Persuasion via Man and Machine

A Living Review

Yaman Kumar Singla

June 2026

Table of contents

1 Persuasion via Man and Machine

A Living Review

Living Review · Draft

This is a **living review** — entries are continuously added and revised as the field evolves. The LaTeX source, bibliography, and compiled PDF are always available on GitHub. Readers are encouraged to open an issue to suggest missing work.

Also see the author's PhD thesis — *Behavior As A Modality: A Framework To Enable Automated Persuasion* — for closely related work.

1.1 Abstract

This living review synthesises the science and engineering of persuasion across five millennia of scholarship and across the disciplinary boundaries that have kept the field fragmented — rhetoric, psychology, linguistics, political science, marketing, and artificial intelligence. We argue that persuasion is best understood as the optimisation of the effect of a message on a receiver, and that machine learning is now a fully-fledged participant in that optimisation — as an instrument, as an object of study, and as a threat.

Because this is a living review, entries are added and revised continuously. Readers are encouraged to open an issue on GitHub to suggest missing work.

1.2 How to Cite

```
@misc{singla2025persuasion,
  author    = {Singla, Yaman Kumar},
  title     = {Persuasion via Man and Machine: A Living Review},
  year      = {2025},
  url       = {https://yamanksingla.github.io},
  note      = {Living review, continuously updated}
}
```

2 Introduction

Communication is defined as optimising the effect of a message on a receiver sent over a channel by a sender at a particular time [4, 5, 7]. This optimisation of altering the receiver's opinions, beliefs, or actions is also referred to as **persuasion**. The study of communication or persuasion as a universal human practice is among the most ancient of human concerns. Persuasion has been looked at as the art and science of how we get anything done. Once, it helped our ancestors to plan and execute hunting expeditions, and today, it helps us plan massive projects like the Large Hadron Collider and even inter-planetary expeditions like Europa Clipper.

Persuasion enables making ourselves understood sufficiently well so that we can coordinate our actions. The ability to persuade is not unique to our species. Persuasion is observed in both conspecific [1] and interspecific [2] scenarios. For example, both monkeys and apes yielded episodes indicating they were able to judge very finely how to obtain or hide desirable objects deceptively from the gaze of others [1]. We humans are exceptional in our capacity to cooperate with strangers. Language allowed our ancestors to cooperate, and helped to resolve conflicts by exchanging information, though this includes invented fictions, social constructions, and other imagined realities.

For example, on encountering the explorer James Cook, two Fuegian islanders stepped forward to display and then throw aside large sticks, gestures that Cook interpreted as indicating peaceful intentions. Indeed, the Fuegians and Cook were soon exchanging gifts and eating together on board ship.

4 years were children very competent in recognizing that others' mental representations of the world could be very different to their own and need to be computed to predict their future actions accurately, as seen most revealingly in the context of deception and false beliefs (Wimmer & Perner 1983).

A second, and no less important, reason involves influencing other Homo sapiens (Mercier & Sperber, 2011). Sloman and Fernbach (2017) assure us of the very real limits of our individual

knowledge and its inadequacy for accomplishing many, perhaps most, of the tasks that we face daily.

One of the oldest books on persuasion, called *Precepts*, written by Ptah-Hotep for the Pharaoh, is 4,500 years old (2375 BCE). Rhetoric — the use of symbols to persuade other humans — began in ancient Greece and its Mediterranean colonies in the fifth century BCE when there were situations of collective decision making that inspired some practitioners to ask how persuading others was best accomplished.

The study of persuasion as a discipline is in its third millennium. It was Plato's student, Aristotle, who provided the first comprehensive theory of rhetorical discourse. He defined rhetoric in terms of "observing in a given case the available means of Persuasion," instructing that the "available means" encompassed a range of appeals, some grounded in logic (*logos*), others in emotion (*pathos*), and still others in the communicator (*ethos*). Aristotle urged communicators to base judgments about the most appropriate means of Persuasion on the nature of the audience. His views had long-lasting impact: "Aristotle's theory of rhetorical discourse has withstood the test of time, furnishing axioms that guide today's practitioners of Persuasion and campaigns."

