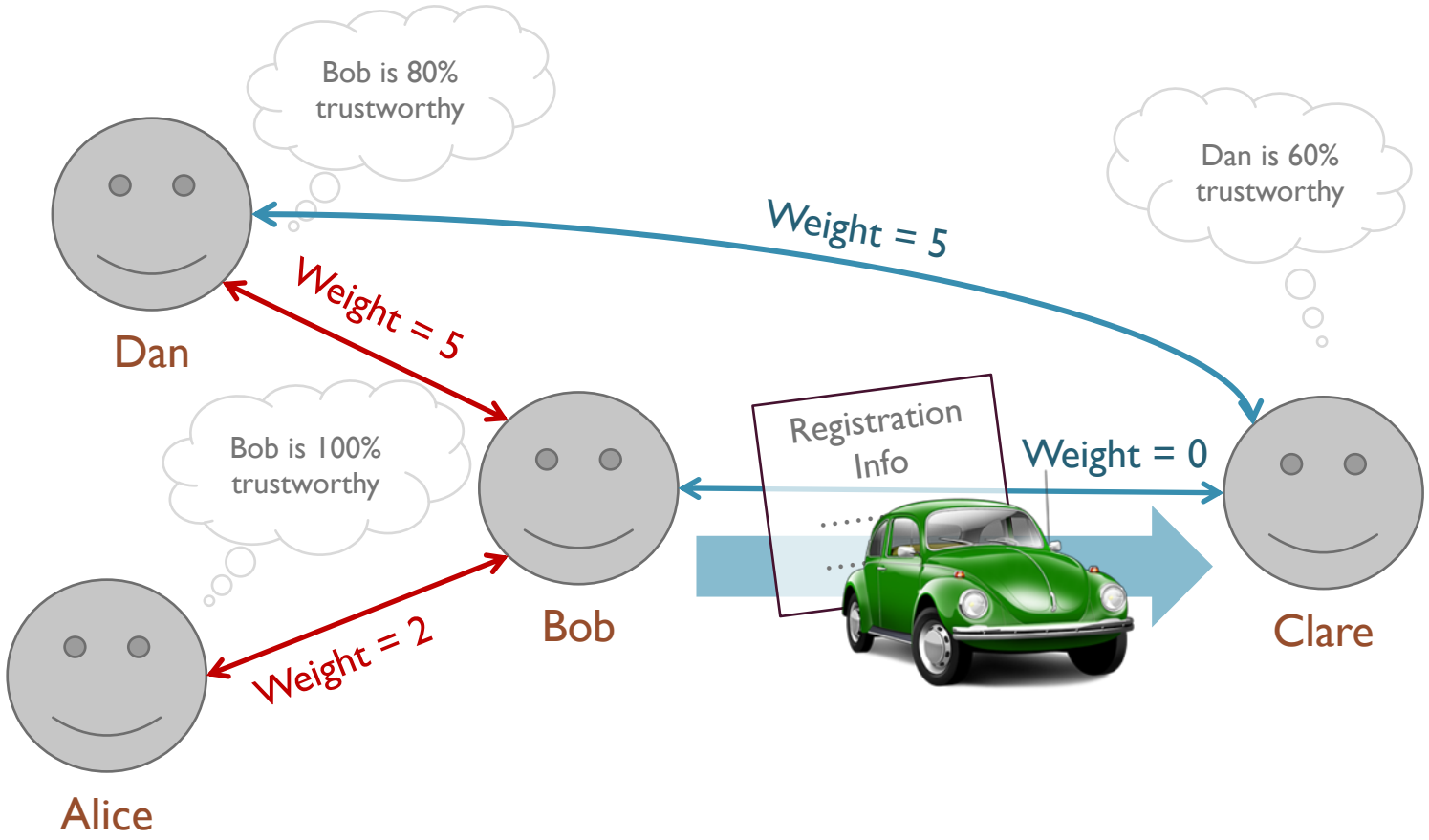
HOW MIGHT WE CALCULATE TRUST?



So how do we combine these in a meaningful way?!

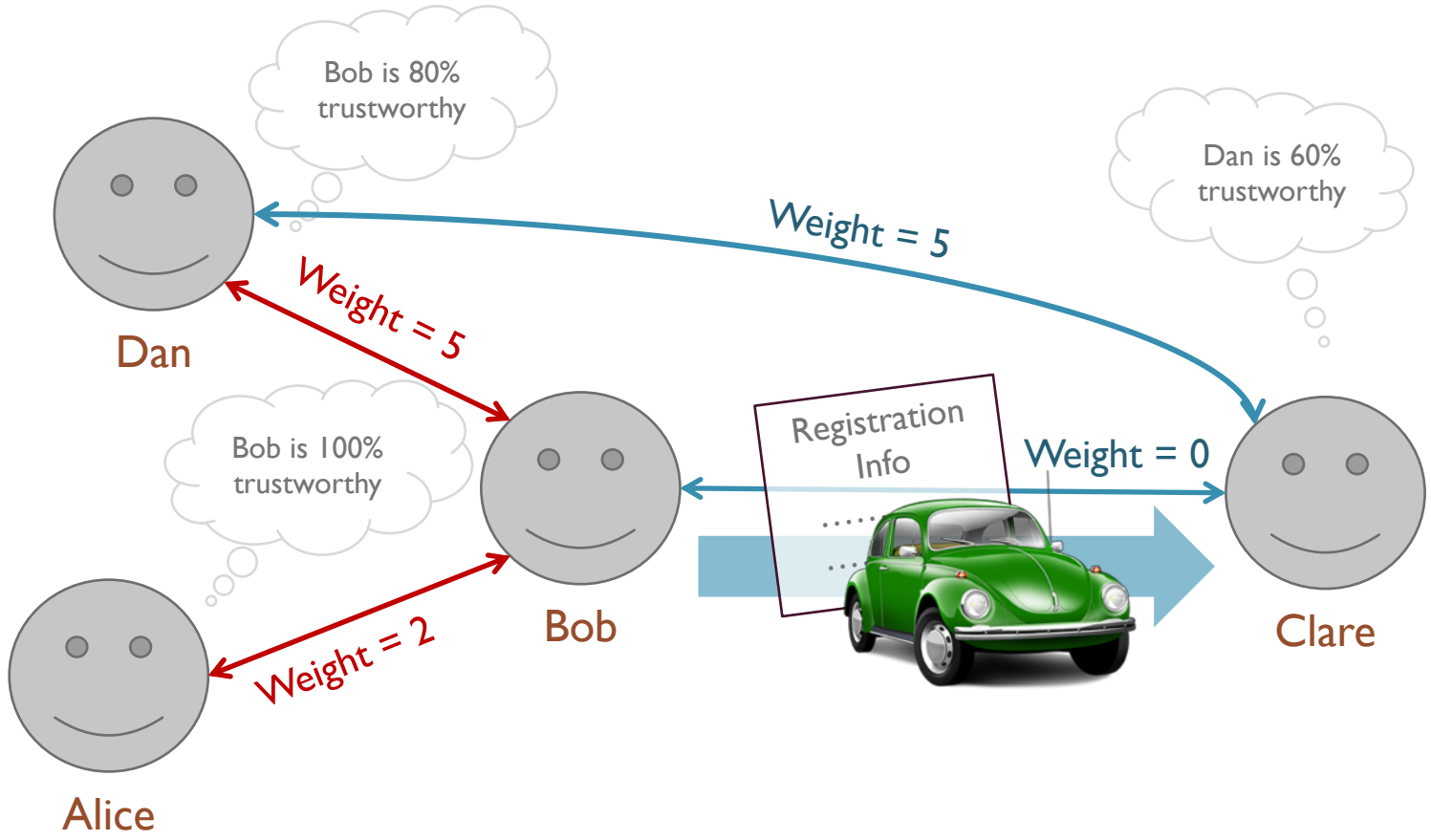
Trust based on personal experience or the experiences of others – their reputation

HOW MIGHT WE CALCULATE TRUST?



