ECN 119: Economics and Psychology
Belief

Jim Campbell

UC Berkeley

Summer 2021

Belief

What happens when new information comes to light? The gold standard of Bayes's Rule is not often a good match for the messy ways that humans process news and form their opinions. The cold, hard math of probability theory can expose the ways in which our brains play tricks on us when we try to understand the world.

In this section we will see evidence on how people learn and respond to what other people do and what they see in the world. Conspiracy theories, wilful ignorance, stubbornness, and gullibility are all on the table. We will also explore in more depth whether we even really understand ourselves. Are even our beliefs about our own selves irredeemably flawed?

Belief

In this section:

1. Bayes' Theorem
2. The law of small numbers
3. Correlation neglect
4. Overconfidence
5. Motivated beliefs
6. The conjunction effect and portioning
7. Dopamine and the response to news
8. Projection bias and diversification bias
9. Persuasion
10. Dialectic belief formation
11. Information cascades and social inference

Belief

In traditional economic theory, beliefs play a crucial role

- In game theory one needs to consider what a person believes to be true about the setting or their opponents and how that belief will change with new information
    - You are about to play a game of darts when your opponent pulls out their own set of fancy-looking darts
    - You are about to launch your new product when you hear that your closest competitor buys a Super Bowl ad
- In macroeconomics one needs to consider what people are thinking about the future
    - If households expect bad times ahead they may tighten their belts making the expectation a self-fulfilling prophecy
    - Inflation expectations of households and investors may influence their consumption and savings decisions
- And we have seen the role of subjective probability assessment already in the course

Bayes' Theorem

Bayes' Theorem gives us the mathematically correct way to update beliefs:

$$\Pr(A|B) = \frac{\Pr(B|A)\Pr(A)}{\Pr(B)} \tag{1}$$

- $\Pr(A|B)$ is notation for conditional probability: what is the probability of event A given that we have observed event B?
- Unfortunately there is a bunch of evidence that people are not good at forming and updating beliefs in this way (see Benjamin 2018 from the reading list)

Bayes' Theorem: an example

- Let's do a really simple example—one that we can easily see the answer to before we do the calculation
- I will roll a fair six-sided die but I won't show it to you
- Instead I will just tell you (truthfully) whether it was even or odd
- If I tell you it came up even, what is the probability it was a six?

Bayes' Theorem: an example

- Let's do a really simple example—one that we can easily see the answer to before we do the calculation
- I will roll a fair six-sided die but I won't show it to you
- Instead I will just tell you (truthfully) whether it was even or odd
- If I tell you it came up even, what is the probability it was a six?

$$\Pr(6|\mathit{even}) = \frac{\Pr(\mathit{even}|6)\Pr(6)}{\Pr(\mathit{even})} \tag{2}$$

Bayes' Theorem: an example

- Let's do a really simple example—one that we can easily see the answer to before we do the calculation
- I will roll a fair six-sided die but I won't show it to you
- Instead I will just tell you (truthfully) whether it was even or odd
- If I tell you it came up even, what is the probability it was a six?

$$\Pr(6|\mathit{even}) = \frac{\Pr(\mathit{even}|6)\Pr(6)}{\Pr(\mathit{even})} \tag{2}$$

- $\Pr(\mathit{even}|6) = 1$; $Pr(6) = \frac{1}{6}$; $\Pr(\mathit{even}) = \frac{1}{2}$

$$\Pr(6|\mathit{even}) = \frac{1 \cdot \frac{1}{6}}{\frac{1}{2}} = \frac{1}{3} \tag{3}$$

Bayes' Theorem: a 'medical test' example

- Say that 3% of people in the population have a disease
- You have a test that correctly identifies those with the disease 98% of the time (i.e. some false negatives happen)
- And it correctly identifies those who don't have the disease 95% of the time (i.e. some false positives happen)
- What is the chance that the person with a positive test has the disease?

Bayes' Theorem: a 'medical test' example

- Say that 3% of people in the population have a disease
- You have a test that correctly identifies those with the disease 98% of the time (i.e. some false negatives happen)
- And it correctly identifies those who don't have the disease 95% of the time (i.e. some false positives happen)
- What is the chance that the person with a positive test has the disease?

$$\Pr(\mathit{disease}|+) = \frac{\Pr(+|\mathit{disease})\Pr(\mathit{disease})}{\Pr(+)} \tag{4}$$

$$= \frac{0.98 \cdot 0.03}{(0.03 \cdot 0.98) + (0.97 \cdot 0.05)} \simeq 0.377 \tag{5}$$

Bayes' Theorem: a 'medical test' example

- Say that 3% of people in the population have a disease
- You have a test that correctly identifies those with the disease 98% of the time (i.e. some false negatives happen)
- And it correctly identifies those who don't have the disease 95% of the time (i.e. some false positives happen)
- What is the chance that the person with a positive test has the disease?

$$\Pr(\textit{disease}|+) = \frac{\Pr(+|\textit{disease})\Pr(\textit{disease})}{\Pr(+)} \tag{4}$$

$$= \frac{0.98 \cdot 0.03}{(0.03 \cdot 0.98) + (0.97 \cdot 0.05)} \simeq 0.377 \tag{5}$$

- This is a big jump in probability from 5% population frequency, but maybe it doesn't look *that* big given how accurate the test is
- The surprisingness here is an example of 'base rate neglect' (Kahneman & Tversky 1973)

Bayes' Theorem: one more example

Here is a neat example from Botond Kőszegi via Vincent Crawford:

- Coin flip; your prior is that it's fair (50% heads) with probability $\frac{2}{3}$ and biased (75% heads) with probability $\frac{1}{3}$
- It comes up heads: how much should you decrease your belief that the coin is fair?

Bayes' Theorem: one more example

Here is a neat example from Botond Kőszegi via Vincent Crawford:

- Coin flip; your prior is that it's fair (50% heads) with probability $\frac{2}{3}$ and biased (75% heads) with probability $\frac{1}{3}$
- It comes up heads: how much should you decrease your belief that the coin is fair?
- Imagine two urns with poker chips
  1. Urn 1: half of the chips are H and half are T
  2. Urn 2: three quarters of the chips are H and one quarter are T
- The flip is a draw from a big box with the two urns inside

Bayes' Theorem: one more example

Here is a neat example from Botond Kőszegi via Vincent Crawford:

- Coin flip; your prior is that it's fair (50% heads) with probability $\frac{2}{3}$ and biased (75% heads) with probability $\frac{1}{3}$
- It comes up heads: how much should you decrease your belief that the coin is fair?
- Imagine two urns with poker chips
    1. Urn 1: half of the chips are H and half are T
    2. Urn 2: three quarters of the chips are H and one quarter are T
- The flip is a draw from a big box with the two urns inside
- The prior of $\frac{2}{3}$ means that we should think of urn 1 as having twice as many total chips as urn 2
- So let's say: fair urn has 8 chips, 4 of each type; biased urn has 4 chips, 3 heads and one tail

Bayes' Theorem: one more example

Here is a neat example from Botond Kőszegi via Vincent Crawford:

- Coin flip; your prior is that it's fair (50% heads) with probability $\frac{2}{3}$ and biased (75% heads) with probability $\frac{1}{3}$
- It comes up heads: how much should you decrease your belief that the coin is fair?
- Imagine two urns with poker chips
  1. Urn 1: half of the chips are H and half are T
  2. Urn 2: three quarters of the chips are H and one quarter are T
- The flip is a draw from a big box with the two urns inside
- The prior of $\frac{2}{3}$ means that we should think of urn 1 as having twice as many total chips as urn 2
- So let's say: fair urn has 8 chips, 4 of each type; biased urn has 4 chips, 3 heads and one tail
- Posterior prob. of the chip coming from the fair urn is $\frac{4}{7}$: 7 ways to get H of which 4 were from the fair urn

Spooky predictability

1. Write down a number between 0 and 9
2. Write down the name of a color

Spooky predictability

1. Write down a number between 0 and 9
2. Write down the name of a color

- Simon (1971), Simon & Primavera (1972): the number 7 and the color blue are far and away the most common answers

Randomization

There is evidence that people are bad at randomizing

- This matters because the extent to which one can mimic randomization might be related to what beliefs we form when we see the results of a randomization
- In general it seems that people are not that great at intuitively grasping what randomization 'looks like'

Gambler's fallacy

Notion that there is negative autocorrelation in a non-autocorrelated sequence

- Croson and Sundali (2005): flipping a fair coin, seeing three heads in a row, and so holding a subjective probability that the chance of heads on the next flip is less than 50%
- 'He hasn't had a hit in a few games, he's due for one'
- Tversky and Kahneman (1971): 'law of small numbers' a.k.a. representativeness heuristic
  - Belief that short sequences 'should' be representative of the whole
  - HHHTTT seems 'not random' and HHHHTH seems 'not representative' relative to HTHTTH
- Clotfelter and Cook (1991, 1993): after a lottery number wins people are less likely to bet on it for a while

Roulette data (Croson and Sundali 2005)

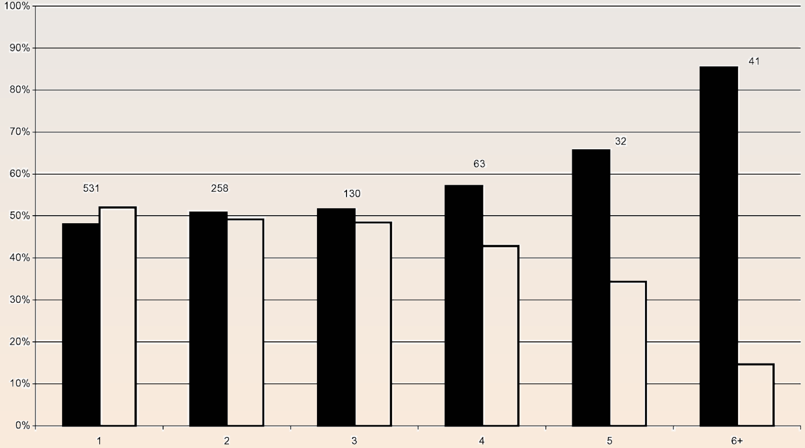


Figure 4. Proportion of gambler's fallacy outside bets after a streak of at least length N.

The law of small numbers

- When asked to produce sequences of random numbers, people switch more often than chance would require
- Rabin (2002) constructs conditional probabilities based on a randomization task experiment in Rapaport and Budescu (1997):

$$\begin{array}{rl} \Pr(A|B) & 58.5\% \\ \Pr(A|AB) & 46.0\% \\ \Pr(A|AAB) & 38.0\% \\ \Pr(A|AAA\ldots) & 29.8\%. \end{array}$$

- Rabin (2002) develops a model of a believer in the law of small numbers
  - Overinfers that the rate of signals is more extreme than it really is; infers more variation among different sources than there really is
  - Application: investors will believe in nonexistent variation in mutual fund manager quality/ability

Field evidence for the gambler's fallacy

Chen, Moskowitz, & Shue (2016): field evidence for the gambler's fallacy

- Three settings:
  1. Refugee asylum court decisions in the US
     - Are judges more likely to deny asylum after granting it to the previous applicant?
     - Random assignment to judges and control for time variation, but don't know about the merits of cases

Field evidence for the gambler's fallacy

Chen, Moskowitz, & Shue (2016): field evidence for the gambler's fallacy

- Three settings:
  1. Refugee asylum court decisions in the US
     - ★ Are judges more likely to deny asylum after granting it to the previous applicant?
     - ★ Random assignment to judges and control for time variation, but don't know about the merits of cases
  2. Loan application reviews
     - ★ Are loan officers more likely to deny a loan application after approving the previous application?
     - ★ Order was randomized in this sample by an experimenter, true loan quality (and thus mistakes) are observed, pay scheme is varied, and payoffs to loan officers depend only on accuracy

Field evidence for the gambler's fallacy

Chen, Moskowitz, & Shue (2016): field evidence for the gambler's fallacy

- Three settings:
  1. Refugee asylum court decisions in the US
     - ★ Are judges more likely to deny asylum after granting it to the previous applicant?
     - ★ Random assignment to judges and control for time variation, but don't know about the merits of cases
  2. Loan application reviews
     - ★ Are loan officers more likely to deny a loan application after approving the previous application?
     - ★ Order was randomized in this sample by an experimenter, true loan quality (and thus mistakes) are observed, pay scheme is varied, and payoffs to loan officers depend only on accuracy
  3. Major League Baseball umpire calls on pitches
     - ★ Are umpires more likely to call a ball after calling a strike?
     - ★ Precise location data is available, but sequencing is not random