In the West, the tumultuous emergence of democracy in fourth-century BCE Athens gave rise to a class of civic intellectuals teaching persuasive speaking — known as Sophists — to the establishment of rhetoric as a field of study, and to clashes with another new field, philosophy, over explanations of the linguistic features running through belief, knowledge, argumentation, and human conduct. The early modern period was another epoch of intensified interest, with the rise of print, the Reformation, technoscience, and the colonial pursuits of Europe. The middle of the twentieth century was still another, when propaganda studies, argumentation studies, and the New Rhetoric all arose from the trauma of the Second World War.

We are in another such period now. Large Language Models have opened a new chapter in the history of persuasion — one in which the sender of a message may be a machine, the message may be tailored to a single individual at a cost approaching zero, and the channel may be every screen simultaneously. This review documents what we know, what we do not, and what the field must now answer.

> "Persuasion is ubiquitous in the political process; it is also the central aim of political interaction. It is literally the stuff of politics."
>
> — Mutz, Sniderman & Brody, 1996

When people think of persuasion they often think of debates over controversial topics, where distinct individuals and groups invested in different viewpoints confront one another. In these situations, successful persuasion amounts to convincing people to cross over from one side to the other. However, most fully formed persuasive discourse is not aimed at people who disagree anyway; in fact it is usually aimed at people who already agree, but whose agreement can be lessened or intensified.

Persuasion is also aimed at forming beliefs in the first place in people who have no stance on an issue at all. Response shaping is roughly the acquisition of an attitude, whereas response reinforcement can be equated with strengthening a preexisting attitude. By contrast, response changing references movement across the midpoint of an attitude scale.

At both a more subtle and a more powerful level, it has been argued that the Persuasion profession "serves not so much to advertise products as to promote consumption as a way of life" (Lasch, 1978, p. 72). But many of these same marketing techniques are also used to help solve pressing social problems such as improvement of the nation's health.

> **i** Case Study: "Don't Mess With Texas"
>
> In 1985, "Don't Mess With Texas" bumper stickers began appearing on cars in Texas, beginning the launch of what turned out to be the most successful anti-littering campaign ever conducted in the United States. The campaign, commissioned by the Texas Department of Transportation, targeted young male truck drivers. The advertising team recruited well-known masculine icons — members of the Dallas Cowboys, Willie Nelson, Matthew McConaughey — who looked sternly into the camera as they crushed beer cans and proclaimed the slogan.
> The campaign capitalised on Texan pride, reduced litter by 72% in six years, and is now a textbook example of **message–audience fit** in persuasion.

Charismatic leaders such as Nelson Mandela, Martin Luther King Jr., and Mahatma Gandhi have stirred the masses and brought about radical social change, even when lacking significant institutional power or money. And in 2014, the "Ice Bucket Challenge" persuaded people to donate, ultimately taking in over $115 million, funding groundbreaking ALS research.

People, however, can be remarkably resistant to persuasion. Many well-designed, well-funded efforts to encourage people to practice safe sex, stop using drugs, or improve their diet have failed [6]. People can be stubbornly resistant to changing their minds, even when their health or economic well-being is affected.

2.1 Why This Survey?

The basic difference between the sciences that study persuasion can be characterised as: the *how* of persuasion vs the *what* of persuasion. Right now, machine learning is not given a seat at the table where rhetoric and persuasion are discussed. Top academic programmes in rhetoric do not discuss the advances that computers have enabled.

Similarly, in the machine learning field, there has been comparatively little research on persuasion. Further, whatever research there has been has been in the context of optimising content for persuasion. On the other hand, persuasion should be seen as optimising across the *source*, *channel*, *time*, *message*, and even the *audience* — the full **SCEMA** framework.

What makes the study of persuasion interesting and difficult at the same time is that every field has tried to answer the same fundamental question from a different vantage point, accumulating incompatible vocabularies, incommensurable measures, and siloed literatures. This review attempts a unification.

2.2 Early Beginnings

- Language itself is believed to have developed largely because of the need for cooperation — therefore, persuasion is coextensive with sociality.
- Aristotle's *Rhetoric* (4th c. BCE) remains the founding document of the field.
- Mythology is an artefact of persuasion: narratives that mobilise groups around shared values.
- Linguistic devices — tropes, metaphors, humour — are the micro-mechanisms through which persuasion operates.

2.3 Common Misconceptions and Ethical Challenges

> **Caution**
>
> This section is under active development. Suggest content via GitHub Issues.