Global Trust means the **whole network** is used to calculate a **single trust value** for each node

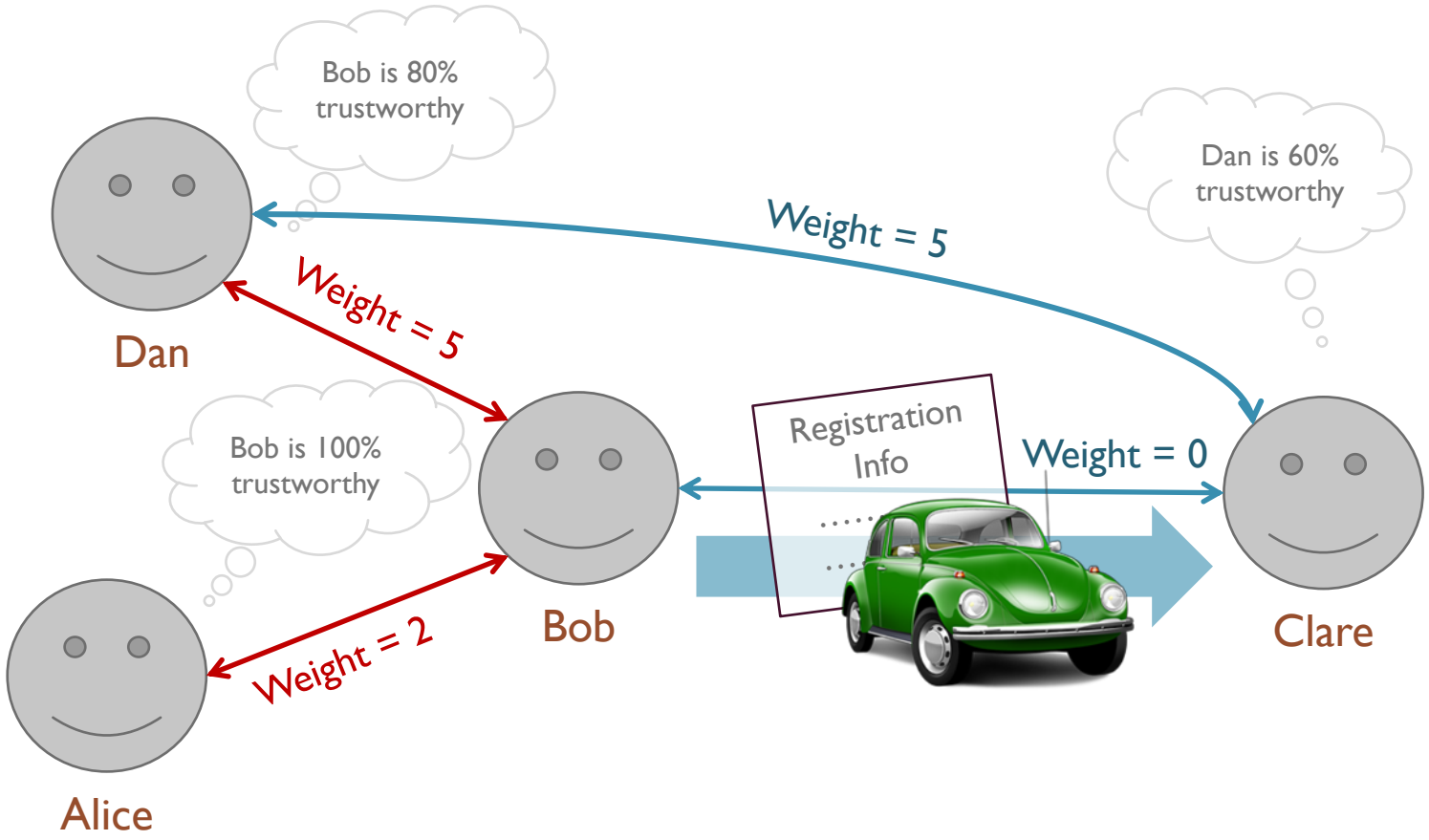
HOW MIGHT WE CALCULATE TRUST?



Global Trust means the **whole network** is used to calculate a **single trust value** for each node

$$\frac{(5 * 0.8) + (2 * 1.0)}{(5 + 2)} = \frac{6}{7} = 85\%$$

HOW MIGHT WE CALCULATE TRUST?



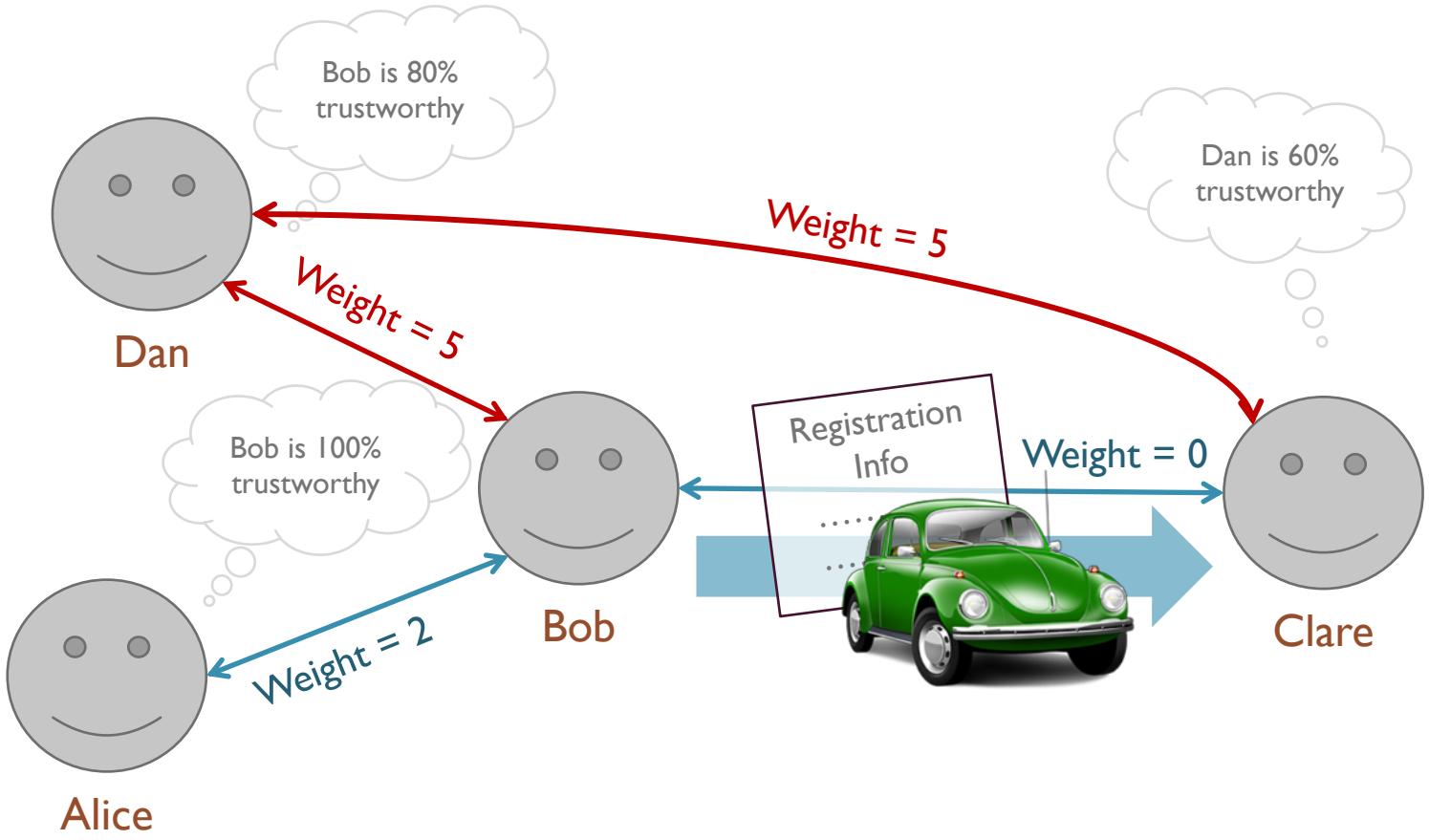
Global Trust means the **whole network** is used to calculate a **single trust value** for each node

Resnick et al. show buyers will pay a STRONG-reputation seller **8.1% more** on average than a NEW seller

A Pretty Poor system that is easy to exploit - Example of **Yhprum's Law**?