Chen, Moskowitz, & Shue (2016)

1. Refugee asylum court decisions in the US
    - Up to 3.3pp more likely to reject if previous case was approved
    - 2% of decisions reversed due to sequencing, stronger effect after longer sequence
    - Similar characteristics or closeness in time of cases makes the effect stronger; experience of the judge lessens it

Chen, Moskowitz, & Shue (2016)

1. Refugee asylum court decisions in the US
   - Up to 3.3pp more likely to reject if previous case was approved
   - 2% of decisions reversed due to sequencing, stronger effect after longer sequence
   - Similar characteristics or closeness in time of cases makes the effect stronger; experience of the judge lessens it
2. Loan application reviews
   - Up to 9% of decisions reversed due to negative autocorrelation
   - Effect is weaker when incentives are stronger and DM less moderate, stronger after a longer sequence
   - Education, age, experience, time spent reviewing all reduce the effect

Chen, Moskowitz, & Shue (2016)

1. Refugee asylum court decisions in the US
   - Up to 3.3pp more likely to reject if previous case was approved
   - 2% of decisions reversed due to sequencing, stronger effect after longer sequence
   - Similar characteristics or closeness in time of cases makes the effect stronger; experience of the judge lessens it
2. Loan application reviews
   - Up to 9% of decisions reversed due to negative autocorrelation
   - Effect is weaker when incentives are stronger and DM less moderate, stronger after a longer sequence
   - Education, age, experience, time spent reviewing all reduce the effect
3. Major League Baseball umpire calls on pitches
   - 1.5pp less likely to call a strike if the previous pitch was a strike, holding location fixed
   - Effect more than doubles on ‘close calls’ and is stronger after a longer sequence

The hot hand fallacy

Notion that a random process sometimes enters a 'hot' state making it likelier than normal

- A *lot* of papers have litigated and relitigated whether a hot hand exists in sports, particularly basketball
- Reconciling with the gambler's fallacy?
  - DM who exhibits the gambler's fallacy would think that a process has too many streaks and so might resort to the hot hand to 'explain' them
  - Or could simply be belief that *people* can get a hot hand but *processes* are negatively autocorrelated
- Miller and Sanjurjo (2018) show that a 'streak selection bias' in analyzing data to look for hot hands has caused a lot of this research to have incorrectly concluded that hot hands don't exist

Kahneman & Tversky (1974): insensitivity to sample size

A certain town is served by two hospitals. In the larger hospital about 45 babies are born each day, and in the smaller hospital about 15 babies are born each day. As you know, about 50 percent of all babies are boys. However, the exact percentage varies from day to day. Sometimes it may be higher than 50 percent, sometimes lower.

For a period of 1 year, each hospital recorded the days on which more than 60 percent of the babies born were boys. Which hospital do you think recorded more such days?

- The larger hospital (21)
- The smaller hospital (21)
- About the same (that is, within 5 percent of each other) (53)

The values in parentheses are the number of undergraduate students who chose each answer.

Correlation neglect

Enke and Zimmerman (2019): series of experiments to test for correlation neglect

- Idea is that people might get many sources of information that are correlated, but treat them as if they're uncorrelated
- Example: news media

Correlation neglect

Enke and Zimmerman (2019): series of experiments to test for correlation neglect

- Idea is that people might get many sources of information that are correlated, but treat them as if they're uncorrelated
- Example: news media
- Experiment 1: subjects estimate unknown state, paid for accuracy
  - Computer-generated unbiased noisy information is provided
  - Between-subjects design: one group gets correlated information and the other uncorrelated; Bayesian posterior identical
- Experiment 2: embed this in a financial asset trading setting
  - Value of asset is equal to the true state in exp. 1; some groups receive correlated signals, some uncorrelated
- Experiment 3: subjects extract information from newspaper articles
  - This is to see how the effect is in a more natural setting

Enke and Zimmerman (2019)

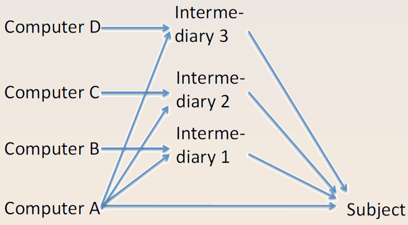


Figure 1: Correlated (left panel) and uncorrelated (right panel) information structure

Enke and Zimmerman (2019)

True state	Computer A	Intermed. 1 uncorr.	Intermed. 2 uncorr.	Intermed. 3 uncorr.	Intermed. 1 corr.	Intermed. 2 corr.	Intermed. 3 corr.	Bayesian belief	Correlation neglect belief
10	12	9	10	0	10.5	11	6	7.75	9.88
88	122	90	68	5	106	95	64	71.25	96.63
250	179	295	288	277	237	234	228	259.75	219.38
732	565	847	650	1'351	706	608	958	853.25	709.13
1'000	1'100	1'060	629	1'100	1'085	870	1'105	974.75	1'042.38
4'698	1'608	7'240	4'866	5'526	4'424	3'237	3'567	4'810.00	3'209.00
7'338	9'950	1'203	11'322	11'943	5'577	10'636	10'947	8'604.50	9'277.25
10'000	2'543	10'780	6'898	8'708	6'662	4'721	5'626	7'232.25	4'887.63
23'112	15'160	21'806	20'607	47'751	18'483	17'884	31'456	26'331.00	20'745.50
46'422	12'340	32'168	49'841	61'293	22'254	31'091	36'817	38'910.50	25'625.25

The reports of intermediaries 1 through 3 in the uncorrelated condition directly reflect the draws of computers B-D. The Bayesian belief is computed by taking the average of the estimates of computers A-D. The correlation neglect belief is computed assuming $\chi = 1$, i.e., full correlation neglect. Thus, this benchmark is given by the average of the estimate of computer A and the reports of intermediaries 2-4 in the correlated condition. Note that subjects faced the ten rounds in randomized order.

Enke and Zimmerman (2019)

True State	Bayesian Belief	Correlation Neglect Belief	Median Estimate Control Treatment	Median Estimate Correlated Treatment	Ranksum Test (p-value)
10	7.75	9.88	7.95	9.2	0.0002
88	71.25	96.63	71.25	88.15	0.0000
250	259.75	219.38	260	250	0.0064
732	853.15	709.13	850	774	0.0063
1'000	974.75	1'042.38	991	1024	0.0180
4'698	4'810	3'209	4'810	4'500	0.0009
7'338	8'604.5	9'277.25	8'653	9'053.15	0.6983
10'000	7'232.25	4'887.63	7'232	6'395	0.0002
23'112	26'331	20'745.5	25'018	21'555	0.0000
46'422	38'910.5	25'625	38'889	30'054	0.0149

See Table 1 for details of the computation of the Bayesian and the correlation neglect benchmarks. Note that subjects faced the ten rounds in randomized order.

Enke and Zimmerman (2019)

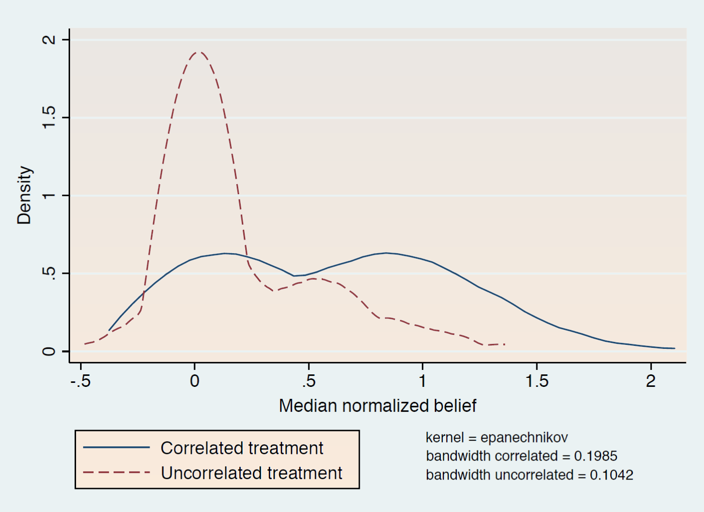


Figure 2: Kernel density estimates of median normalized beliefs

Correlation neglect

- Higher cash incentives for accuracy don't make much difference: subjects try harder but don't do much better
- Higher cognitive ability subjects get closer to Bayes

Correlation neglect

- Higher cash incentives for accuracy don't make much difference: subjects try harder but don't do much better
- Higher cognitive ability subjects get closer to Bayes
- In the market experiment, price bubbles or crashes arise in the correlated treatment depending on the signals
- Individual trading behavior and success reflects correlation neglect: people who have stronger correlation neglect earn worse payoffs

Correlation neglect

- Higher cash incentives for accuracy don't make much difference: subjects try harder but don't do much better
- Higher cognitive ability subjects get closer to Bayes
- In the market experiment, price bubbles or crashes arise in the correlated treatment depending on the signals
- Individual trading behavior and success reflects correlation neglect: people who have stronger correlation neglect earn worse payoffs
- The newspaper article experiment also finds a difference between beliefs in the correlated and uncorrelated cases

Overconfidence in precision

Soll and Klayman (2004) find experimental evidence that people are overconfident in how precise their knowledge is

- 32 subjects, U Chicago undergrads and grad students
- Paid $9 for 45 min experiment (note no direct incentives on the questions answered...)
- Asked to give 80% confidence interval for a numerical question, or for both a number that they are 90% sure the value is above and a number that they are 90% sure the value is below
- For example: given characteristics of a car and had to give those bounds on its price
- The true answer is in the range far less than 80% of the time

Overconfidence in precision

			Interval size relative to norm		
		Hit rate[a]	MEAD/MAD	Gamma simulation	Log simulation
Format					
	Range	39	.45	.46	.44
	Two-point	57	.66	.71	.69
Domain					
	Movie grosses	39	.44	.46	.43
	Basketball wins	46	.57	.55	.54
	Car prices	51	.59	.63	.61
	College ratings	57	.62	.70	.68
Gender					
	Female	58	.70	.74	.72
	Male	40	.45	.47	.45

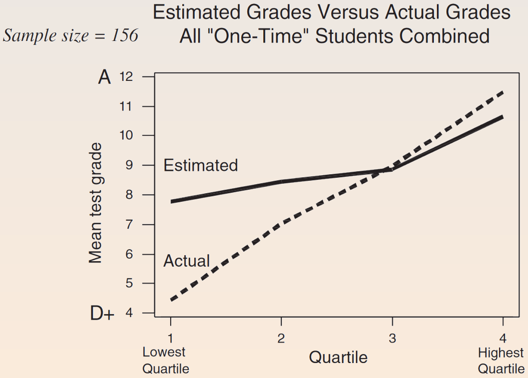
Overconfidence in one's own ability

Soll and Klayman (2004) find experimental evidence that people are overconfident in how precise their knowledge is

- 32 subjects, U Chicago undergrads and grad students
- Paid $9 for 45 min experiment (note no direct incentives on the questions answered...)
- Asked to give 80% confidence interval for a numerical question, or for both a number that they are 90% sure the value is above and a number that they are 90% sure the value is below
- For example: given characteristics of a car and had to give those bounds on its price
- The true answer is in the range far less than 80% of the time

Overconfidence in one's own ability

Kennedy, Lawton, and Plumlee (2002) for student's predicted vs. actual grades:



Overconfidence, competitiveness, and gender

Niederle and Vesterlund (2007): lab experiment on whether men and women of the same ability differ in their selection into a competitive environment

- Groups of 4, public identity but no public performance info until very end
- Task 1: five minute piece rate addition task; task 2: 5 minute addition tournament (only winner in each group paid); task 3: choice of whether to do piece rate or tournament
- No gender differences in task performance, yet 73% of men enter the tournament and only 35% of women
- Why do men enter the tournament more?
    1. Preference?
    2. More overconfidence?
    3. Less risk aversion?
    4. Less averse to feedback?

