

From the Masses Towards the Individual Social Engineering in Digital Communication

Abstract

This textbook examines social engineering in digital communication. Social engineering is understood here as the deliberate design and modification of collective behaviour through control over information flows and the steering of the communication environment. Digital communication is understood as communication mediated by digital technologies, particularly in those forms in which the content, mode of distribution or architecture of the communication environment serves to shape collective opinions, attitudes and behaviour. This meaning must be distinguished from the two other senses of the term: the philosophical and political one, in which social engineering denotes a programme of transforming the social and institutional order, and the one used in cybersecurity, which includes phishing, obtaining data by deception and gaining unauthorised access to systems.

The analysis encompasses three interdependent pillars of social engineering: the design of collective behaviour, control over information flows and the construction of narratives. A given practice qualifies as social engineering only if four features are jointly present: deliberate and systematic action, an orientation towards collective effects, the design not only of content but also of message distribution, and limited shared observability of the influence exercised. The principal question guiding the textbook is: how have digital tools transformed the conditions under which social engineering is applied?

The textbook's central premise is that the underlying logic of social engineering exhibits historical continuity. Mechanisms of influence known from the propaganda of the past century remain recognisable in digital communication because they rest on relatively stable properties of human cognition. What has changed is the environment in which it is practised: the scale and speed of distribution, the marginal cost of reaching an additional recipient, the granularity of segmentation, the scope for personalisation, automated testing and message optimisation, and the degree to which messages are publicly observable. As a result, influence shifts from the mass towards increasingly narrow segments and, potentially, towards the individual recipient, while retaining mass reach.

The textbook does not assume that digital messages change attitudes more effectively than press, radio or television propaganda. The distinction between the conditions of influence and its effects structures the entire argument. The mere use of

a particular technique does not constitute evidence of its effectiveness. Documenting that a technique was used and demonstrating that it actually influenced recipients' opinions, attitudes or behaviour belong to two distinct evidentiary domains. A lack of evidence confirming effectiveness is not evidence of ineffectiveness. Nor do platform-recorded responses, such as clicks, shares or viewing time, automatically constitute evidence of changes in opinion or behaviour beyond the platform.

The textbook draws on a critical, problem-oriented analysis of literature and sources, including scholarly publications, legal acts, institutional documents, reports by specialised organisations and documented case studies. The analysis is guided by demythologisation, understood as the critical examination of commonly held claims about digital manipulation, and above all about its power and effectiveness. This makes it possible to distinguish findings supported by research from sensational narratives, commonplace assumptions and claims made by actors with an interest in exaggerating the effectiveness of their own tools. Demythologisation is conducted symmetrically. The textbook rejects both technopanic, which portrays digital manipulation as virtually omnipotent, and trivialisation, according to which digital technologies have introduced no significant change. Human beings have not changed. What has changed are the conditions under which people can be reached, their reactions observed, message variants tested and the messages directed at them differentiated.

The textbook performs three principal functions. First, it organises the dispersed concepts, techniques and mechanisms of social engineering. Second, it critically distinguishes established findings from myths, technopanic and commonplace claims. Third, it provides students, teachers, journalists, creators and recipients of digital content, those engaged in public communication and the authors of political and social campaigns with tools for understanding and recognising practices designed to shape collective behaviour.

At the core of the textbook is an original model combining four processes that are most often discussed separately in the literature: the platformisation of gatekeeping, personalisation, automation and scale. They are conceptualised as overlapping aspects of a single transformation: the limitation of the shared observability of messages. Gatekeeping has not disappeared; it has been transformed and partly taken over by platforms, recommendation systems, moderators, users and advertisers. Platformisation makes the criteria governing content selection more difficult to identify. Personalisation weakens the commonality of the message because different recipients may receive different versions of it. Automation disperses agency among the campaign organiser, the system designer, the platform and the algorithmic model. Scale, in turn, makes it more difficult to reconstruct the operation as a whole, since the number of message variants, impressions and interactions may exceed the capacities of any single observer.

Observability functions as a descriptive category. It refers to the possibility of seeing a message, comparing its variants, identifying the criteria governing its distribution and attributing responsibility for its dissemination. Transparency, by contrast, is a normative requirement: the essential purpose, source and mode of influence must not be concealed from the recipient. Considered together with the truthfulness of the message, respect for the recipient's autonomy and the accountability of the sender, this requirement helps define the boundary between journalism and manipulation. A message is manipulative especially when its effect depends on the recipient's failure to recognise the sender's intention or the influence technique being used.

The material is organised into thirty problem-oriented topics divided into three parts. The argument retains the form of questions and answers, brief definitions, examples, contrasts and practical guidance. Each topic closes with a didactic block comprising an exercise, a one-sentence takeaway, self-check questions and suggestions for further reading. The first part examines the foundations and history of social engineering, demonstrating the continuity of its psychological mechanisms and the transformation of the conditions in which they are applied. The second analyses contemporary techniques and tools and critically assesses the effects attributed to them. The third addresses artificial intelligence as a factor automating known techniques of influence, together with the verification tools and strategies of resistance available to journalists and audiences.

Generative artificial intelligence is presented as a further stage in the same process. It does not create entirely new foundations of human susceptibility to influence, but increases the scale, speed and automation of message production and expands the scope for personalisation. It can generate numerous variants of a message, tailor them to particular recipients, simulate a sender and sustain interaction in real time. In these areas, quantitative change may be taken so far that it begins to verge on qualitative change.

This diagnosis points to a fundamental direction for defence. It is not enough to correct each falsehood individually or to focus on successive tools that may soon be replaced by new ones. An effective response requires restoring the conditions under which messages can be seen, compared, attributed to a sender and subjected to public scrutiny. Defence against contemporary social engineering is therefore, to a significant extent, a defence of the observability of communication. The textbook reflects the state of research, technology and regulation as of July 2026.

Keywords: social engineering, digital communication, message observability, platformisation, personalisation, microtargeting, disinformation, journalism, manipulation, transparency, artificial intelligence.