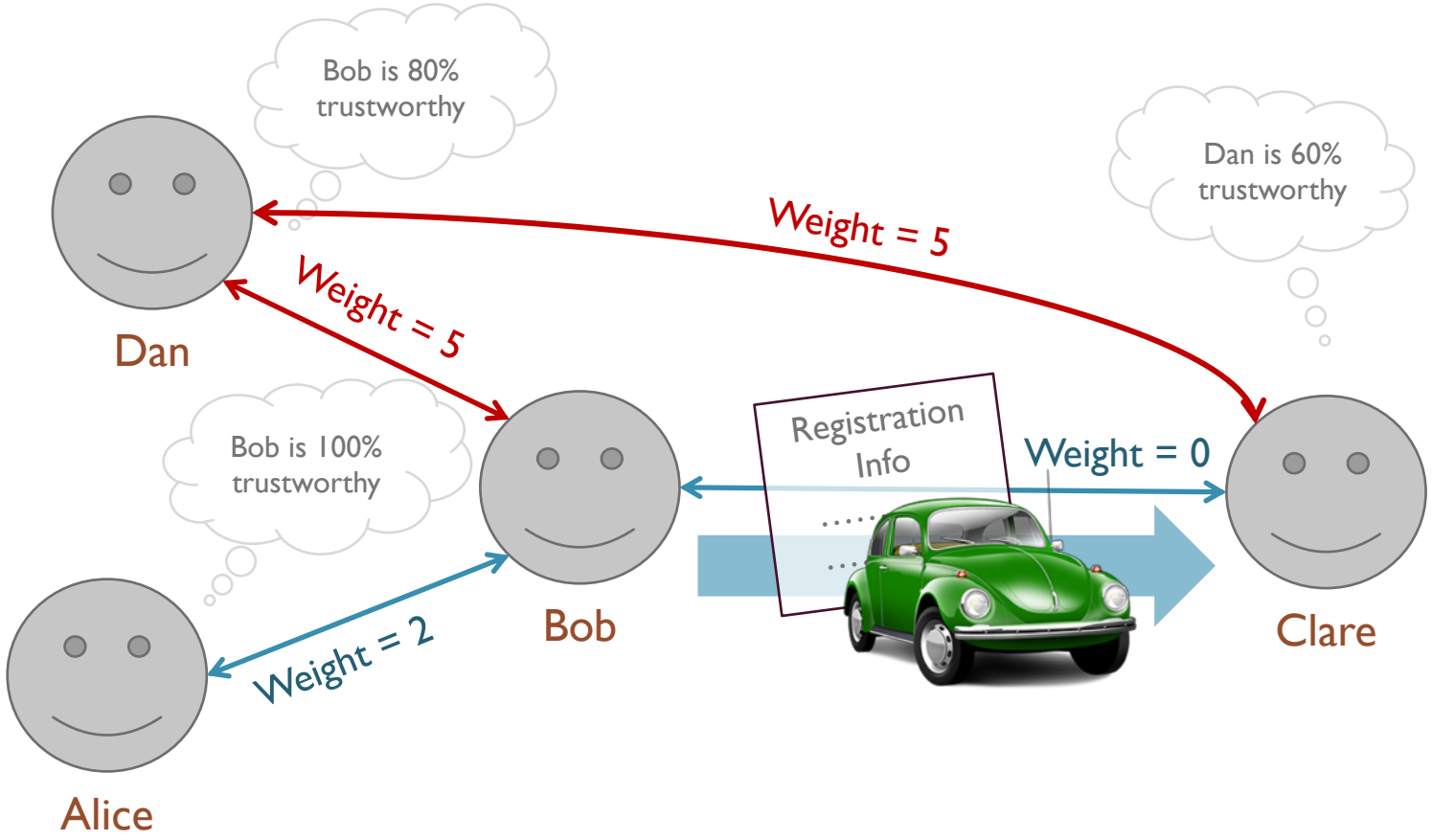
Resnick et al. The value of reputation on eBay: A controlled experiment. Experimental Economics (2006) vol. 9 (2) pp. 79-101

HOW MIGHT WE CALCULATE TRUST?



Local Trust means Trust is held from the view of a particular person – it is thus **personalised**

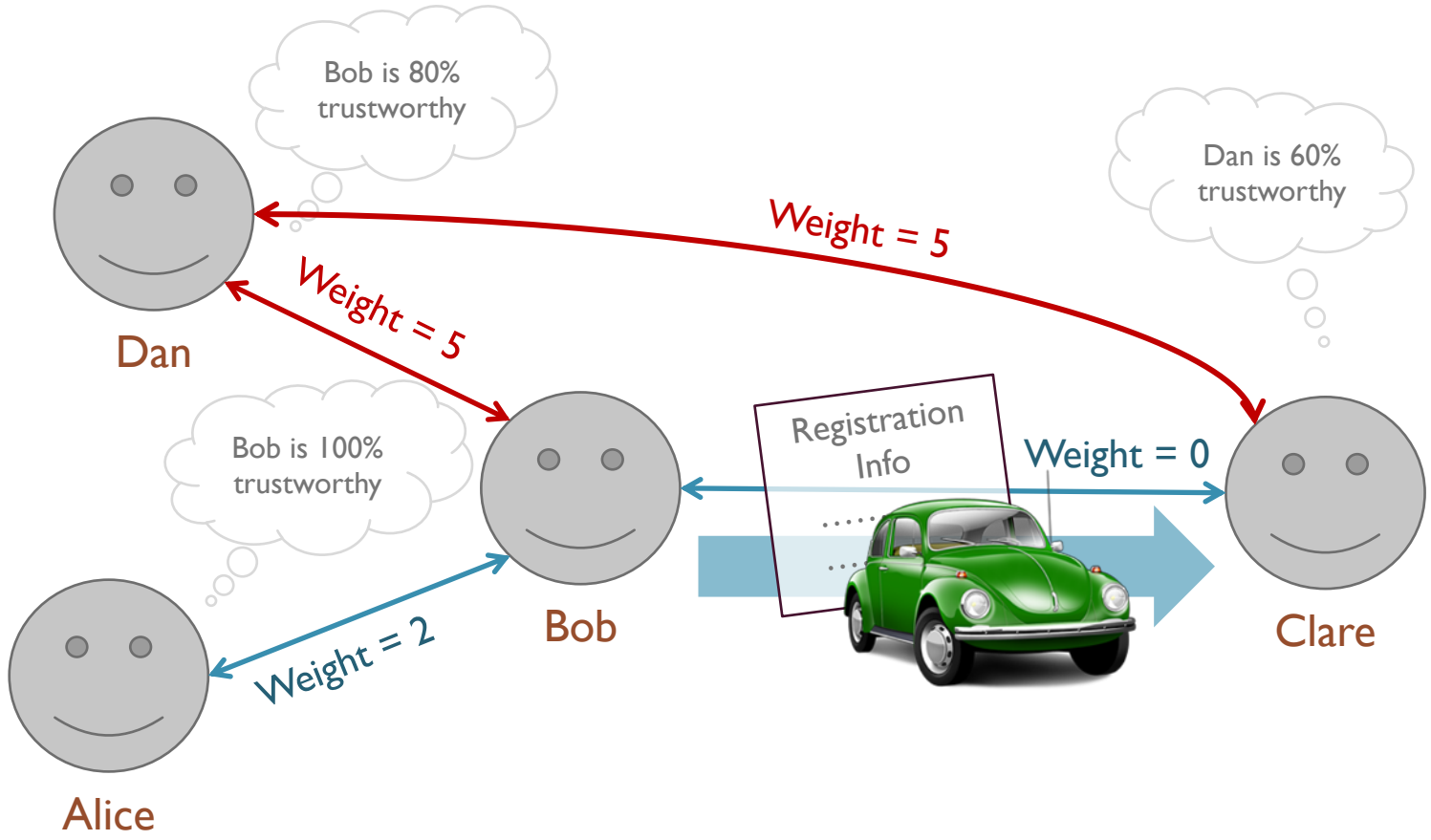
HOW MIGHT WE CALCULATE TRUST?



Local Trust means Trust is held from the view of a particular person – it is thus **personalised**

$$\frac{(5 * 0.6)}{5} * \frac{(5 * 0.8)}{5} = \frac{3}{5} * \frac{4}{5} = 48\%$$

HOW MIGHT WE CALCULATE TRUST?



Local Trust means Trust is held from the view of a particular person – it is thus **personalised**

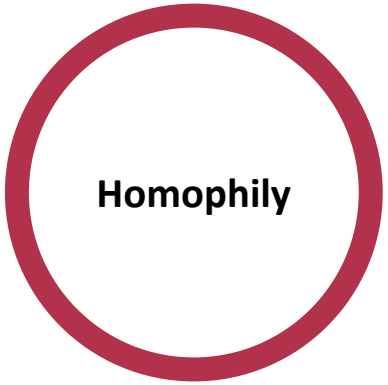
Relies on **Transitivity**
If A trusts B
and B trusts J
then A will also trust J

Assumes **Asymmetry**
If A trusts B
it doesn't follow that
B trusts A

PART 2

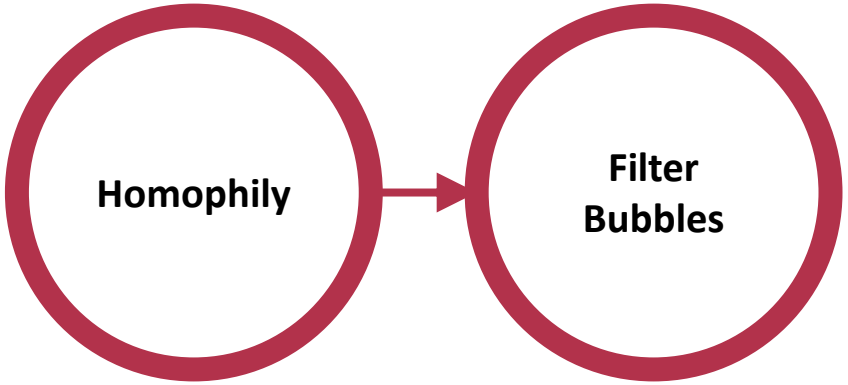
FAKE NEWS – A CRISIS OF TRUST?