3 Views of Different Fields

Persuasion is one of those rare phenomena that every human discipline claims as its own. The table below maps the dominant lens, the canonical measure, and the key open question that each field brings to the common problem.

Table 3.1: Overview of disciplinary perspectives on persuasion

Field	Dominant Lens	Key Construct
Psychology	Attitude change, dual-process	ELM, dissonance
Linguistics	Phonology, syntax, framing	Powerless speech, hedging
Sociology	Social norms, networks	Diffusion, conformity
Economics	Incentives, biases	Nudge, loss aversion
Political Science	Propaganda, agenda	Framing effects
Marketing	Ad effectiveness	Memorability, recall
CS & AI	NLP, argumentation, LLMs	Claim detection, stance
Neuroscience	Neural correlates	fMRI of attitude change

3.1 Psychology

Psychology has produced the most influential theoretical frameworks in persuasion research. Three dominate the modern literature:

Cognitive Dissonance Theory [festinger1957theory?] People experience discomfort when holding inconsistent beliefs and are motivated to reduce it, making them susceptible to persuasion that resolves the inconsistency.

Matching Hypothesis Messages are more persuasive when they match the audience's regulatory focus (promotion vs. prevention orientation).

Elaboration Likelihood Model [6] The dual-process model distinguishing the *central route* (argument-quality driven) from the *peripheral route* (heuristic driven). The route taken depends on motivation and ability to process the message.

3.2 Linguistics

Hosman neatly subdivides the linguistics of persuasion in terms of *phonology*, *syntax*, *lexicon*, and *text/narrative* before turning to an examination of the effects of each on judgements of the message source, recall, and attitude change. Key findings include: powerful versus powerless speech styles, hedging, intensifiers, rhetorical questions, and the surprising persuasive force of narrative over argument in some populations.

3.3 Sociology

> **Caution**
>
> Section under active development.

3.4 Economics & Behavioural Economics

Behavioural economics treats persuasion through the lens of systematic cognitive biases — the anchoring effect, loss aversion, default effects, and social proof — that make humans predictably susceptible to certain message architectures. Thaler & Sunstein's *nudge* framework formalised this: changing the choice architecture can shift behaviour without mandate, a form of structural persuasion.

3.5 Political Science

> "Persuasion is ubiquitous in the political process; it is also the central aim of political interaction. It is literally the stuff of politics."
>
> — Mutz, Sniderman & Brody, 1996

Political science studies persuasion at the level of populations: propaganda, agenda-setting, framing effects, and the conditions under which people update political beliefs. Electoral persuasion effects are famously small but real — and AI micro-targeting threatens to amplify them at industrial scale.

3.6 Marketing & Advertising

Marketing operationalises persuasion at commercial scale. The discipline has produced rich empirical literatures on message memorability, the effectiveness of emotional vs. rational appeals, celebrity endorsement, scarcity and social proof, and the varying persuasive power of different media channels. The rise of programmatic advertising and data-driven micro-targeting has brought marketing into direct dialogue with AI.

3.7 Computer Science & AI

CS and AI engage persuasion through multiple sub-fields:

- **Formal argumentation** — Dung's abstract argumentation frameworks, computational models of argument strength.
- **Argumentation mining in NLP** — automatic identification of claims, premises, and attack/support relations in text.
- **Computer vision** — recognising persuasive intent in images and video; visual rhetoric; ad-effectiveness prediction.
- **Large Language Models** — generating and detecting persuasive text; the frontier where the field is now accelerating most rapidly.
- **Agentic systems** — AI agents that negotiate, persuade, and adapt their communication strategy to an interlocutor.

3.8 Neuroscience

> **Caution**
>
> Section under active development.

4 Types of Work in Automated Persuasion

We organise the computational literature on persuasion into three capability classes, ordered by the direction of the information flow relative to the persuasion act:

Capability	Direction	Goal
Descriptive	Observation → Explanation	Explain and measure persuasion in existing content
Simulative	Content → Predicted Response	Predict human persuasive response before it occurs
Generative	Target → Optimised Content	Create maximally persuasive content for a given audience

4.1 Descriptive Capabilities

Descriptive systems explain the mechanics of persuasion in observed content. The core task is understanding what properties of content correlate with persuasive outcomes.