Overconfidence, competitiveness, and gender

Niederle and Vesterlund (2007): lab experiment on whether men and women of the same ability differ in their selection into a competitive environment

- Groups of 4, public identity but no public performance info until very end
- Task 1: five minute piece rate addition task; task 2: 5 minute addition tournament (only winner in each group paid); task 3: choice of whether to do piece rate or tournament
- No gender differences in task performance, yet 73% of men enter the tournament and only 35% of women
- Why do men enter the tournament more?
  1. Preference?
  2. More overconfidence?
  3. Less risk aversion?
  4. Less averse to feedback?
- Explanations 1 and 2 are found to be important, 3 and 4 negligible

Niederle and Vesterlund (2007)

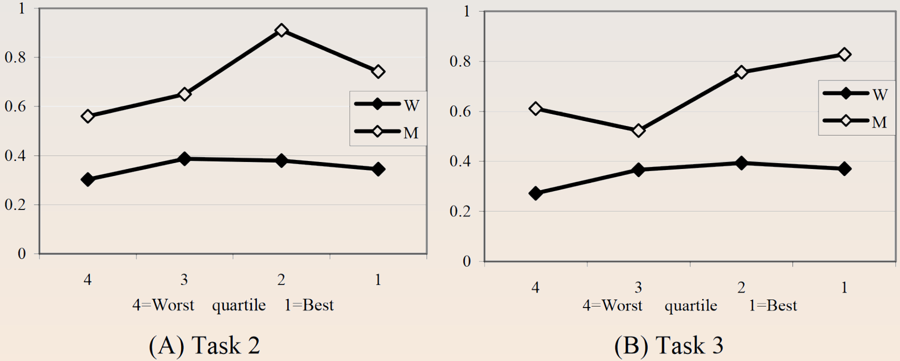


FIGURE II
Proportion of Participants Entering Tournament Conditional on Task-2 Tournament Performance Quartile (panel A) and Task-3 Performance Quartile (panel B).[19]

Motivated beliefs

- A person might manipulate their own beliefs in a way that makes them feel better about their self-image or maintain a core conviction

Motivated beliefs

- A person might manipulate their own beliefs in a way that makes them feel better about their self-image or maintain a core conviction
  - Desire to see oneself as skilful at something
  - Desire to see oneself as morally good
- Why would a DM do this?

Motivated beliefs

- A person might manipulate their own beliefs in a way that makes them feel better about their self-image or maintain a core conviction
  - Desire to see oneself as skilful at something
  - Desire to see oneself as morally good
- Why would a DM do this?
  - Consumption utility from optimism about oneself or the future (Kőszegi 2006, Brunnermeier and Parker 2005)
  - Motivation from optimism to overcome self-control problems (Bénabou and Tirole 2002)
  - Social signaling (Burks et al. 2013, Schwardmann and van der Weele 2017)
- How might a DM do this?

Motivated beliefs

- A person might manipulate their own beliefs in a way that makes them feel better about their self-image or maintain a core conviction
  - Desire to see oneself as skilful at something
  - Desire to see oneself as morally good
- Why would a DM do this?
  - Consumption utility from optimism about oneself or the future (Kőszegi 2006, Brunnermeier and Parker 2005)
  - Motivation from optimism to overcome self-control problems (Bénabou and Tirole 2002)
  - Social signaling (Burks et al. 2013, Schwardmann and van der Weele 2017)
- How might a DM do this?
  - Selective memory (Bénabou and Tirole 2002) or imperfect memory (Gennaioli and Shleifer 2010, 2017)
  - Conservative updating of beliefs (Möbius et al. 2013)
  - Asymmetric updating that puts more weight on good news (Eil and Rao 2011, Sharot et al. 2011)
  - Avoid information that might be bad (Oster et al. 2013, Ganguly and Tasoff 2017)

Motivated beliefs: experiment

Zimmermann (2020) conducts experiments designed to check for motivated beliefs and selective recall

Motivated beliefs: experiment

Zimmermann (2020) conducts experiments designed to check for motivated beliefs and selective recall

- Subjects complete an IQ test, then randomly placed in groups of 10 and asked about their belief on where they rank in the group
- Next, told whether they are ranked higher or lower than each of three randomly chosen members of the group
- For classification in the paper (i.e. not for subjects) feedback is called 'positive' if told they ranked above at least two of the three randomly chosen and 'negative' otherwise

Motivated beliefs: experiment

Zimmermann (2020) conducts experiments designed to check for motivated beliefs and selective recall

- Subjects complete an IQ test, then randomly placed in groups of 10 and asked about their belief on where they rank in the group
- Next, told whether they are ranked higher or lower than each of three randomly chosen members of the group
- For classification in the paper (i.e. not for subjects) feedback is called 'positive' if told they ranked above at least two of the three randomly chosen and 'negative' otherwise
  - Noisy feedback—subjects at the same rank in different groups will, by chance, get different feedback
- Ask again for belief about their rank; first treatment asked immediately, second treatment asked one month later

Motivated beliefs: experiment

Zimmermann (2020) conducts experiments designed to check for motivated beliefs and selective recall

- Subjects complete an IQ test, then randomly placed in groups of 10 and asked about their belief on where they rank in the group
- Next, told whether they are ranked higher or lower than each of three randomly chosen members of the group
- For classification in the paper (i.e. not for subjects) feedback is called 'positive' if told they ranked above at least two of the three randomly chosen and 'negative' otherwise
    - Noisy feedback—subjects at the same rank in different groups will, by chance, get different feedback
- Ask again for belief about their rank; first treatment asked immediately, second treatment asked one month later
- Other treatments run: paid to simply recall the feedback in one month; told in advance they will be paid if they can accurately assess their rank in one month

Zimmermann (2020) findings

Some main findings:

1. Measured right after feedback, beliefs adjust in the correct directions
2. Measured a month after feedback, beliefs after positive feedback are still adjusted upwards, but beliefs after negative feedback is substantially mitigated
   - e.g. you might learn you have a self-control problem in the cold light of day, then lapse back into naivety
3. Negative feedback is recalled with significantly lower accuracy
   - Interventions to correct biases or misperceptions could be undermined by people's ability to forget or suppress threatening information
4. Announcement of future task to assess rank reduces the extent to which negative feedback is 'forgotten'
5. For high enough incentives, subjects are willing to recall negative feedback
   - Suppression rather than erasing memories?

Motivated beliefs to dodge morals

Exley (2016) studies whether potential donors to charity use risk as an excuse not to give

- This would be a motivated belief that 'justifies' selfish behavior
- Subjects: 100 undergraduates at Stanford
- First: 'normalization' price list to elicit the x such that DM is indifferent between \$10 for themselves or x for charity
    - (up to \$30 for charity; 42 of 100 participants chose \$10 for themselves always and were dropped in the main analysis)

Motivated beliefs to dodge morals

Exley (2016) studies whether potential donors to charity use risk as an excuse not to give

- This would be a motivated belief that 'justifies' selfish behavior
- Subjects: 100 undergraduates at Stanford
- First: 'normalization' price list to elicit the x such that DM is indifferent between \$10 for themselves or x for charity
  - (up to \$30 for charity; 42 of 100 participants chose \$10 for themselves always and were dropped in the main analysis)
- Next: 'buffer' that repeats the same with \$5 for self
- Then: complete 28 price lists, with 2x2 treatment (4 blocks of 7)
  1. Option A is the same on all rows: a self lottery (\$10 for self with pr. p else 0) or a charity lottery (\$x for charity with pr. p else 0)
  2. Option B: either self-certain amount or charity-certain amount increasing down the rows of the price list (\$0 to \$10 in 50 cent increments for self; $\frac{x}{20}$ to x in equal increments for charity)

Exley (2016)

- 2x2 with {self lottery, charity lottery} $\times$ {self certain, charity certain}
- Within each block of 7 lists, only difference is the probability involved in the lottery: $p \in \{0.95, 0.9, 0.75, 0.5, 0.25, 0.1, 0.05\}$

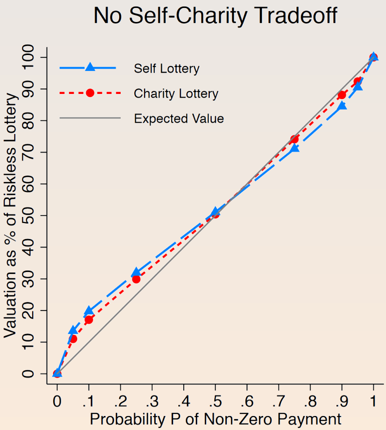
Exley (2016)

- 2x2 with $\{$self lottery, charity lottery$\} \times \{$self certain, charity certain$\}$
- Within each block of 7 lists, only difference is the probability involved in the lottery: $p \in \{0.95, 0.9, 0.75, 0.5, 0.25, 0.1, 0.05\}$
- Switch point between the lottery and the sure thing turned into valuation using the midpoint between the two sure things where the switch occurred
- One randomly selected decision is implemented for payment
- 'No self-charity tradeoff': self lottery vs. self certain or charity lottery vs. charity certain
- 'Self-charity tradeoff': self lottery vs. charity certain or charity lottery vs. self certain

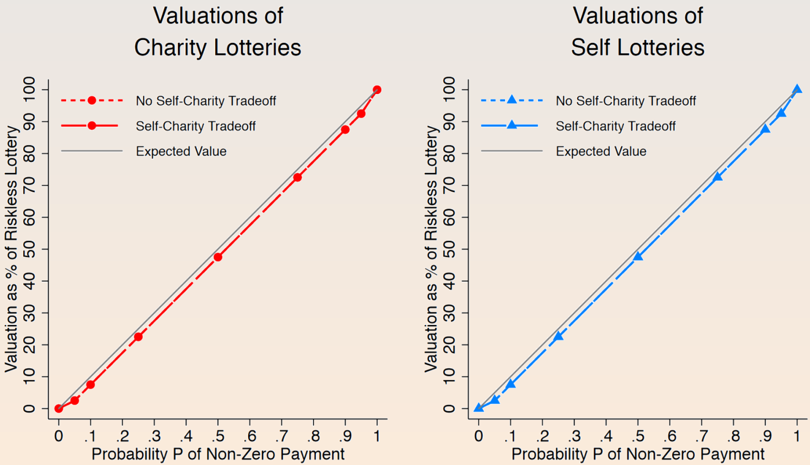
Exley (2016)

- 2x2 with {self lottery, charity lottery} $\times$ {self certain, charity certain}
- Within each block of 7 lists, only difference is the probability involved in the lottery: $p \in \{0.95, 0.9, 0.75, 0.5, 0.25, 0.1, 0.05\}$
- Switch point between the lottery and the sure thing turned into valuation using the midpoint between the two sure things where the switch occurred
- One randomly selected decision is implemented for payment
- 'No self-charity tradeoff': self lottery vs. self certain or charity lottery vs. charity certain
- 'Self-charity tradeoff': self lottery vs. charity certain or charity lottery vs. self certain
- The whole thing is repeated in a partner study to test robustness to generalized pro-social preferences
  - Left Group and Right Group; one randomly selected decision is implemented for each participant in a randomly selected group
  - DMs choose between self payoffs and partner payoffs (reciprocity shouldn't matter since only a partner's decision will be implemented)

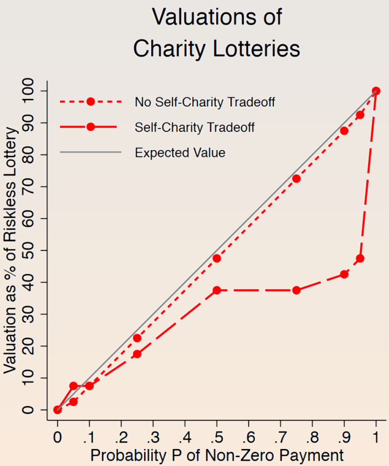
Exley (2016) results

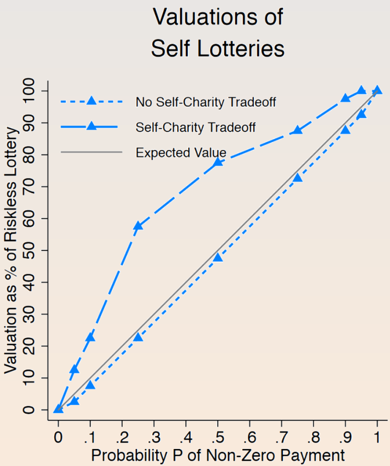


Exley (2016) example subject, not excuse-driven

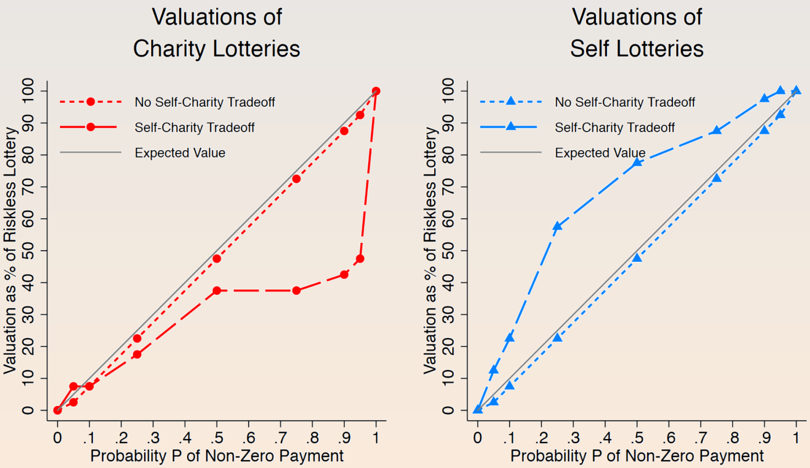


Exley (2016) example subject, excuse-driven





Exley (2016) example subject, excuse-driven



Participants with higher X values ('more selfish') are significantly more likely to be excuse-driven