FAKE NEWS



Tendency to create connections to similar people (gender, age, class, politics, religion, etc.)

FAKE NEWS



Tendency to create connections to similar people (gender, age, class, politics, religion, etc.)

Algorithms designed to please, will learn from this network and reflect your views back to you

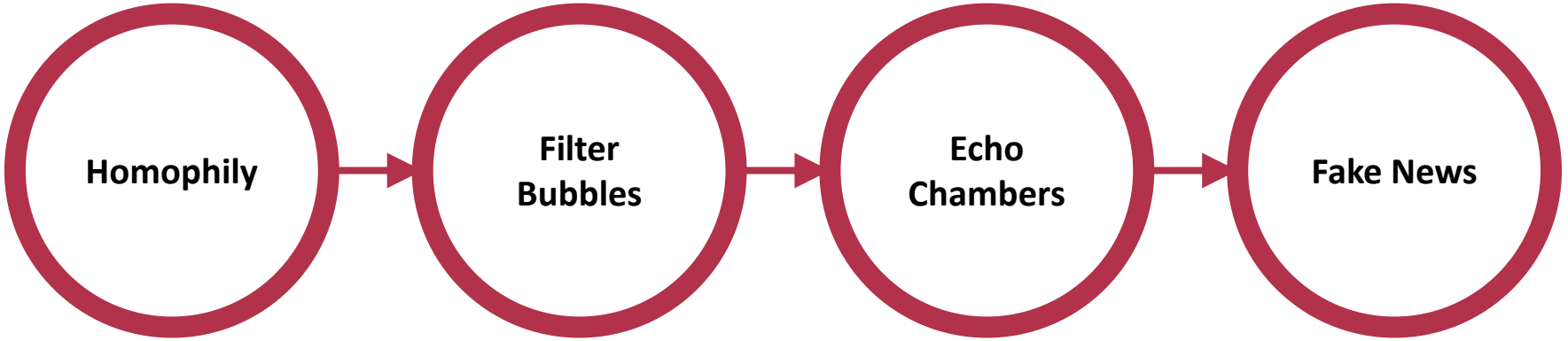
FAKE NEWS

Tendency to create connections to similar people (gender, age, class, politics, religion, etc.)

Algorithms designed to please, will learn from this network and reflect your views back to you

Reflected views reinforce existing beliefs and normalise extreme views

FAKE NEWS



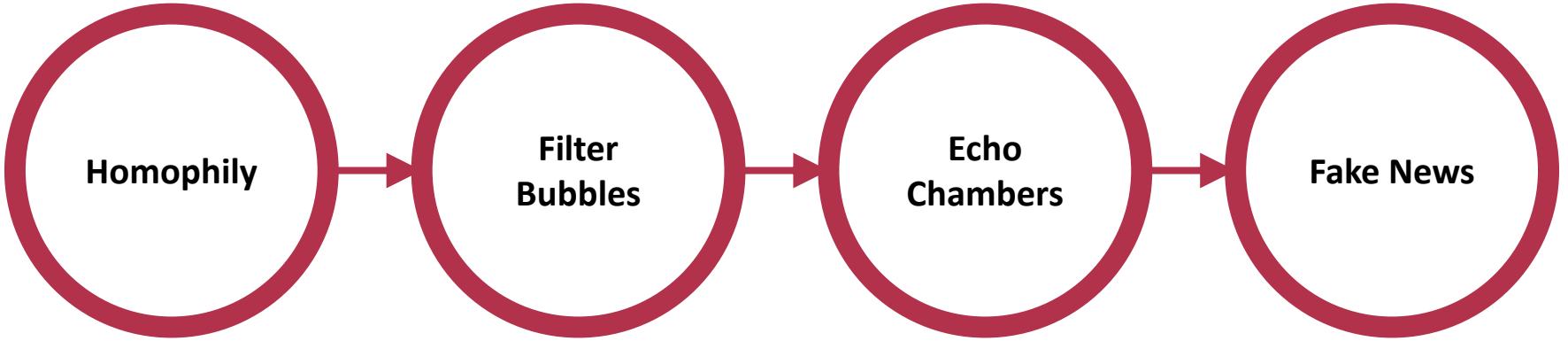
Tendency to create connections to similar people (gender, age, class, politics, religion, etc.)

Algorithms designed to please, will learn from this network and reflect your views back to you

Reflected views reinforce existing beliefs and normalise extreme views

Articles designed to propagate in this environment (clickbait) – **with monetary or political aims**

THE POST-TRUTH WORLD



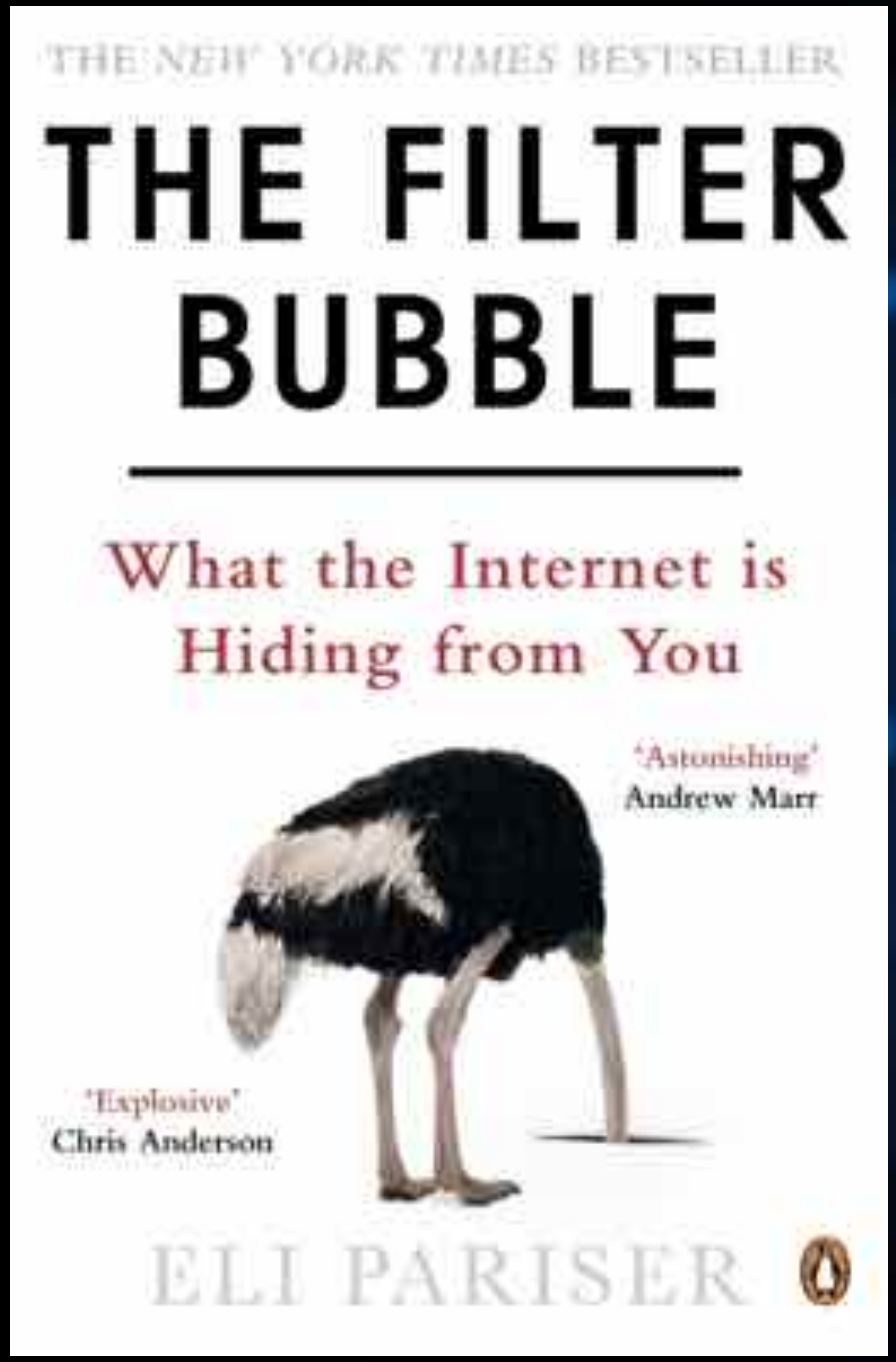
Tendency to create connections to similar people (gender, age, class, politics, religion, etc.)

