Conscious Moral Change

The Beginnings and Theories of Conscious Moral Formation
with help of the 6th version of ChatGPT Astra ultra powerful version

by Béla Pokol

Budapest
2026

Content

Comprehensive introduction……...3

Jeremy Bentham…….19

Rober Owen……30

Auguste Comte…….41

Émile Durkheim……52

Lévy Bruhl…….64

Johen Dewey…….75

B. F. Skinner……..86

Howard S. Becker……..97

John Rawls……107

Jürgen Habermas……116

Philip Kitcher…….127

Buchanan and Powell…….137

Sally Haslanger……...147

Inghmar Persson ans Julian Savulescu……159

Markus Gabriel……...171

American Moral Majority Movement…….180

Rudolf von Jhering and Nicolai Hartmann against the conscious change of moral……..191

Comprehensive Introduction

The deliberate shaping of morality in Western civilization can be traced back at least to ancient Greece. Already at that time, we find the critical examination of inherited moral beliefs, the systematic cultivation of human character, and even the organization of social institutions in ways intended to produce particular forms of moral conduct. The principal innovations of the modern era have been the expansion of these endeavours, their grounding in the social sciences, and the increasingly explicit recognition that moral rules and the institutions sustaining them can be subjected to deliberate revision.

Your parallel with legal history is particularly fruitful, but three different operations need to be distinguished:

- **Moral criticism:** demonstrating that an accepted norm is wrong, inconsistent, or harmful.
- **Moral reform:** securing social acceptance for new evaluations, practices, and expectations of conduct.
- **Institutional moral norm-making:** using designated bodies, written rules, and procedures of review to establish—and subsequently modify—the norms to be followed.

The first two are very old. The third also has early precedents, but its contemporary forms most closely resembling legislation are found in professional codes of ethics, ethics committees, and the development of international ethical standards. Meanwhile, the designation "more modern" is always part of a reform programme's justification: **the deliberate creation of a norm, its social success, and its moral rightness are three separate questions.**

One further conceptual clarification is necessary: **"positive morality,"** as used by Austin and Hart, means the morality actually accepted within a society, which may have developed largely through custom. This differs from what, in the sense of your question, we might call the *positivization of morality*: its deliberate formulation, institutionalization, and incorporation into arrangements that permit its revision.

In ancient Greece, the decisive starting point was that inherited customs became open to questioning. The fifth-century BCE Sophists' debate over *nomos* and *physis*—human custom or law and nature—drew attention to the possibility that existing prescriptions did not necessarily follow from an immutable order of things. If different communities judged the same action differently, one could ask which rule better served coexistence, and who maintained it and in whose interests.

Protagoras is particularly important in this connection. According to his view that virtue can be taught, the survival of a political community depends on capacities that must be developed through education. In Plato's *Theaetetus*, the position attributed to him holds that the wise

person can replace a community's disadvantageous beliefs with more advantageous ones. Here we already encounter the idea of intentionally transforming moral and political opinions. Nevertheless, the Sophists did not constitute a unified school, and much of what we know about Protagoras's views comes through his intellectual opponent, Plato. An overview of the Sophists' theories.

Socrates made self-examination the starting point of moral reform. Inherited opinions about courage, justice, or piety must be tested through questioning. This establishes a model in which the authority of prevailing social morality is no longer sufficient: a belief must also be capable of justification. The primary site of change is the individual's capacity for judgment, and its instruments are public conversation and the exposure of contradictions.

Plato developed this into a comprehensive institutional programme. The *Republic* subordinates education, poetry, music, physical training, political roles, and the living arrangements of the guardians to the formation of the desired character. Morality is not a separate subject of instruction here: the entire social environment must produce appropriate desires, attachments, and evaluations.

In the *Laws*, the connection with legislation is even more direct. The legislator's task is to educate people for a virtuous life; the preambles accompanying the laws serve the purpose of persuasion. Public festivals, shared practices, and educational arrangements for different ages also shape emotions. **Thus, Plato already brings together the formulation of norms, the formation of character, and institutional implementation.** In his own understanding, all this rests on knowledge of the objective good, so he does not regard change as freely chosen norm creation. The moral and political programme of Plato's *Laws*.

Aristotle identified habituation as the mechanism through which morality is shaped. According to Book II of the *Nicomachean Ethics*, moral virtue develops through repeated actions. Legislators seek to make citizens good by cultivating good habits; one of the differences between political arrangements is precisely how successfully they accomplish this.

An enduring implication is that changing norms cannot consist merely in promulgating new statements. Anyone seeking to reshape morality must also alter everyday practices, rewards, role models, and the educational environment. For Aristotle, the political community has a substantial role in moral education, but virtuous action also requires an appropriate inner disposition. Aristotle: *Nicomachean Ethics*, Book II.

The Hellenistic schools represented other approaches to the deliberate transformation of ways of life. The Cynics challenged conventions concerning wealth, rank, and social propriety; the Stoics emphasized virtue independent of external success and membership in a wider human community. In these traditions, philosophy was a practice of life, self-discipline, and the reordering of the hierarchy of values. They often sought to advance changes in social norms through exemplary conduct, communities of disciples, and personal exercises.

Epicurus explicitly opened up the possibility of revising rules of justice. Propositions 31–38 of his *Principal Doctrines* connect justice with mutually advantageous coexistence. As circumstances change, a rule may lose its just character if it no longer serves that coexistence. This is an important precedent: a norm's previous acceptance does not guarantee its continuing rightness. Epicurus did not, however, develop an institution of continuously operating moral legislation. Epicurus: *Principal Doctrines*.