Surprise donation ask

Exley and Petrie (2018) study whether alerting DM to an upcoming prosocial ask changes the way they respond

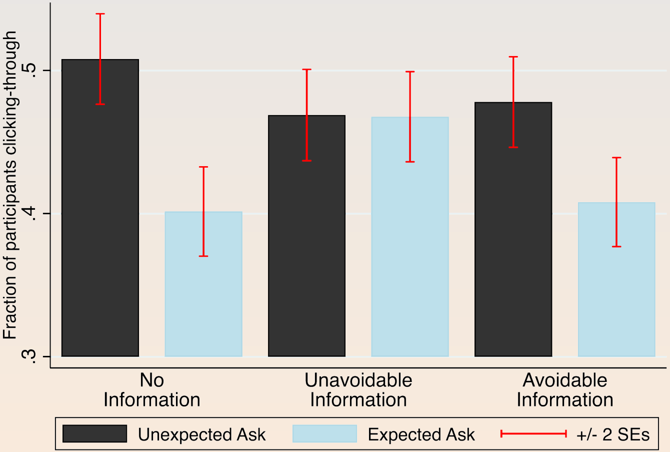
- Online contest: vote for your favorite animal group to get a big cash donation; 6,000 participants
  1. Step 1: vote for favorite group
  2. Step 2: provide info on how they heard about the group and may or may not learn they are about to be asked if they want to make a donation
  3. Step 3: DM is asked whether they want to click through to the group's donation page

Surprise donation ask

Exley and Petrie (2018) study whether alerting DM to an upcoming prosocial ask changes the way they respond

- Online contest: vote for your favorite animal group to get a big cash donation; 6,000 participants
    1. Step 1: vote for favorite group
    2. Step 2: provide info on how they heard about the group and may or may not learn they are about to be asked if they want to make a donation
    3. Step 3: DM is asked whether they want to click through to the group's donation page
- Condition 1 varies the expectation of the ask; adds a tiny '... and if you want, donate to them!' to the instructions for step 2
- Condition 2 varies whether in the second step DM sees no info, unavoidable info, or avoidable info about a dog's adoption story

Exley & Petrie (2018)



In no info case, expectation of the ask reduces click through rates from 0.51 to 0.40, a 22% decline

Motivated beliefs about social norms

Bicchieri, Dimant, and Sonderegger (2020) studies whether DMs distort their beliefs about social norms

- Experiment varies nature of elicited beliefs
  1. Descriptive about what others to
  2. Normative about what others approve of
- And varies whether DM is aware of forthcoming opportunity to lie

Motivated beliefs about social norms

Bicchieri, Dimant, and Sonderegger (2020) studies whether DMs distort their beliefs about social norms

- Experiment varies nature of elicited beliefs
  1. Descriptive about what others to
  2. Normative about what others approve of
- And varies whether DM is aware of forthcoming opportunity to lie
- Compared to DMs who form beliefs before awareness of the lying task, DMs who are aware of upcoming lying task appear to distort their beliefs in the descriptive beliefs case:
  1. More likely to believe that lying is widespread
  2. More likely to lie
- But no distortion in normative beliefs case

Conjunction effect (Tversky & Kahneman 1983)

"The Bill problem"

Bill is 34 years old. He is intelligent, but unimaginative, compulsive, and generally lifeless. In school, he was strong in mathematics but weak in social studies and humanities.

Conjunction effect (Tversky & Kahneman 1983)

"The Bill problem"

Bill is 34 years old. He is intelligent, but unimaginative, compulsive, and generally lifeless. In school, he was strong in mathematics but weak in social studies and humanities.

Rank the following according to their probability:

- Bill is a physician who plays poker for a hobby
- Bill is an architect
- Bill is an accountant (A)
- Bill plays jazz for a hobby (J)
- Bill surfs for a hobby
- Bill is a reporter
- Bill is an accountant who plays jazz for a hobby (A&J)
- Bill climbs mountains for a hobby

Conjunction effect (Tversky & Kahneman 1983)

"The Bill problem"

Bill is 34 years old. He is intelligent, but unimaginative, compulsive, and generally lifeless. In school, he was strong in mathematics but weak in social studies and humanities.

Rank the following according to their probability:

- Bill is a physician who plays poker for a hobby
- Bill is an architect
- Bill is an accountant (A)
- Bill plays jazz for a hobby (J)
- Bill surfs for a hobby
- Bill is a reporter
- Bill is an accountant who plays jazz for a hobby (A&J)
- Bill climbs mountains for a hobby

87% of a sample of 88 undergraduates ranked A over A&J over J

Conjunction effect

An even more stark demonstration:

"The Linda problem"

Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in antinuclear demonstrations.

Conjunction effect

An even more stark demonstration:

"The Linda problem"

Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in antinuclear demonstrations.

Rank the following according to their probability:

- Linda is a bank teller (T)
- Linda is a bank teller and is active in the feminist movement (T&F)

Conjunction effect

An even more stark demonstration:

"The Linda problem"

Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in antinuclear demonstrations.

Rank the following according to their probability:

- Linda is a bank teller (T)
- Linda is a bank teller and is active in the feminist movement (T&F)

85% of a sample of 142 undergraduates ranked T&F as more probable than T, in violation of the rule that a conjoint event cannot be more probable than any of the individual events that constitute it

Conjunction effect

One class of explanations for the conjunction effect focuses on the fuzzy meaning of ‘probable’ in natural language

- Participants might assume that what they are being told is relevant, which is quite a natural thing to assume about conversations in real life
- If you add more detail and elaboration to a conjecture, it may be found more convincing by your audience—this is a possibility that could be useful for good or for evil...
- Tversky & Kahneman (1983) goes to increasingly elaborate lengths to ‘erase’ the conjunction effect (different ways of posing the question, different levels of statistical education among participants)
- Asking participants which they would rather bet on reduced the violations of conjunction to 56%
- Among 64 social science grad students at Berkeley and Stanford (who had taken statistics courses) the violation was displayed by 36%

Conjunction effect

Another example described in Kahneman (2011):

Tennis

Rank four possible outcomes of the next Wimbledon tournament from most to least probable:

- Borg will win the match
- Borg will lose the first set
- Borg will lose the first set but win the match
- Borg will win the first set but lose the match

Conjunction effect

Another example described in Kahneman (2011):

Tennis

Rank four possible outcomes of the next Wimbledon tournament from most to least probable:

- Borg will win the match
- Borg will lose the first set
- Borg will lose the first set but win the match
- Borg will win the first set but lose the match

72% of subjects assigned a lower probability to the second option than the third

Conjunction effect

Hertwig & Chase (1998) ask a Linda Problem with options B, F, and B&F and vary the response mechanism

- Subjects: 72 students from University of Chicago
- Randomly assigned to a ranking group or estimation group
- Ranking group asked to rank the event probabilities, estimation group asked to estimate them

Conjunction effect

Hertwig & Chase (1998) ask a Linda Problem with options B, F, and B&F and vary the response mechanism

- Subjects: 72 students from University of Chicago
- Randomly assigned to a ranking group or estimation group
- Ranking group asked to rank the event probabilities, estimation group asked to estimate them
- In the ranking group, 78% violated the conjunction rule
- In the estimation group, 42% violated the conjunction rule

Conjunction effect

Hertwig & Chase (1998) ask a Linda Problem with options B, F, and B&F and vary the response mechanism

- Subjects: 72 students from University of Chicago
- Randomly assigned to a ranking group or estimation group
- Ranking group asked to rank the event probabilities, estimation group asked to estimate them
- In the ranking group, 78% violated the conjunction rule
- In the estimation group, 42% violated the conjunction rule
- Primed to think about probability as 'math-y'?
- Hertwig & Gigerenzer (1997) asked German-speaking participants in the Linda problem to paraphrase 'probability' for non-native speakers; 82% of the paraphrases were non-mathematical
- Participants may be evaluating the hypotheses rather than the evidence

Ameliorating conjunction effects

In Charness, Karni, and Levin (2008) communication among subjects and small cash incentives reduce the likelihood of observing the conjunction fallacy

- Is the effect just an artifact of casual consideration of a hypothetical problem?
- Again we run into the hypothetical questions problem
- And issues of learning / teaching in external validity of experimental regularities

Portioning probability (Tversky & Koehler 1994)

- Participants: 196 Stanford undergraduates
- Task: write down the last digit of your telephone number, and evaluate what percentage of U.S. married couples have exactly that many children
- Prize: the three most accurate respondents get $10 each

Portioning probability (Tversky & Koehler 1994)

- Participants: 196 Stanford undergraduates
- Task: write down the last digit of your telephone number, and evaluate what percentage of U.S. married couples have exactly that many children
- Prize: the three most accurate respondents get \$10 each
- Total of the means assigned by each group: 199%
- Total of the medians: 180%
- Overestimated every category except for 0 children
- Largest overestimate was for 2 children
- Sum of the means for 0, 1, 2, and 3 children: 145%

Portioning (Tversky & Koehler 1994)

- Participants: 139 undergraduates
- Task: estimate the percentage of U.S. married couples with "less than 3", "3 or more", "less than 5", "5 or more" children (each subject was assigned only one of these)

Portioning (Tversky & Koehler 1994)

- Participants: 139 undergraduates
- Task: estimate the percentage of U.S. married couples with "less than 3", "3 or more", "less than 5", "5 or more" children (each subject was assigned only one of these)
- Total of the estimates: 97.5% for the first pair of hypotheses
- 96.3% for the second pair of hypotheses
- This 'binary complementarity' does not display the subadditivity effect

Dopamine and reward prediction error

Glimcher (2011) surveys empirical evidence for the link between the activity of dopamine neurons and reinforcement learning mechanisms

- Activity of midbrain dopamine neurons as a mechanism for reward-driven learning in animal behavior
- Believed to signal a misestimation of value of current or future events
- Signal adjusts synaptic strength until the value of current and future events is accurately encoded

Dopamine and reward prediction error

Glimcher (2011) surveys empirical evidence for the link between the activity of dopamine neurons and reinforcement learning mechanisms

- Activity of midbrain dopamine neurons as a mechanism for reward-driven learning in animal behavior
- Believed to signal a misestimation of value of current or future events
- Signal adjusts synaptic strength until the value of current and future events is accurately encoded
- Pavlov type experiments; e.g. Schultz et al. (1993)
    - Monkey seated at two levers
    - One gives apple juice reward, the other nothing
    - At first, push levers erratically after start cue, neurons respond to reward
    - After training, push the reward lever only and neurons respond to the start cue but not the juice

Dopamine and reward prediction error

Glimcher (2011) surveys empirical evidence for the link between the activity of dopamine neurons and reinforcement learning mechanisms

- Activity of midbrain dopamine neurons as a mechanism for reward-driven learning in animal behavior
- Believed to signal a misestimation of value of current or future events
- Signal adjusts synaptic strength until the value of current and future events is accurately encoded
- Pavlov type experiments; e.g. Schultz et al. (1993)
  - Monkey seated at two levers
  - One gives apple juice reward, the other nothing
  - At first, push levers erratically after start cue, neurons respond to reward
  - After training, push the reward lever only and neurons respond to the start cue but not the juice
- Schultz et al. (1997)
  - Reward cued and delivered, neurons respond to cue not reward
  - Reward cued and not delivered, neurons respond to cue and neuron activity suppressed when reward would have arrived

Schultz et al. (1993)

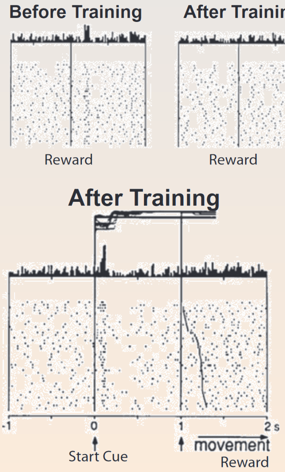


Fig. 2. "Raster plot of dopamine neuron activity. Upper panel shows response of dopamine neuron to reward before and after training. Lower panel shows response of dopamine neuron to start cue after training" (26). [Reproduced with permission from ref. 26 (Copyright 1993, Society for Neuroscience).]