Algorithms designed to please, will learn from this network and reflect your views back to you

Reflected views reinforce existing beliefs and normalise extreme views

Articles designed to propagate in this environment (clickbait) – **with monetary or political aims**

Public narrative is disconnected from science or policy, and instead framed emotionally or ideologically


THE NEW YORK TIMES BESTSELLER
THE FILTER BUBBLE
What the Internet is Hiding from You
'Astonishing'
Andrew Marr
'Explosive'
Chris Anderson
ELI PARISER

FILTER BUBBLES?

Bakkshy et al. 2015 study into exposure to diverse views on Facebook

10.1 million active US users with self-declared affiliation

7 million links shared over a 6 month period

Eytan Bakkshy, Solomon Messing, Lada Adamic (2015). Exposure to ideologically diverse news and opinion on Facebook. Science. May 2015

FILTER BUBBLES?

Bakkshy et al. 2015 study into exposure to diverse views on Facebook

10.1 million active US users with self-declared affiliation

7 million links shared over a 6 month period

This is a study by Facebook

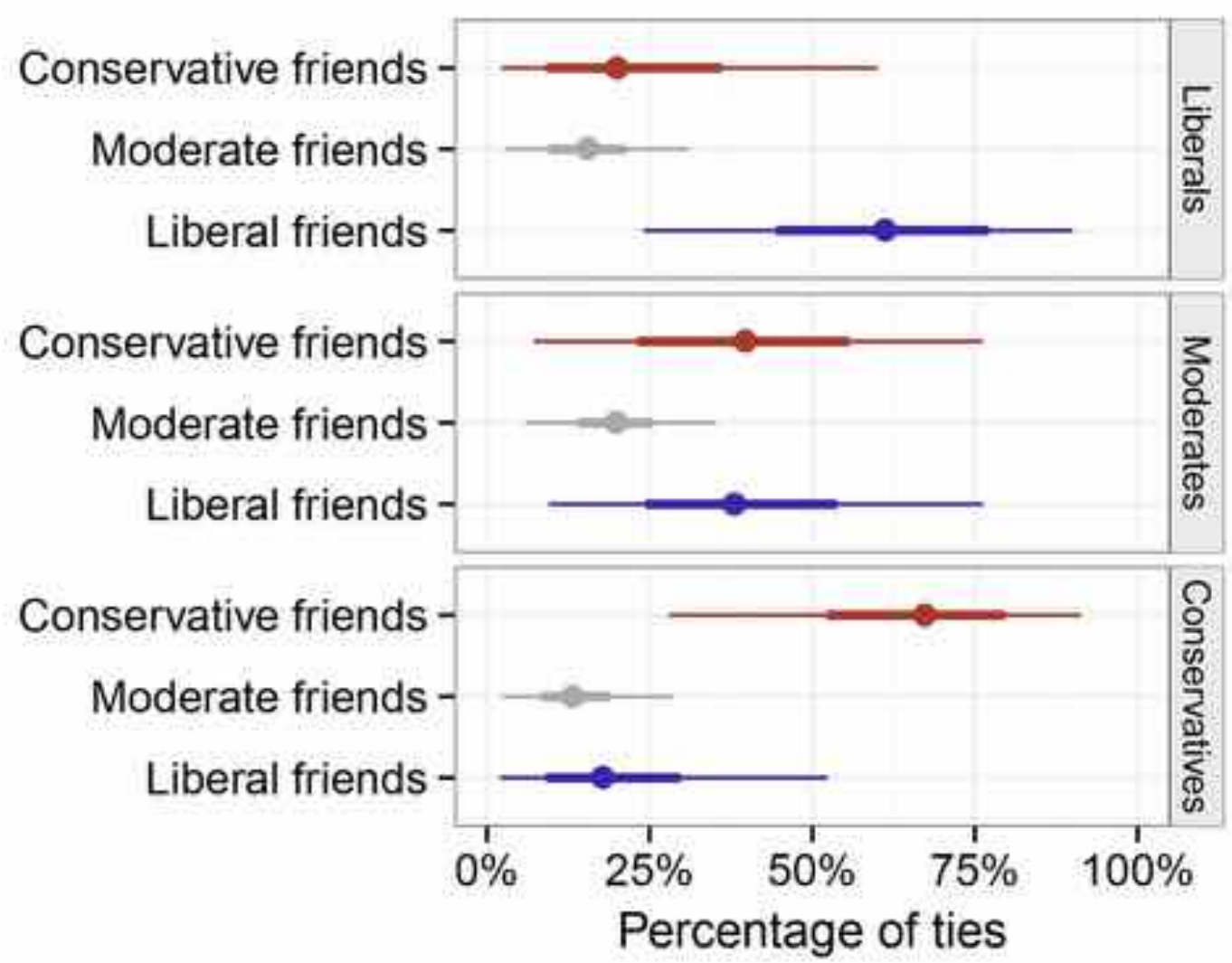
Eytan Bakkshy, Solomon Messing, Lada Adamic (2015). Exposure to ideologically diverse news and opinion on Facebook. Science. May 2015

FILTER BUBBLES?

Bakkshy et al. 2015 study into exposure to diverse views on Facebook

10.1 million active US users with self-declared affiliation

7 million links shared over a 6 month period



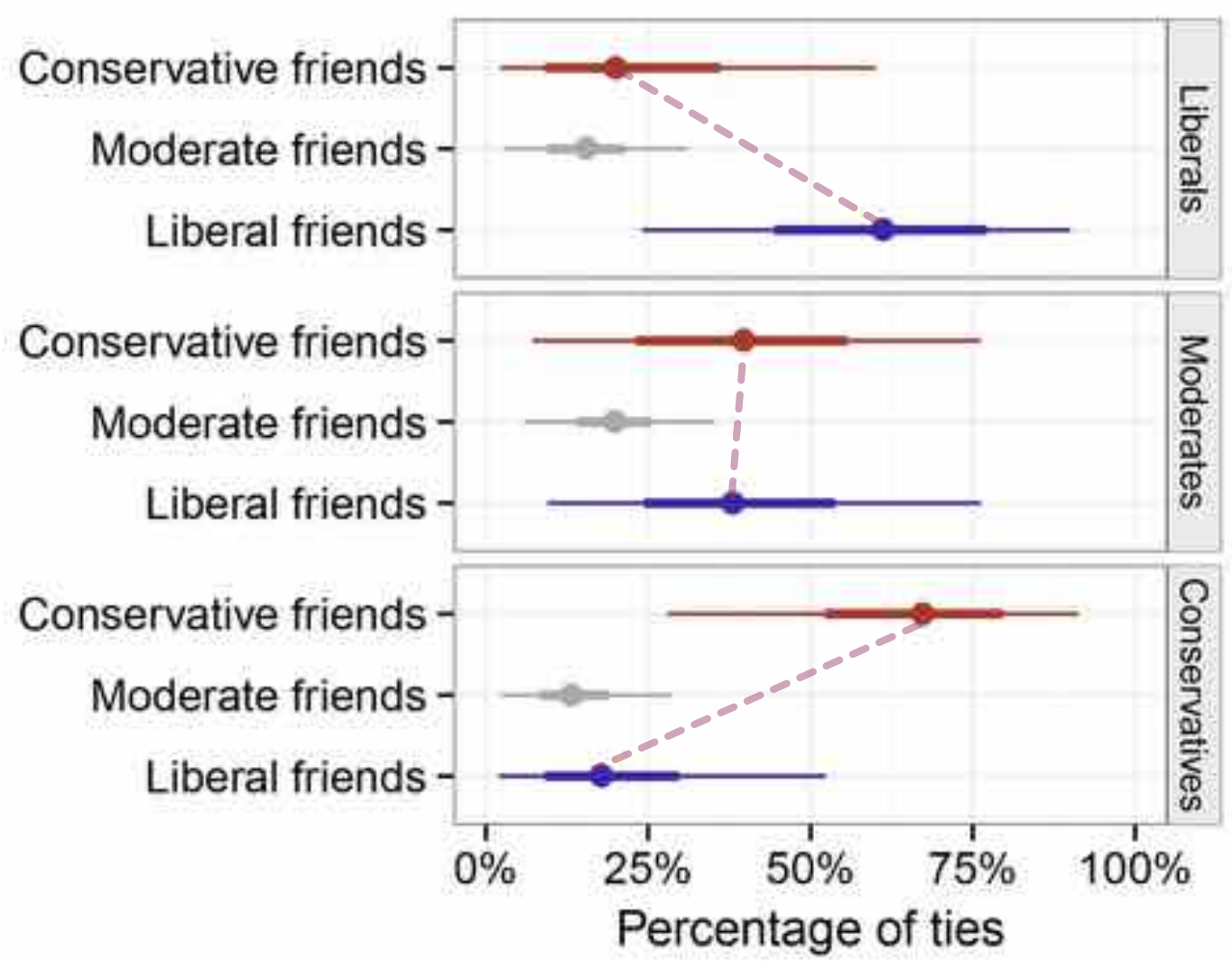
Eytan Bakkshy, Solomon Messing, Lada Adamic (2015). Exposure to ideologically diverse news and opinion on Facebook. Science. May 2015

FILTER BUBBLES?