4.1.1 Content Capabilities

Understanding persuasive content requires parsing three intertwined signals:

Content and Logic The quality of arguments, the presence of logical fallacies, the structure of reasoning chains. Computational argumentation mining extracts and classifies these automatically.

Advertisements Visual and multimodal persuasion; predicting ad memorability, click-through, and brand recall from image and video features.

Emotion and Affect Sentiment analysis, emotion recognition, and the mapping of affective content to persuasive outcomes across different demographics.

4.2 Simulative Capabilities

Simulative systems predict how a given person or population will respond to a given message before the message is deployed. This is hard because persuasive effects are moderated by audience characteristics that are rarely directly observed.

Representative lines of work:

- Can Language Models Recognise Convincing Arguments? — testing LLMs as judges of argument quality
- Measuring and Increasing Persuasiveness of Large Language Models [3]
- Memorability prediction from content and context

4.2.1 Single Person

Micro-targeting and personalisation — predicting individual response. Key challenge: the privacy vs. effectiveness trade-off.

4.2.2 Interpersonal Interaction

> **Caution**
>
> Section under active development.

4.2.3 Societal

- Propaganda detection
- Predicting opinion shifts in populations

4.2.4 Opinion Dynamics

Agent-based models of belief propagation; echo chambers and filter bubbles.

4.2.5 Content Recommendation

Given an audience, select the content that maximises a persuasive objective.

4.2.6 Audience Selection

Given a piece of content, identify the audience for which it will be most persuasive.

4.3 Generative Capabilities

Generative systems produce persuasive content optimised for a particular configuration of **audience**, **time**, **channel**, **sender**, and **topic**. This is the most powerful — and ethically fraught — capability class.

Large language models have dramatically lowered the cost of generating fluent, targeted persuasive text at scale. This raises questions that the research community is only beginning to confront: How persuasive are LLM-generated messages compared to human-written ones? Under what conditions do they outperform human writers? What disclosure obligations do deployers have?

> ⚠ Open Ethical Question
>
> If an AI system can generate a message that is more persuasive than any human writer could produce, targeted at a single individual, delivered via the channel and at the time of greatest receptivity — what governance structures should constrain its deployment?

5 Resources for Persuasion Research

Caution

This chapter catalogues datasets, benchmarks, and toolkits used in computational persuasion research. It is under active development. Suggest a resource via GitHub Issues.

6 Future Trends and Unsolved Questions

- **Micro-targeting at LLM scale** — as generation costs approach zero, personalised persuasion will become the norm. What individual rights are at stake?
- **Multi-modal persuasion** — integrating text, image, audio, and video into unified persuasion models.
- **Longitudinal effects** — most studies measure immediate attitude change; long-run effects of AI-generated persuasion are almost completely unstudied.
- **Cross-cultural persuasion** — what works in one cultural context often fails in another; LLMs trained predominantly on English data may encode WEIRD (Western, Educated, Industrialised, Rich, Democratic) persuasion norms.
- **Detection & provenance** — as generative AI scales, detecting AI-generated persuasive content becomes a critical counterbalancing capability.
- **Regulation** — the EU AI Act and proposed US legislation are beginning to address AI-mediated persuasion; the regulatory landscape is rapidly evolving.

This Review is Living

Entries are added and revised as the field evolves. The source and compiled PDF are always available on GitHub. The PDF updates automatically on every commit via GitHub Actions.

7 References

[1] Hare, B. et al. 2000. Chimpanzees know what conspecifics do and do not see. *Animal Behaviour*. 59, 4 (2000), 771–785.

[2] Krebs, J.R. and Dawkins, R. 1984. Animal signals: Mind-reading and manipulation. (1984).

[3] Kumar, Y. et al. 2023. Persuasion strategies in advertisements. *Proceedings of the AAAI Conference on Artificial Intelligence*. (2023).

[4] Lasswell, H.D. 1971. *Propaganda technique in world war i*. MIT press.

[5] Lasswell, H.D. 1948. The structure and function of communication in society. *The communication of ideas*. 37, 1 (1948), 136–139.

[6] Petty, R.E. and Cacioppo, J.T. 1986. The elaboration likelihood model of persuasion. *Communication and persuasion*. Springer. 1–24.

[7] Shannon, C.E. and Weaver, W. 1949. *The mathematical theory of communication*. University of Illinois Press.