Schultz et al. (1997)

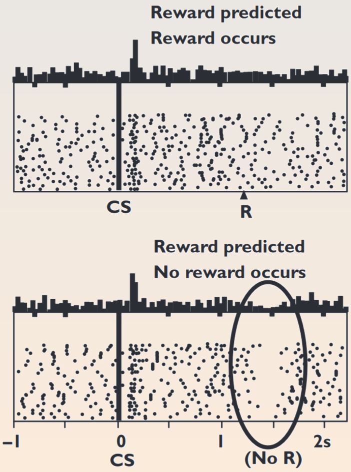


Fig. 3. "When a reward is cued and delivered, dopamine neurons respond only to the cue. When an expected reward is omitted after a cue the neuron responds with a suppression of activity as indicated by the oval" (29). [Reproduced with permission from ref. 29 (Copyright 1997, American Association for the Advancement of Science).]

Response to news

Any significance here for economics?

- Is value/enjoyment/utility derived from levels or changes?
- Possible relationship to the importance of reference points?
- Along similar lines is 'focusing illusion in affective forecasting: "nothing in life matters quite as much as it does while you are thinking about it" (Schkade & Kahneman 1998)
- Examples: how will you feel if you move from the midwest to California (Schkade & Kahneman 1998); how will you feel if you don't get tenure (Gilbert et al. 1998)
- "... as the new state loses its novelty it ceases to be the exclusive focus of attention, and other aspects of life again evoke their varying hedonic responses."

Projection bias

Projection bias is when one's beliefs are biased towards how things currently are relative to how things will be

- The canonical example here is the hungry shopper
- Grocery shoppers tend to purchase food as if their current hunger level will last forever
  - Shoppers who get a muffin to eat before they shop are better at only buying the things on their shopping list (Gilbert, Gill, and Wilson 1998)
- Read and Van Leeuwen (1998): office workers could choose a snack to be delivered later in the week; more likely to choose an unhealthy snack when they made the choice while hungry than right after lunch
- Read, Loewenstein, aand Kalyanaraman (1999): subjects could choose a free movie rental either tonight or tomorrow; choose highbrow movies for the next day but lowbrow movies for tonight

Diversification bias

Along similar lines there is a tendency for people to choose more variety if they make combined choices of quantities for future consumption than if they choose sequentially immediately before consuming

- e.g. if I ask you to choose an ice cream flavor each week for the next 10 weeks vs. ask you to choose *now* what ice cream flavor you want for each of the next 10 weeks...

Diversification bias

Along similar lines there is a tendency for people to choose more variety if they make combined choices of quantities for future consumption than if they choose sequentially immediately before consuming

- e.g. if I ask you to choose an ice cream flavor each week for the next 10 weeks vs. ask you to choose *now* what ice cream flavor you want for each of the next 10 weeks...
- If you choose more variety in the second case, you have exhibited diversification bias
- Simonson (1990) experiment: snack choice for 3 days
  - Simultaneous choosers: 64% chose three different items
  - Sequential choosers: 9% chose three different items
- Simonson & Winer (1992) got similar effects in a real-world choice setting

Diversification bias

Is this 'standard model' economic behavior?

- Standard: convex indifference curves means preference for variety
- Standard: preferences are stable
- The conjunction is were things break down
- Read & Loewenstein (1995): why diversification?
  - No-bias hypotheses
    1. Risk aversion plus preference uncertainty
    2. Information seeking
    3. Cognitive limitations e.g. imperfect recall
  - Bias hypotheses
    1. Time contraction: shrink the interconsumption period, anticipate too much satiation
    2. Choice bracketing: framing as a portfolio choice
  - Experiment 1: 7 different treatments designed to disentangle explanations

Read & Loewenstein (1995)

Variety Seeking in All Conditions of Experiment 1

Group	One %	One n	Two %	Two n	Three %	Three n	M
	Number of different kinds chosen						
Sequential choice	46	19	46	19	8	3	1.61
Sequential–aware	50	25	34	17	16	8	1.66
Prediction	38	18	31	15	31	15	1.94
Simultaneous choice	18	10	38	21	45	25	2.27
Simultaneous–pretaste	17	8	38	18	46	22	2.29
Simultaneous–post-taste	20	9	45	20	34	15	2.14
Simultaneous after immediate	17	9	31	16	52	27	2.35

Read & Loewenstein (1995)

- Experiment 2: deeper into the time contraction hypothesis:
  - First treatment: first imagine choosing for one per week, then for one per day
  - Second treatment: order reversed
- Little evidence of sensitivity to the length of the interval unless made salient by highlighting the contrast
- Experiment 3: identical except day-week condition was hypothetical per-day choices followed by real per-week choices, and week-only condition chose only for the per-week case

Read & Loewenstein (1995)

- Experiment 2: deeper into the time contraction hypothesis:
  - First treatment: first imagine choosing for one per week, then for one per day
  - Second treatment: order reversed
- Little evidence of sensitivity to the length of the interval unless made salient by highlighting the contrast
- Experiment 3: identical except day-week condition was hypothetical per-day choices followed by real per-week choices, and week-only condition chose only for the per-week case
- Experiment 4: deeper into choice bracketing hypothesis
- Trick-or-treaters: neighboring houses; two big piles of candy bars (Three Musketeers and Milky Way)
- Some chose two candy bars at one house; others chose one candy bar at both houses
- All children in the combined choice case chose different candy bars; only 48% did in the separate choice case

Read & Loewenstein (1995)

Variety Seeking in All Conditions of Experiments 2 and 3: Tests of Time Contraction Hypothesis

	Number of different kinds chosen by interconsumption interval													
	Day							Week						
	One		Two		Three			One		Two		Three		
Condition	%	*n*	%	*n*	%	*n*	*M*	%	*n*	%	*n*	%	*n*	*M*
Experiment 2														
Day → week	13	8	42	26	45	28	2.32	35	22	39	24	26	16	1.90
Week → day	25	15	28	17	47	28	2.22	27	16	30	18	43	26	2.17
Experiment 3														
Day → week	20	6	33	10	47	14	2.27	33	10	33	10	33	10	2.00
Week only								16	5	38	12	45	14	2.29

Projection bias: a simple model

Loewenstein, O'Donoghue, and Rabin (2003):

$$\hat{u}(c, s) = (1 - \alpha)u(c, s) + \alpha u(c, s') \tag{6}$$

- State right now is s' and the individual is trying to predict their utility in future state s
- Higher α means more projection bias

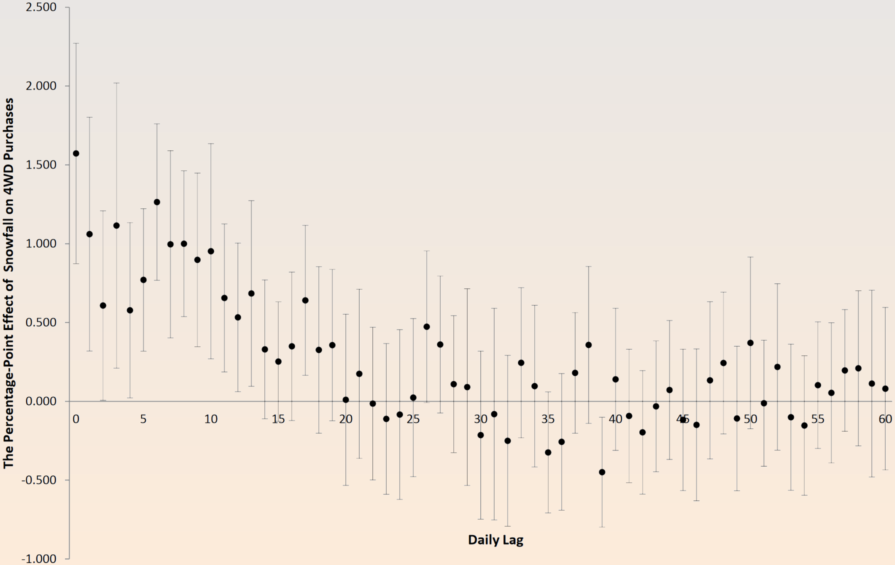
Projection bias: a simple model

Loewenstein, O'Donoghue, and Rabin (2003):

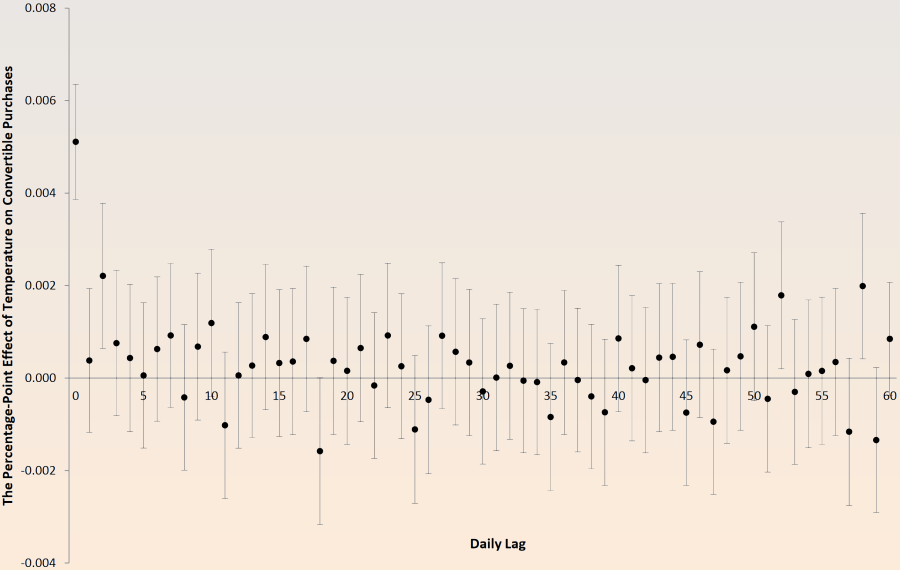
$$\hat{u}(c,s) = (1-\alpha)u(c,s) + \alpha u(c,s') \tag{6}$$

- State right now is s' and the individual is trying to predict their utility in future state s
- Higher α means more projection bias
- Chang, Huang, and Wang (2018): health insurance take-up significantly higher and policy cancellations significantly lower on days with higher air pollution
- In that paper: 180 day waiting period for insurance to kick in means that there 'should' be no impact of day-to-day fluctuations
- In general, though: how can you disentangle projection bias from other (possibly rational) explanations?

Busse et al. (2015)



Busse et al. (2015)



Persuasion

Following DellaVigna & Kaplan (2007), the persuasion rate in percent terms for a binary behavioral outcome can be expressed

$$f = 100 \times \frac{y_T - y_C}{e_T - e_C} \frac{1}{1 - y_0} \tag{7}$$

where

- T and C refer to the treatment and control groups
- e_i is the share of group i that receives the message
- y_i is the share of group i that adopts the behavior
- y_0 is the share that would adopt without the message (can be approximated by y_C if y_0 not known, as long as e_C or $(y_T - y_C)$ are small)

Persuasion rate example

From DellaVigna & Gentzkow (2010):

- Say that get-out-the-vote (GOTV) mailer increases turnout among treated by 1 percentage point relative to control
- Say 10% of voters in treatment group got the mailer
- Say the targeted population already had an 80% turnout rate

$$f = 100 \times \frac{y_T - y_C}{e_T - e_C} \frac{1}{1 - y_0} \tag{8}$$

$$= 100 \times \frac{0.01}{0.1} \frac{1}{0.2} \tag{9}$$

$$= 50\% \tag{10}$$

A change in behavior that seems quite small can actually reflect a quite high impact of persuasion

Models for persuasion

1. Belief-based
   - Bayesian or non-Bayesian updating (e.g. limited memory models)
   - Larger effects when receiver's prior was weaker
   - Inferences in Bayesian models depend on assessment of sender's credibility
2. Preference-based
   - Advertising entering the utility function directly
   - Framing, salience, signaling models all in a kind of grey area between belief and preference models
   - Message content is often not informative
   - Receivers might take costly steps to avoid exposure to persuasion
3. Demand for information and limited attention
   - Effectiveness of persuasion depends on how the receiver allocates their attention
   - Could be high costs of acquiring information
   - Could be irrational limited attention

Deception and self-deception

Schwardmann and van der Weele (2019): test the idea of strategic self-deception (cf. Trivers 1985, 2011 in evolutionary biology)

- 288 subjects, recruited via Munich Experimental Laboratory for Economic and Social Sciences. Two stages:
  1. Self-deception stage: intelligence test; small payment depending on score relative to three other randomly selected subjects
     - ★ Treatment: 'contestant' group told they can earn 15 euros later by persuading 'employers' about how well they did on the test in a short face-to-face interaction
     - ★ Control group: no opportunity to 'persuade' and given no information about the next experiment stage
  2. Elicit 'confidence' in performance on the test
     - ★ BDM elicitation: asked for prob. such that indifferent between cash prize with that prob. and same prize for sure if in the top 2 of their group
     - ★ Contestants now get noisy feedback on whether they're actually top 2 (that is: informative but sometimes wrong)
  3. Persuasion stage: control group are the 'employers' and they 'interview' the contestants

Deception and self-deception

The interactions followed a speed-dating protocol. Equipped with a pen and evaluation sheets, employers leave their computer stations and sit down in front of the contestants. A group of four employers is matched with a group of four contestants. There are four rounds of interviews so that each of the four employers in a group interviews each of the four contestants in a group. Each interview takes place behind a partition to assure some level of privacy. On the ring of a bell, contestants say one sentence: 'I believe that my performance was in the top 2 of my group with ... per cent probability'. In the blank, each contestant verbally fills in a number between 0 and 100.