Bakkshy et al. 2015 study into exposure to diverse views on Facebook

10.1 million active US users with self-declared affiliation

7 million links shared over a 6 month period



Clear evidence of political homophily

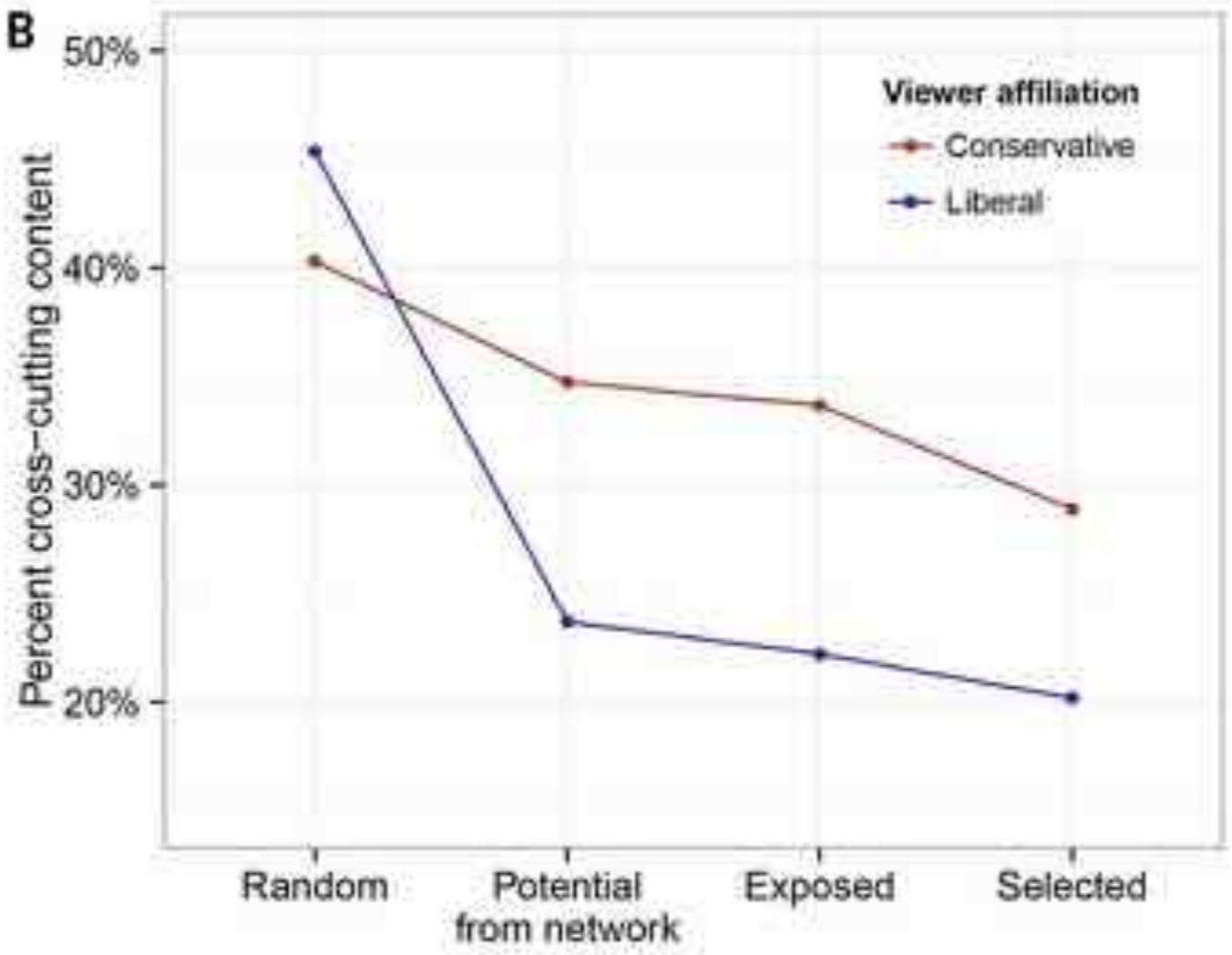
Eytan Bakkshy, Solomon Messing, Lada Adamic (2015). Exposure to ideologically diverse news and opinion on Facebook. Science. May 2015

FILTER BUBBLES?

Bakkshy et al. 2015 study into exposure to diverse views on Facebook

10.1 million active US users with self-declared affiliation

7 million links shared over a 6 month period



Eytan Bakkshy, Solomon Messing, Lada Adamic (2015). Exposure to ideologically diverse news and opinion on Facebook. Science. May 2015

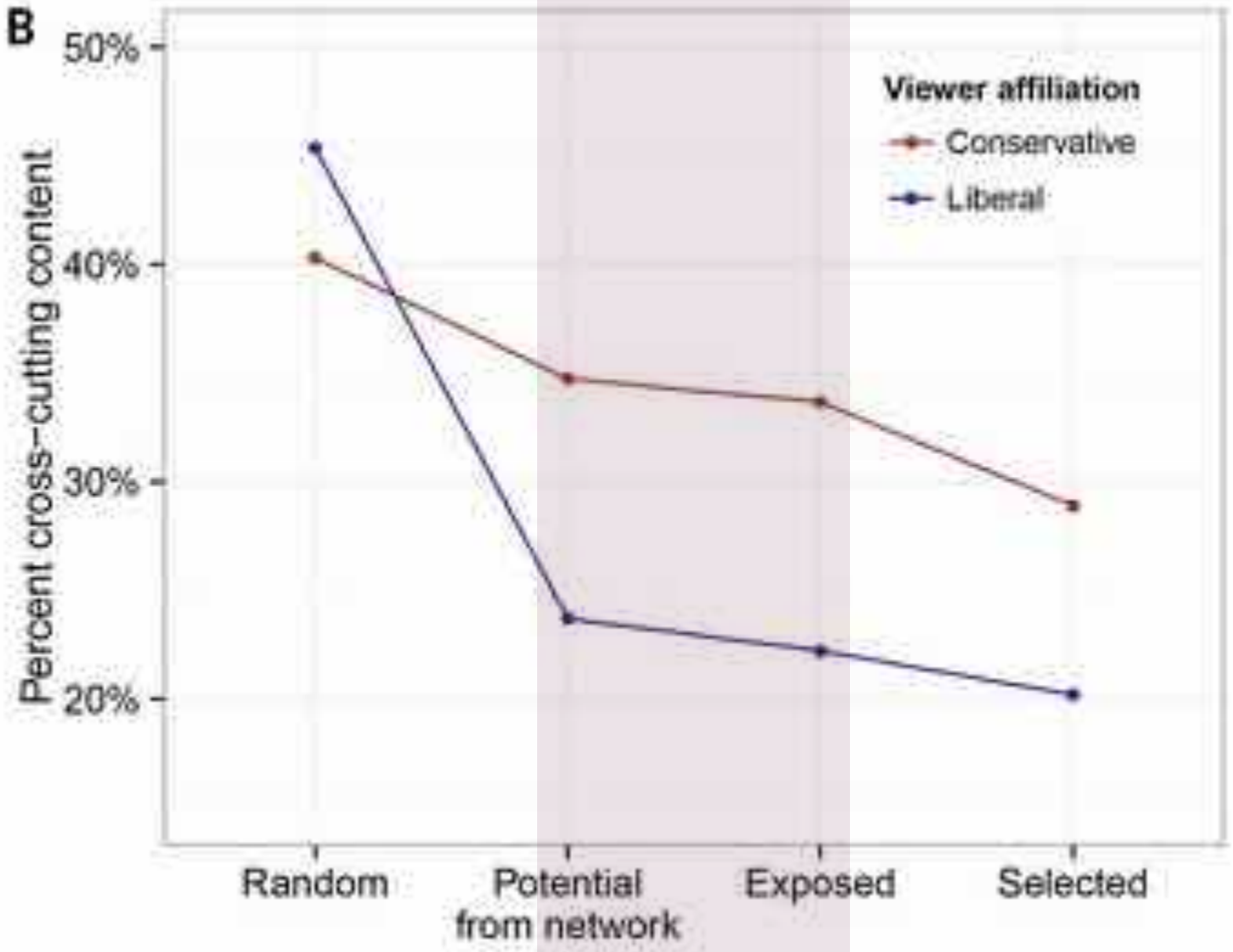
FILTER BUBBLES?