After the interviews, employers return to their computer stations and enter their evaluations as well as the verbal messages of the contestants into the computer. Each employer states the probability that they associated with each of the four contestants that they interviewed being in the top 2 of the contestant group. One of these evaluations is selected at random and the employer is paid according to the Becker–DeGroot–Marschak mechanism described above, with a chance to win €10 in case of a favourable outcome. A contestant is paid €15 if he or she received one of the two best evaluations by a randomly selected employer and nothing otherwise.

Deception and self-deception

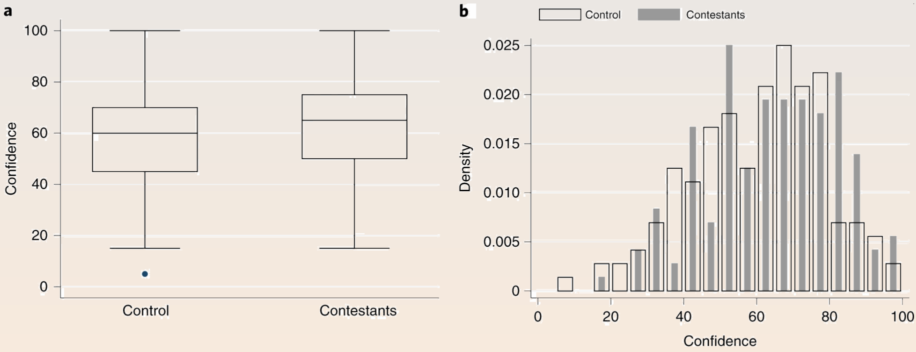
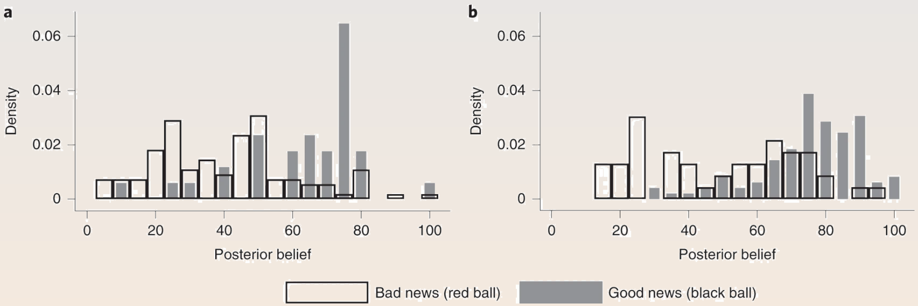


Figure: 'Contestants' think they did better than control: 'self-deception'

Deception and self-deception



a, Posterior beliefs of participants who were not in the top 2 ($n = 144$). **b**, Posterior beliefs of participants in the top 2 ($n = 144$). The *x* axis is the subjective probability of being in the top 2 of the group (scale, 0–100).

Figure: The noisy signal meaningfully shifted the subjects' posterior beliefs in their performance; subjects that became more confident got higher evaluations from 'employers' *controlling for their actual performance*

Field evidence on self-persuasion

Schwardmann, Tripodi, and van der Weele (2019): international debating competition randomly assigns competitors to be for or against a motion—does it change their beliefs?

- A big picture question: impressing and persuading others is part of life; does it change how we form beliefs?
- March 2019: two international debate tournaments in Munich and Rotterdam
- Rules: random assignment of for/against, revealed 15 minutes ahead of the debate
- Debaters answered survey questions: at the start of the event, just before each debate (after 15 minute prep time following random assignment of role), right after each debate, at the end of the first round of the tournament
    1. Factual beliefs: related to the motion, predict true or false
    2. Attitudes: assign money to a neutral charity or one aligned with one side of the motion
    3. Confidence: subjective probability that *other* teams' debates on the same motion will be won by the 'for' side

Field evidence on self-persuasion

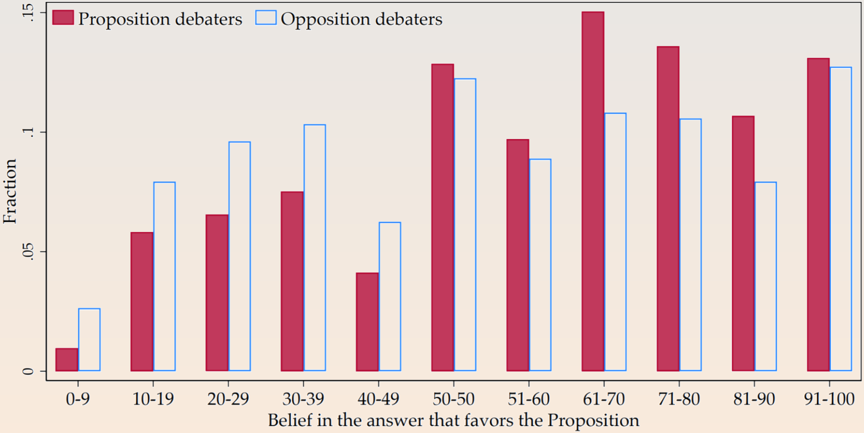


Figure: Debaters are more likely to believe in a given factual statement if it favors the position they were randomly assigned

Field evidence on self-persuasion

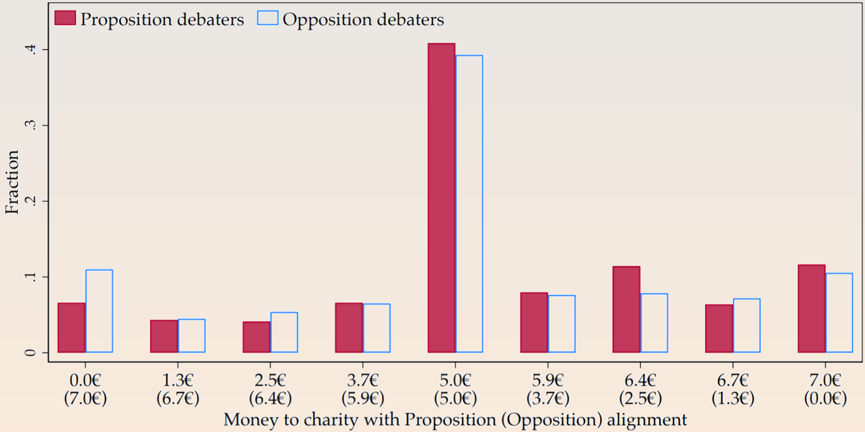


Figure: Debaters slightly more likely to allocate money to a charity that aligns with their side, though the effect is small

Field evidence on self-persuasion

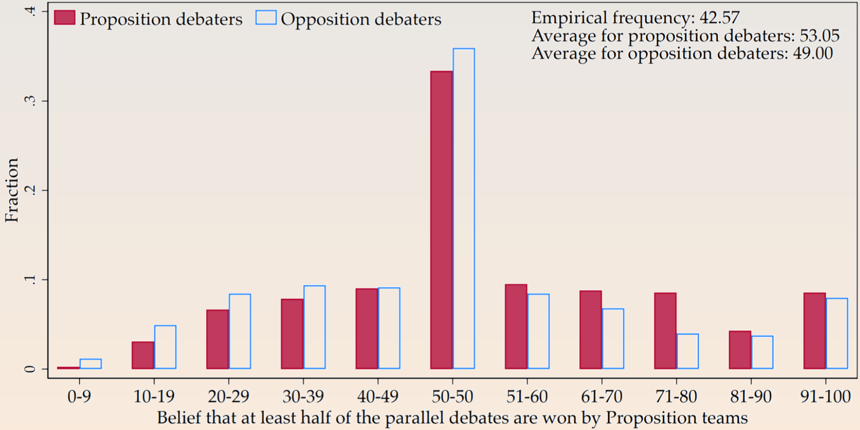


Figure: Debaters slightly more likely to believe that other teams assigned to their side will win their debates

‘Strategically delusional’

Another paper in a similar vein is Soldà, Ke, Page, and von Hippel (2019)

- Very complex lab experiment with 2x3 structure:

	Information conditions		
Treatments	*No Information*	*Given Information*	*Self-Chosen Information*
Accuracy-first	*NI* x *Acc.1st*	*GI* x *Acc.1st*	*SCI* x *Acc.1st*
Persuasion-first	*NI* x *Per.1st*	*GI* x *Per.1st*	SCI x *Per.1st*

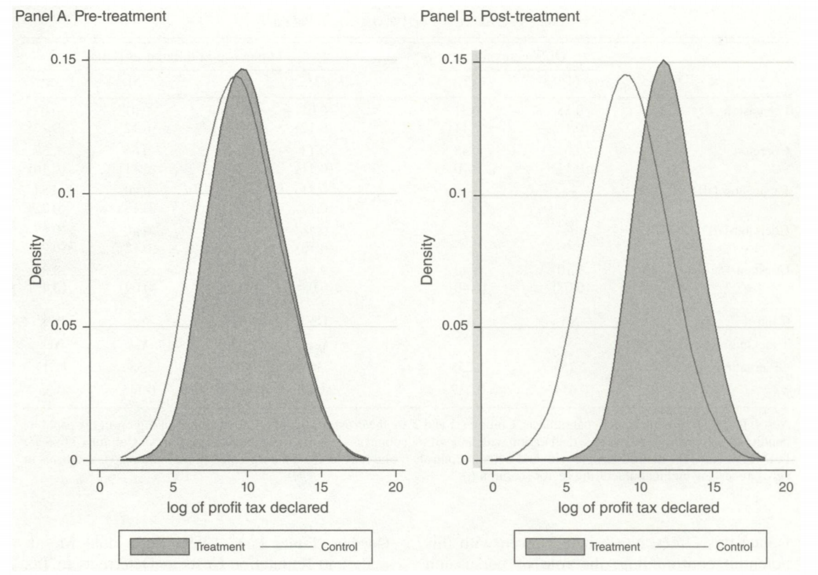
- Subjects:
  1. More likely to overestimate their performance when they'll have to persuade others afterwards
  2. Search for information about their performance as if searching out more positive feedback
  3. Are better at persuading others because of their increased confidence from 1 and 2

Persuasion vs. coercion in the field

Shimeles, Gurara, and Woldeyes (2017): how can a tax collecting authority increase compliance?

- Randomized experiment with the Ethopian Revenue and Customs Authority in Addis Ababa
- 3,120 randomly selected businesses, targeted one month before taxes are due
    1. Group 1: threatening letter; warning about audits, civil and criminal penalties, part of the tax code printed
    2. Group 2: persuasive letter; patriotic duty, loyalty and honesty, projects financed by tax revenue, compliance incentives
    3. Control: no letter
- Threats: profit tax payable up 38%
- Persuasion: profit tax payable up 32%
- Is it persuasion or a polite threat?

Persuasion vs. coercion in the field



Using information

A dialectic model of belief formation:

1. Propose a decision-making heuristic that provides a unified explanation for well-established behavioral anomalies across different
    - Bad at updating beliefs (Ouwersloot et al. 1998, Benjamin 2018)
    - Bad at interpreting probabilities (Kahneman & Tversky 1979)
2. Embed that heuristic in a game of competitive spin that can explain patterns of political messaging and its reception by voters
    - '[T]he best test of truth is the power to get itself accepted in the competition of the market' (Oliver Wendell Holmes 1919)

Model inspired by Froeb, Ganglmair, and Tschantz (2016)

Dialectic model

DM makes a ‘best guess’ about the state of the world

- Domain of possible states of the world $[0, 1]$
- DM puts two frames on the data: two continuous uniform distributions over a subset of the domain
- Frames must be consistent with available evidence but are unconstrained otherwise
  - Optimistic vs. pessimistic
  - Good cop / bad cop
  - The angel and devil on your shoulders
- DM is good at imagining extreme scenarios but less good at shades of grey or probabilistic scenarios
- DM will grasp for a conclusion that resolves both frames
- Hegelian dialectic as a DM heuristic

Dialectic model

- Mean of the two frames μ_L and μ_H; width of the two frames w_L and w_H
- DM's belief is a weighted average of the two means witha penalty for width:

$$\hat{\mu} = \frac{w_H^s}{w_L^s + w_H^s}\mu_L + \frac{w_L^s}{w_L^s + w_H^s}\mu_H \tag{11}$$

- Skepticism parameter $s > 0$
  - Large s: DM heavily discounts the less plausible frame
  - Small s: DM weights the two means more equally and considers plausibility less
- Interpretations? Uncertainty aversion; naivety; gullibility

Application 1: belief formation

- DM observes signals—evidence that must be explained
- Optimistic and pessimistic explanations: optimistic 'prefers' DM to believe closer to 1, pessimistic closer to 0
- Example: 2 pieces of evidence at 0.6 and 0.8
- Frames include the evidence and stretch to 'preferred' endpoints

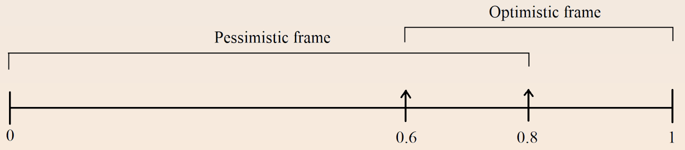
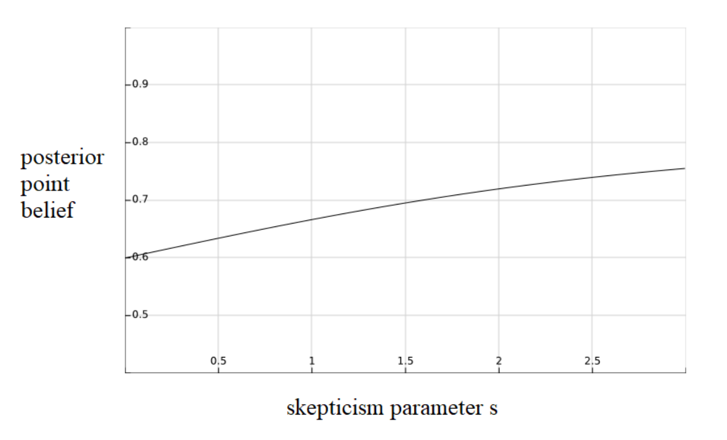


Figure: Pessimistic and optimistic frames with two observations

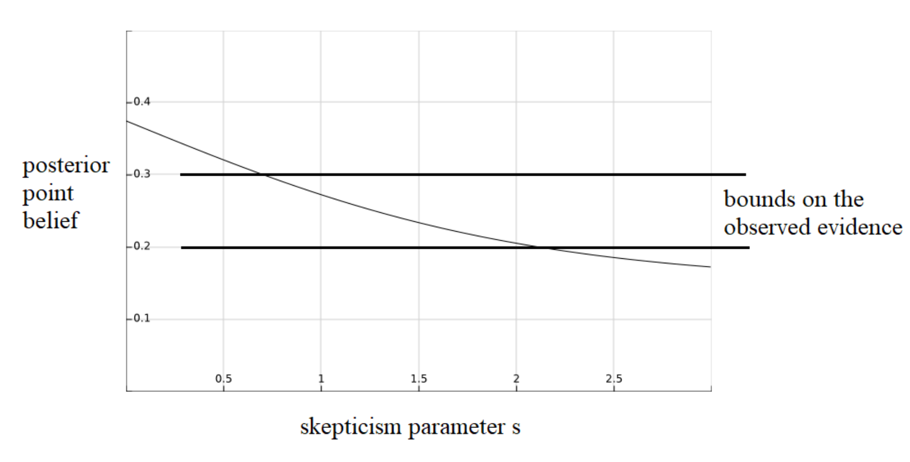
Application 1: belief formation

- DM may over- or under-shoot 0.7, the mean of the ML uniform distribution implied by the evidence



Application 1: belief formation

- Example with evidence 0.2 and 0.3; DM compelled by stories, not evidence: belief can lie outside bounds of the evidence



Application 1: belief formation

Some implications:

- Irrelevance of new reinforcing evidence
  - Consistent with no backlash effect in counter-attitudinal messaging (Guess & Coppock 2018)
- Everyone makes mistakes: there is no universally 'correct' s
  - Some experimental subjects overreact and some underreact relative to Bayes (Epstein et al. 2010)
- If s distributed symmetrically around 'accuracy', aggregate belief is systematically far-fetched
- Beliefs are stubborn
  - Low skepticism makes DM too close to 0.5 than ML estimate
  - Finance application: consistent with disposition effect (Weber & Camerer 1998) and underreaction to news / overreaction to series of news (Barberis et al. 1998)

Application 2: decision weights

- DM observes a stated probability and interprets it

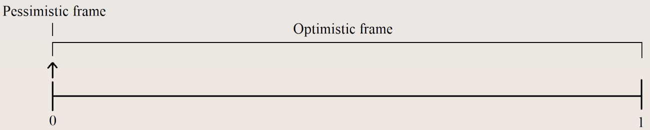


Figure: For stated probability 0 or 1, one frame is a point

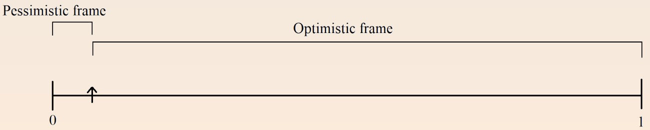


Figure: For stated probability close to an endpoint, both frames have width

Application 2: decision weights

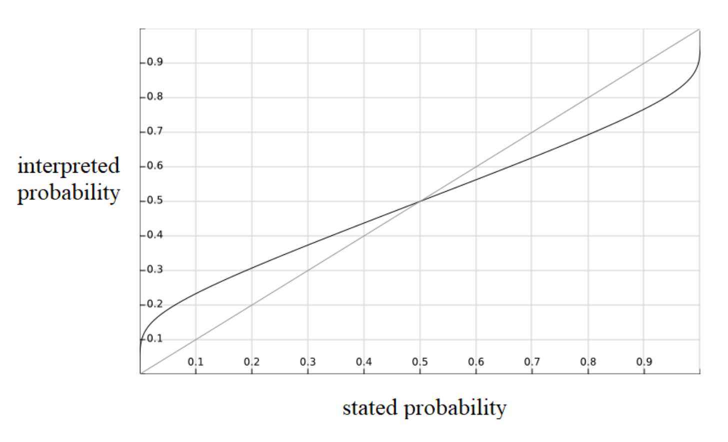


Figure: Interpreted probabilities for $s = 0.25$; S shape well founded in e.g. Gonzalez & Wu (1999), Tversky & Kahneman (1992), Camerer & Ho (1994)

Application 2: decision weights

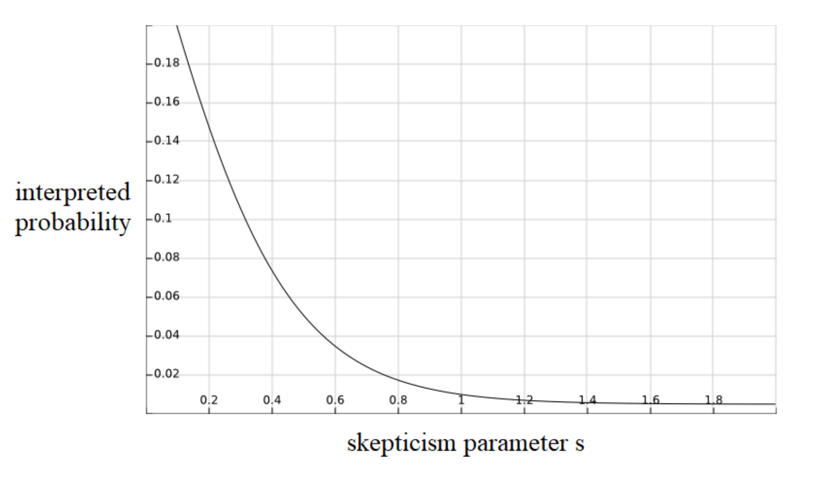


Figure: Interpreted probability for a stated probability of 0.01, by skepticism parameter s

Credulity

- Calibration: excessive credulity best explains evidence for behavioral anomalies
- Similar skepticism parameter in both cases
- Belief formation mistakes are ‘easier’ than decision weight mistakes in the sense of less credulity required

Application 3: political spin

Strategic frame selection by interested parties

- Two parties L and R
- Evidence represents commonly agreed-upon ‘facts’
- Parties choose (contiguous, non-degenerate, uniform distribution) frames; must include the facts
- Receiver is non-strategic and forms belief by the dialectic model (in this context DM’s assessment is similar to Skaperdas and Vaidya 2012 and Hirshleifer and Osborne 2001)
- Payoff to parties is how close belief is to their preferred endpoint

Application 3: political spin

Result

For a sufficiently credulous receiver, with a skepticism parameter of 1 or lower, the unique Nash equilibrium of the political spin game has both senders choose the most extreme possible frame.

- First: never pays to enlarge frame away from own endpoint (increases width, worsens mean)
- Other direction: enlarge frame toward own endpoint?
- Tradeoff is higher width versus moving mean closer to preferred pole
- Credulous DM doesn't discount width much, so tradeoff encourages playing the most extreme possible frame that includes the evidence
- The more extreme the evidence, the more credulity required for the result to pass, but $s = 1$ always gets it done for any evidence

Application 3: political spin

Result

As long as the midpoint of the evidence is not precisely at the midpoint of the spectrum, receivers' ex post positions will be distributed according to their skepticism parameter. That is, one party will enjoy the support of receivers who are very credulous and the other party will enjoy relatively more support from receivers who are very skeptical.

- Watts et al. (1999), Lee (2005): media's 'liberal bias' is the result of conservative elites' claims
- Fessler, Pisor, and Holbrook (2017): political orientation associated with credulity
- Stephen Colbert (2006): "[i]t is a well known fact that reality has a liberal bias"
  - To the credulous receiver, an objective media or the evidence itself seems biased away from the 'fair and balanced' conclusion of the dialectic

Application 3: political spin

Result

Reinforcing evidence, in the sense that it falls within the scope of the existing evidence, has no effect on the receiver's opinion.

- Taber and Lodge (2006): 'disconfirmation bias'
- Climate change example: evidence is overwhelming in one direction
  - Explanation required to dismiss the evidence is a tortured one
  - Credulous receivers accept the conspiratorial denial
  - Campbell and Kay (2014): Republicans discount new scientific evidence on climate change more than Democrats

Application 3: political spin

Result

The marginal value to a party of releasing evidence that forces their opponent to expand their frame is always positive. In the case in which the parties use maximally extreme frames, the marginal value of these actions is higher when their opponent's existing frame is narrower in scope.

- 'Dirty tricks' / oppo research / ratf*cking
- Sender is expanding their opponents frame—seems bad if the receiver is credulous?
- Intuition is that you have shifted the mean of your opponent's frame in a bad way for them, so the DM's credulity for wide frames hurts
- Implication: piece of evidence #2 is the most valuable to dig up: (i) control the narrative, (ii) obfuscate, (iii) change the subject
- PR / news cycle theory

Summing up

- Excessively credulous receivers display a basket of well-known behavioral anomalies and responses to persuasion
- Short, superficial news cycles and ‘both sides’ discourse rationalized by credulity of receivers
- Polarization from personality parameter (complementary to motivated reasoning theory)
- Implications for conspiracy theories, ‘both-sides-ism’, public service announcements
- Dialectic model can be ported and embedded in other models
  - Effort provision / perseverance / hubris entrepreneurship
  - Financial contracts / hype / overinvestment in low quality / underinvestment in high quality
  - Life cycle consumption / insufficient saving on bad news / excessive saving on good news

Learning

The notion of learning by decision makers has cropped up a lot in our course

- Once we relax the 'standard' assumption that the DM has perfect information we have to address how their learning dynamics play out

Learning

The notion of learning by decision makers has cropped up a lot in our course

- Once we relax the 'standard' assumption that the DM has perfect information we have to address how their learning dynamics play out
- In individual choice problems, you may learn about your tastes or about what options are available
- In games, you may learn about how well your strategy performs and what you could do better in the next iterations of the interaction
- Can you brainstorm some ways that you could model learning by a DM in these or other situations? They could be Bayesian or non-Bayesian, 'standard' or behavioral

Information cascades

People ‘close’ to each other frequently display local conformity. Bikhchandani, Hirshleifer and Welch (1992) offer a model that captures one motivation for this conformity, and why it may be fragile.