Bakkshy et al. 2015 study into exposure to diverse views on Facebook

10.1 million active US users with self-declared affiliation

7 million links shared over a 6 month period

Filter Bubble

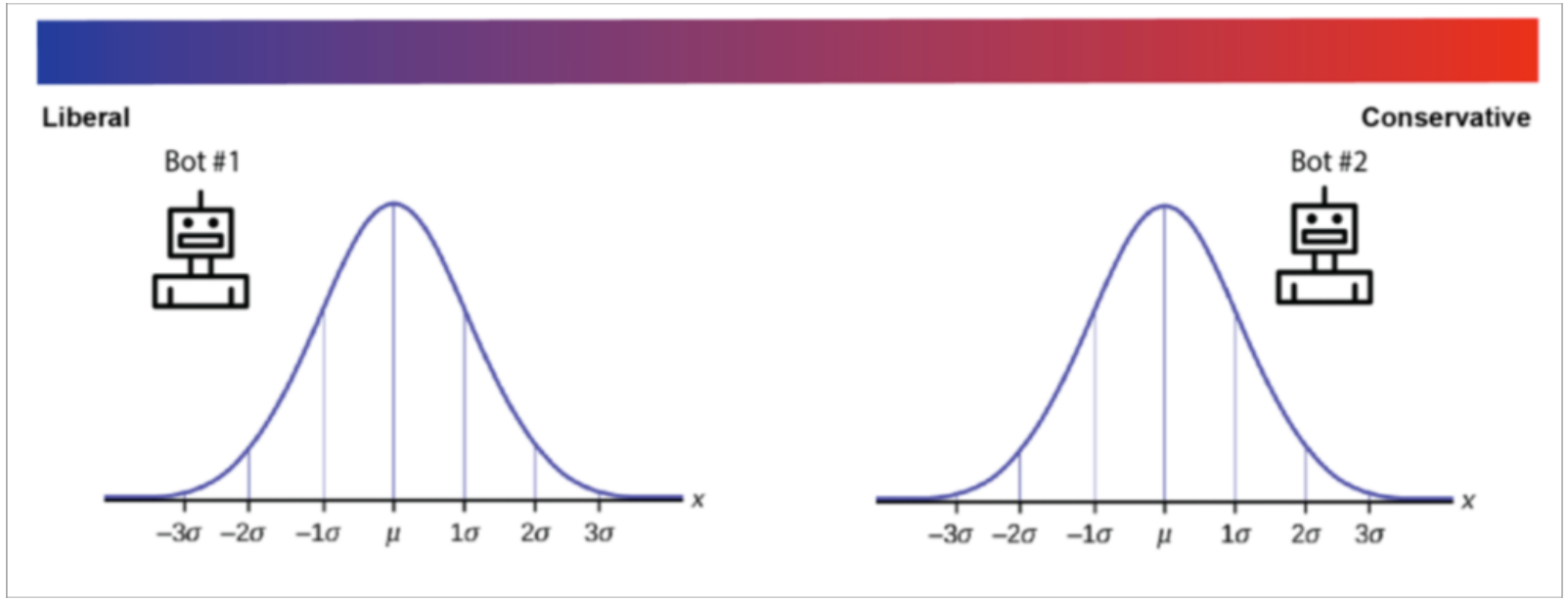


Clear evidence of filter bubble – but filter bubble is smallest effect

Eytan Bakkshy, Solomon Messing, Lada Adamic (2015). Exposure to ideologically diverse news and opinion on Facebook. Science. May 2015

ECHO CHAMBERS?

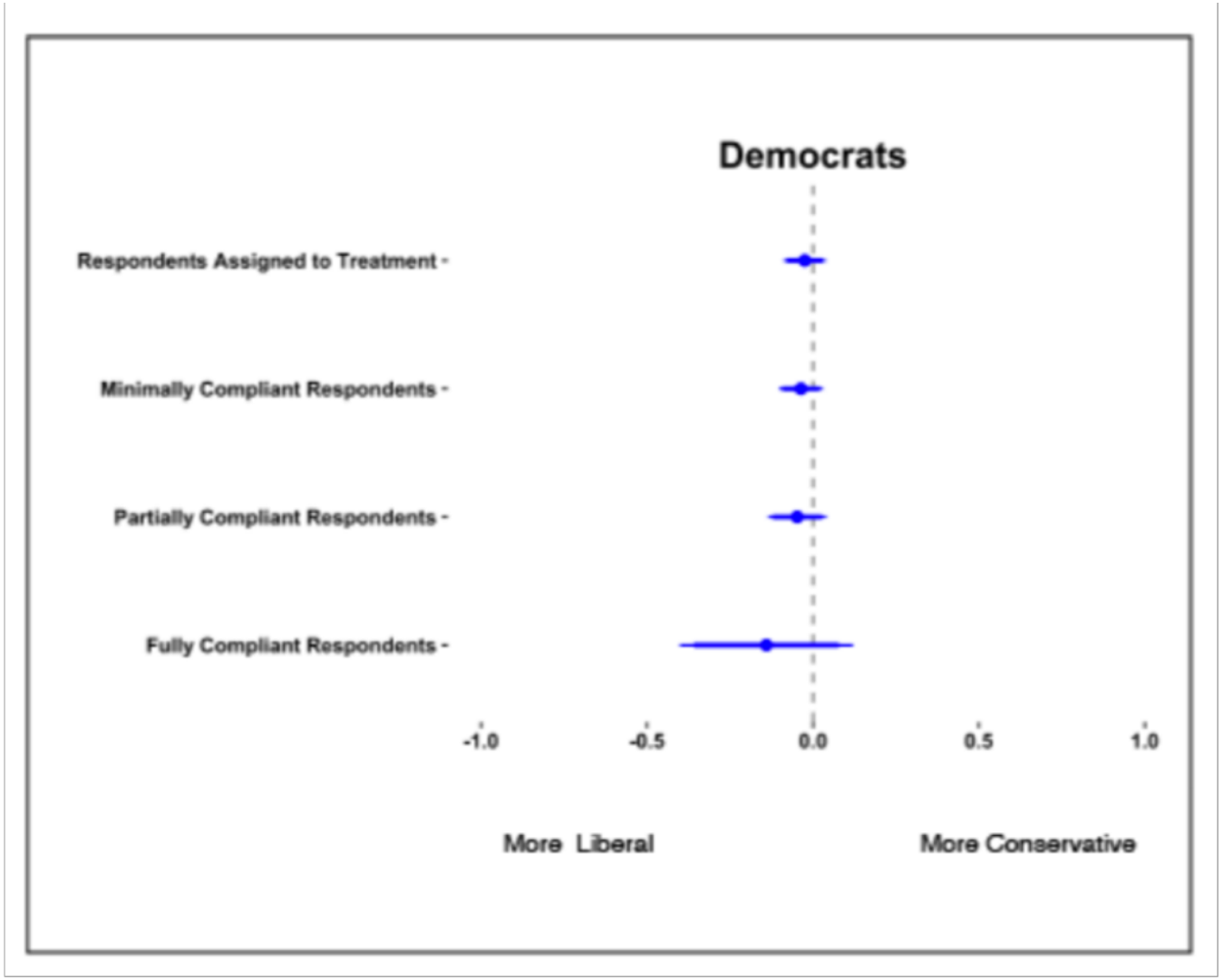
Bail et al. (2018) paid a sample of 697 Democrats and 542 Republicans to follow a political bot that counter retweeted political messages into their Twitter feeds – and measure the impact on their views



Bail et al. (2018) Exposure to Opposing Views can Increase Political Polarization: Evidence from a Large-Scale Field Experiment on Social Media, SocARXiv

ECHO CHAMBERS?