Information cascades

People ‘close’ to each other frequently display local conformity. Bikhchandani, Hirshleifer and Welch (1992) offer a model that captures one motivation for this conformity, and why it may be fragile.

- Say there is some action, option or behavior.
- Each individual in (publicly known) order must decide whether to adopt or reject the behavior. Everyone gets to observe the decisions of those prior.
- Let the cost of adopting be $\frac{1}{2}$.
- The value v of adopting is the same for all, and is either zero (‘low’) or one (‘high’) with equal probability.

Information cascades

- Each individual receives a private signal about the value of adoption.
- Call individual i's signal x_i; it can be either H or L.
- The probability that the signal is correct is $p > \frac{1}{2}$: if the value of adoption is 1, each individual receives a signal H with probability p and a signal L with probability $1 - p$ (and vice versa if the value of adoption is 0).
- Say when it is an individual's turn to make a decision she believes, given her signal and the decisions of prior individuals, that the probability that the value of adoption is high is γ. The expected value of adoption is thus $\mathrm{E}[v] = \gamma * 1 + (1 - \gamma) * 0 = \gamma$.
- If the individual is indifferent between adopting and rejecting, she follows her private signal.

Information cascades

What will happen in this model?

- Individual 1 adopts if his signal is H. Why? By Bayes' rule:

$$\Pr(\text{high}|H) = \frac{\Pr(H|\text{high})\Pr(\text{high})}{\Pr(H)} \tag{12}$$

$$= \frac{p * \frac{1}{2}}{\underbrace{\frac{1}{2}p}_{\text{pr. high times pr. H if high}} + \underbrace{\frac{1}{2}(1-p)}_{\text{pr. low times pr. H if low}}} \tag{13}$$

$$= p \tag{14}$$

- Since $p > \frac{1}{2}$, the expected net value of adoption is positive. A high signal means high value is more likely to be the true state, so individual 1 prefers to adopt.

Information cascades

What about the second individual?

- The second individual can infer what signal the first saw by his action.
- Say 1 saw H and 2 also saw H. 1's action reinforces the signal 2 sees; she believes even more strongly that adoption is beneficial. The opposite is true if both saw L.
- What if 1 saw H and 2 saw L?

$$\Pr(\text{high}|HL) = \frac{\Pr(HL|\text{high})\Pr(\text{high})}{\Pr(HL)} \tag{15}$$

$$= \frac{p(1-p) * \frac{1}{2}}{p(1-p)} = \frac{1}{2} \tag{16}$$

- Then the expected value of adoption is $\frac{1}{2}$: the two signals cancel. 2 is indifferent between adopting and rejecting, and so follows her private signal.

Information cascades

What about the third individual?

1. If **both predecessors adopted**:
   - Even if 3 received a signal L, the posterior probability on H is higher than $\frac{1}{2}$; he adopts.

Information cascades

What about the third individual?

1. If **both predecessors adopted**:
   - Even if 3 received a signal L, the posterior probability on H is higher than $\frac{1}{2}$; he adopts.
2. If **both predecessors rejected**:
   - Even if 3 received a signal H, the posterior probability on L is higher than $\frac{1}{2}$; he rejects.

Information cascades

What about the third individual?

1. If **both predecessors adopted**:
   - Even if 3 received a signal L, the posterior probability on H is higher than $\frac{1}{2}$; he adopts.
2. If **both predecessors rejected**:
   - Even if 3 received a signal H, the posterior probability on L is higher than $\frac{1}{2}$; he rejects.
3. If **the predecessors chose different actions**:
   - The posterior probability for 3 is higher on whatever action 3 received a signal for. He follows his private signal.

Information cascades

- If both 1 and 2 have adopted, 3 adopts *regardless* of his signal!
- Intuitively: say 1 and 2 adopt and 3 sees L. If 3 was to only consider 1's decision (which reveals that 1 saw H) and his own signal L, he believes that the value of adoption is equally likely to be high or low. But 2 adopted; it is more likely that 2 saw H than L, and so 3 places extra weight on the value of adoption being high.
- The same is true if both 1 and 2 have not adopted.

Information cascades

- If both 1 and 2 have adopted, 3 adopts *regardless* of his signal!
- Intuitively: say 1 and 2 adopt and 3 sees L. If 3 was to only consider 1's decision (which reveals that 1 saw H) and his own signal L, he believes that the value of adoption is equally likely to be high or low. But 2 adopted; it is more likely that 2 saw H than L, and so 3 places extra weight on the value of adoption being high.
- The same is true if both 1 and 2 have not adopted.
- But if this is true then 3's choice conveys no information about 3's signal: all subsequent individuals have the same information as 3 did, and so *every* subsequent individual will follow 1 and 2!
- If 1 and 2 choose different actions, 3's problem is the same as the original problem for 1.

Information cascades

Notice that this doesn't depend on the true state. Say the value of adoption is actually low, but both 1 and 2 happened to get the signal H; this happens with probability $(1 - p)^2$.

- Then *all* individuals end up adopting a bad behavior.
- This is a **cascade**, in which individuals 'herd' on adoption even though on average more individuals see a signal L.
- A cascade happens when the informational value of my signal is outweighed by my inference on others' signals given their choices.
- These can herd the population to either the 'right' or 'wrong' behavior. In the model, the more fuzzy the private signal (p closer to $\frac{1}{2}$), the more likely is the 'wrong' cascade.

Cascade probabilities

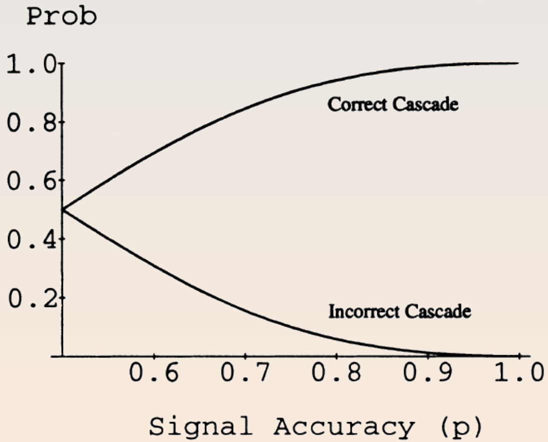


FIG. 1.—Probability of a correct and an incorrect cascade as a function of p (p is the probability that the signal is high [H] given that the true value is high [eq. (1)]). Even for large p, the probability of ending up in the wrong cascade is considerable.

Information cascades

The fact that individuals' behavior was public and their signal private is the key.

- In this case 'bad' outcomes can perpetuate entirely due to rational inference over what others' actions suggest is 'right'.
- Endogenizing the timing of individuals' decisions and enriching the signal space complicate matters. For example, what if individuals with better signals moved first? What if some individuals are stubborn and ignore what happened before them?
- In general the problem gets *worse* if individuals with more precise signals move first.

Information cascades in the lab

Anderson and Holt (1997) conduct a lab experiment designed to resemble the cascades model setup

- 72 subjects in undergrad economics courses at University of Virginia; $5 show up fee and average $20 earnings
- Session was 15 periods long, about 90 minutes
- Each period, die roll determined which of two urns would be used: A contains 2 'a' balls and 1 'b' ball; B contains 2 'b' balls and 1 'a' ball
- Urn contents put in a neutral container
- Subjects chosen in random order; in turn they get to see one private draw with replacement from the container
- Subject records the private info and writes down their assessment or what urn they think was used
- This is relayed to the announcer, who doesn't know the urn used or the private info; it's announced and all subjects record it on their record sheet

Anderson & Holt (1997)

- After all subjects have moved, true urn announced; \$2 paid for a correct guess, \$0 else

Anderson & Holt (1997)

- After all subjects have moved, true urn announced; \$2 paid for a correct guess, \$0 else

Table 2—Data for Selected Periods of Session 2

Period	Urn used	Subject number: Urn decision (private draw) 1st round	2nd round	3rd round	4th round	5th round	6th round	Cascade outcome
5	B	S12: A (*a*)	S11: B (*b*)	S9: B (*b*)	S7: B (*b*)	**S8: B (*a*)**	**S10: B (*a*)**	cascade
6	A	S12: A (*a*)	S8: A (*a*)	**S9: A (*b*)**	**S11: A (*b*)**	S10: A (*a*)	S7: A (*a*)	cascade
7	B	S8: B (*b*)	S7: A (*a*)	S10: B (*b*)	S11: B (*b*)	S12: B (*b*)	**S9: B (*a*)**	cascade
8	A	S8: A (*a*)	S9: A (*a*)	S12: B* (*b*)	S10: A (*a*)	**S11: A (*b*)**	S7: A (*a*)	cascade
9	B	S11: A (*a*)	S12: A (*a*)	**S8: A (*b*)**	**S9: A (*b*)**	**S7: A (*b*)**	**S10: A (*b*)**	reverse cascade

Notes: Boldface—Bayesian decision, inconsistent with private information.
*—Decision based on private information, inconsistent with Bayesian updating.

Anderson & Holt (1997)

- Cascades form in 41 of 56 periods in which the pattern of private signals made them possible
- Extensions to try to see if status quo bias, representativeness, or a counting heuristic are present
    - Twist 1: added two public draws after subject 4's move; makes representativeness possible for subjects 5 and 6 since they now have 3 additional draws to work with
    - Subjects followed Bayes rather than representativeness in all relevant cases
    - Twist 2: use asymmetric urns (6-1 and 5-2) that mean can't just count signals to form posterior
    - Counting explains one third of non-Bayesian decisions in this treatment
- Across all 12 sessions of all types, cascades formed in 87 of 122 possible periods

Information cascades and institutions

Hung and Plott (2001) conduct experiments to see how institutions matter in cascade formation

1. Treatment 1: individualistic institution
   - As in Anderson-Holt subjects are paid according to whether their guess is correct
   - Aside from information spillovers, there is no effect of one subject's decision on anyone else
2. Treatment 2: majority rule institution
   - Subjects paid according to whether the group decision was correct or not; each subject's guess is a 'vote' with ties broken randomly
3. Treatment 3: conformity-rewarding institution
   - Subjects paid a little according to whether their guess is correct, but more according to whether their guess matches the group decision

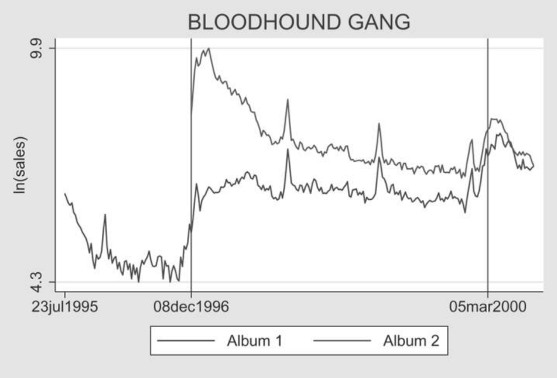
Hung & Plott (2001)

TABLE 3—EXPERIMENTAL RESULTS: OCCURRENCE OF INFORMATION CASCADES AND MEASURES OF SYSTEM PERFORMANCE

		Cascades	Efficiency	*IPQ*
Institution	$\frac{\beta}{\gamma}$	Percent occurrence	Means	
Individualistic	$\frac{\beta}{\gamma} = 0.3749$	77.5	0.817	0.85
Majority rule	$\frac{\beta}{\gamma} = 0.1344$	39	0.943	0.98
Conformity rewarding	$\frac{\beta}{\gamma} = 1.2338$	96.7	0.75	0.70

Inferring from others

Hendricks & Sorensen (2009) illustrates a different type of information spillover: product awareness. They study the effect of a new album release on sales of an older album.



Berger, Sorensen and Rasmussen (2010) demonstrate that while bad reviews of established authors hurt sales, bad reviews of unknown authors increase their sales, for similar reasons.