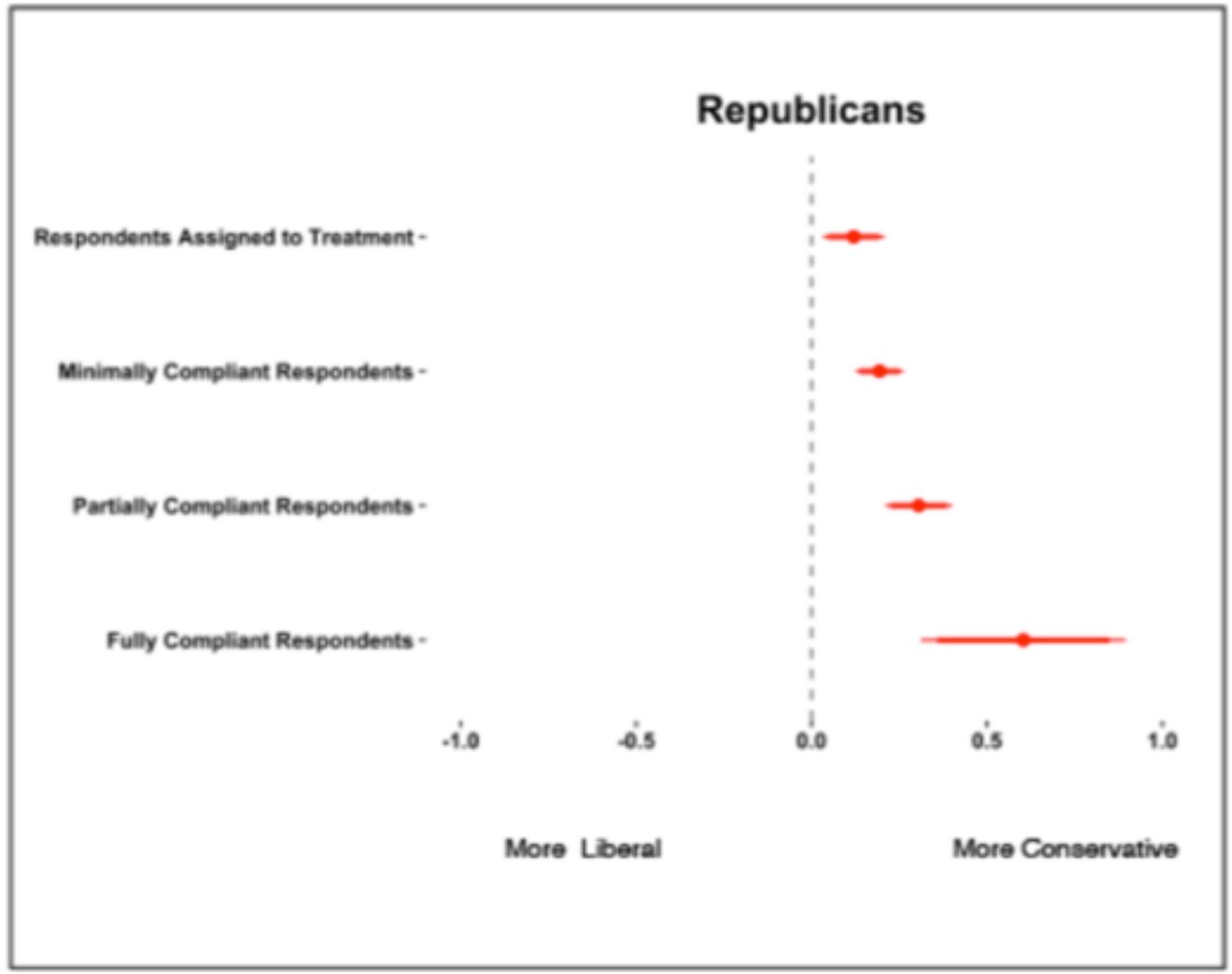
Bail et al. (2018) paid a sample of 697 Democrats and 542 Republicans to follow a political bot that counter retweeted political messages into their Twitter feeds – and measure the impact on their views



Bail et al. (2018) Exposure to Opposing Views can Increase Political Polarization: Evidence from a Large-Scale Field Experiment on Social Media, SocARXiv

ECHO CHAMBERS?

Bail et al. (2018) paid a sample of 697 Democrats and 542 Republicans to follow a political bot that counter retweeted political messages into their Twitter feeds – and measure the impact on their views



Bail et al. (2018) Exposure to Opposing Views can Increase Political Polarization: Evidence from a Large-Scale Field Experiment on Social Media, SocARXiv

THE BACKFIRE EFFECT

"our study indicates that well- intentioned attempts to introduce people to opposing political views on social media might not only be ineffective, but **counter-productive**—particularly if they are initiated by Democrats."

Bail et al. (2018) Exposure to Opposing Views can Increase Political Polarization: Evidence from a Large-Scale Field Experiment on Social Media, SocARXiv

PART 3

CREDIBILITY – HOW DO PEOPLE MAKE TRUST JUDGEMENTS?

Credibility

Trust is typically framed as a quality between individuals (interpersonal)

Credibility is a measure of how "believable" a piece of information is

Fogg, B. J. and Tseng, H. (1999). The elements of computer credibility. In ACM CHI '99, pg 80–87, New York, NY, USA.

SOURCE, MESSAGE, AND MEDIA CREDIBILITY

Hellmueller, L. and Trilling, D. (2012). The credibility of credibility measures: A meta- analysis in leading communication journals, 1951 to 2011. In WAPOR 65th Annual Conference in Hong Kong.

Metzger, et al. (2003). Credibility for the 21st century: integrating perspectives on source, message, and media credibility in the contemporary media environment. In *Communication Yearbook 27*

SOURCE, MESSAGE, AND MEDIA CREDIBILITY

Where it **comes from**.

Influenced by:

- Perceived competence
- Trustworthiness
- Goodwill

Hellmueller, L. and Trilling, D. (2012). The credibility of credibility measures: A meta- analysis in leading communication journals, 1951 to 2011. In WAPOR 65th Annual Conference in Hong Kong.
Metzger, et al. (2003). Credibility for the 21st century: integrating perspectives on source, message, and media credibility in the contemporary media environment. In *Communication Yearbook 27*

SOURCE, MESSAGE, AND MEDIA CREDIBILITY

The **quality**.

Influenced by:

- Information quality
- Language intensity
- Message discrepancy

Hellmueller, L. and Trilling, D. (2012). The credibility of credibility measures: A meta- analysis in leading communication journals, 1951 to 2011. In WAPOR 65th Annual Conference in Hong Kong.
Metzger, et al. (2003). Credibility for the 21st century: integrating perspectives on source, message, and media credibility in the contemporary media environment. In *Communication Yearbook 27*

SOURCE, MESSAGE, AND MEDIA CREDIBILITY

The **media type**.

Influenced by:

- by perception of media type

Hellmueller, L. and Trilling, D. (2012). The credibility of credibility measures: A meta- analysis in leading communication journals, 1951 to 2011. In WAPOR 65th Annual Conference in Hong Kong.
Metzger, et al. (2003). Credibility for the 21st century: integrating perspectives on source, message, and media credibility in the contemporary media environment. In *Communication Yearbook 27*

THE CORE OF THE PROBLEM

- Elaboration Likelihood Model (ELM) (Petty and Cacioppo, 1986):
  - **Central cues** – careful and thoughtful consideration of the content of the message (high effort)
  - **Peripheral cues** – surface-level features such as design and usability (low effort)

- Examining central cues results in stable, long-term changes to attitude.
- In cases where users interact with topics where they lack the requisite interest or knowledge, they are more likely to rely on peripheral cues.

Petty, R. E. and Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. Adv in Experimental Social Psychology, 19(C):123–205.

COMBATTING FAKE NEWS

Emine Saner, From Wikitribune to StopFake: the battle against fake news, The Guardian, 25 Apr 2017

COMBATTING FAKE NEWS

- **Google**
  - Modified their algorithm to foreground more reliable sources
  - New tools for users to report problems with infoboxes and autocomplete

Emine Saner, From Wikitribune to StopFake: the battle against fake news, The Guardian, 25 Apr 2017

COMBATTING FAKE NEWS

- **Google**
  - Modified their algorithm to foreground more reliable sources
  - New tools for users to report problems with infoboxes and autocomplete

- **Facebook**
  - Users can tag stories as 'disputed'
  - Sent to independent fact checkers (e.g. Snopes)
  - Other users are warned if they try to share a disputed story

Emine Saner, From Wikitribune to StopFake: the battle against fake news, The Guardian, 25 Apr 2017

COMBATTING FAKE NEWS

- **Google**
  - Modified their algorithm to foreground more reliable sources
  - New tools for users to report problems with infoboxes and autocomplete

- **Facebook**
  - Users can tag stories as 'disputed'
  - Sent to independent fact checkers (e.g. Snopes)
  - Other users are warned if they try to share a disputed story

- **WikiTribune**
  - Professional Journalists producing transparent stories (with sources)
  - Augmented by community members

Emine Saner, From Wikitribune to StopFake: the battle against fake news, The Guardian, 25 Apr 2017

COMBATTING FAKE NEWS

- **Google**
  - Modified their algorithm to foreground more reliable sources
  - New tools for users to report problems with infoboxes and autocomplete

Reputation

- **Facebook**
  - Users can tag stories as 'disputed'
  - Sent to independent fact checkers (e.g. Snopes)
  - Other users are warned if they try to share a disputed story

Policy

- **WikiTribune**
  - Professional Journalists producing transparent stories (with sources)
  - Augmented by community members

Provenance

Emine Saner, From Wikitribune to StopFake: the battle against fake news, The Guardian, 25 Apr 2017

The Liar's Dividend

"When nothing is true then the dishonest person will thrive by saying what's true is fake."

- Danielle Citron

https://www.theguardian.com/technology/2018/nov/12/deep-fakes-fake-news-truth

Chesney, R. and Citron, D., Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security (July 14, 2018). 107 California Law Review (2019)

?