In the Roman world, Augustus's moral policies illustrate deliberate change justified as restoration. His regulation of marriage, childbearing, and adultery sought to shape the conduct of the upper social strata through state action. In section 8 of the *Res Gestae*, Augustus himself connects new laws with the restoration of ancestral customs and the creation of examples to be followed.

This is a recurrent historical pattern: **the reforming authority invokes the restoration of an older morality while creating new forms of supervision and new expectations of conduct.** A self-understanding centred on preserving tradition can therefore coexist with active moral transformation. Augustus: *Res Gestae*.

Christianity made moral transformation a general programme of conversion. Drawing on its Jewish religious antecedents, it connected love of one's enemies, forgiveness, humility, responsibility towards the vulnerable, and sexual self-discipline within an order of life intended to transform people's inner motives as well.

The early Christian *Didache* organized desirable and undesirable forms of conduct into a teachable framework. Communal admonition, almsgiving, confession, ritual, and repeated instruction worked together. This went considerably beyond the formulation of moral propositions: it sought to form a different kind of person through new communal practices. Its own justification appealed to revealed truth. Historically, however, the transformation remained uneven; Christianity did not immediately abolish slavery, hierarchies, or violence. The text of the *Didache*.

Augustine placed the transformation of the inward orientation of love at the centre. The same external action can arise from different motives; moral renewal therefore concerns the will and what a person regards as the highest good. At the same time, the necessity of grace limits claims about the unlimited power of human education.

Monasticism supplied detailed institutional arrangements for the conduct of life. From the Rule of Benedict onwards, the regulation of daily routines, work, obedience, self-examination, and communal discipline served the lasting formation of moral personality. Here, within a bounded community, we encounter an explicitly organized moral environment.

Medieval ecclesiastical reforms established a durable organizational system for shaping morality. Penitential books, confession, preaching, ecclesiastical jurisdiction, and visitations made everyday conduct observable, assessable, and correctable. The Cluniac, Gregorian, and mendicant reforms sought to purify conduct and institutions in different areas.

The Fourth Lateran Council of 1215 provides a particularly clear example. Canon 21 required confession at least once a year from believers who had reached the age of discretion; Canon 6 connected the reform of morals with regular provincial councils. **Moral supervision thus acquired recurrent occasions, procedures, and responsible officials.** The truth of the norms continued to be derived from their divine origin, but their application was organized through human institutions. The canons of the Fourth Lateran Council.

Thomas Aquinas also articulated the limits of moral education and legal coercion. According to the *Summa theologiae*, human law must take account of the community's actual moral condition. It cannot prohibit every vice: it should primarily restrain actions that seriously threaten coexistence. Excessive demands can lead to greater evils.

This is an important early argument against comprehensive moral legislation. It is desirable for society to become more virtuous, but this requires gradual education; legal prohibition has different capacities from moral instruction. Thomas Aquinas: the power of human law.

In the sixteenth and seventeenth centuries, the Reformation and Catholic renewal organized religious reform into moral education on a mass scale. Luther's catechisms of 1529 set out duties associated with familial, occupational, and communal roles in an accessible form suitable for repeated instruction. Religious reform abolished certain earlier institutions and prohibitions, while in other areas seeking to establish more intensive discipline in everyday life. Luther: *Small Catechism*.

In Calvin's Geneva, moral supervision assumed a particularly definite institutional form. The Ecclesiastical Ordinances of 1541 distinguished the functions of pastors, teachers, elders, and deacons. The consistory regularly examined disorders within the community and intervened through admonition and discipline. On the Catholic side, the Tridentine reform and the development of clerical training, catechisms, and episcopal supervision likewise increased the organization of deliberate moral formation. The outcomes differed between confessions, but both sides strengthened the institutional shaping of conduct. Calvin's Ecclesiastical Ordinances.

From the seventeenth century onwards, the shaping of morality increasingly proceeded from theories of human nature and social institutions. Thomas Hobbes's *Leviathan*—1651—makes the teaching of political obligations and the management of doctrines that undermine the community conditions of the state's survival. Peaceful coexistence requires a shared framework of interpretation, appropriate education, and incentives that guide conduct.

Hobbes does not, however, regard every moral standard as an arbitrary creation of the sovereign: he considers the laws of nature immutable. The principal domain of deliberate shaping is the determination, teaching, and maintenance of the order to be followed collectively. Hobbes: *Leviathan*, especially Chapters XV, XVIII, and XXX.

Spinoza developed both the idea of transforming beliefs and that of limiting coercion over conscience. The *Theological-Political Treatise*—1670—and the *Ethics*—1677—examine the natural connections among fears, passions, and religious ideas. Moral reform can therefore proceed through understanding people's actual motivations and creating the conditions for a more rational life.

At the same time, he considers complete state control of thought impossible and harmful. A coerced profession of belief can produce hypocrisy and resistance. For him, the shaping of morality is therefore closely connected with the question of how freedom of thought can be preserved. Spinoza's political philosophy.

The thesis that morality can be socially produced emerged with particular force in the French Enlightenment. In *De l'esprit*—1758—Helvétius emphasized the formative role of interests, education, and institutions. If human conduct is largely a consequence of its environment, improving prevailing social morality requires transforming that environment. The legislator should arrange incentives so that the pursuit of private interest is connected with socially beneficial conduct.

This is already an explicitly design-oriented perspective: the desired social outcome must be specified, followed by the educational and institutional conditions that facilitate it. A substantial proportion of traditional virtues may remain; the major change occurs in their explanation and in the methods of cultivating them. The moral programmes of the Enlightenment.

D'Holbach's "ethocracy" was an explicit programme of government founded on morality. *Éthocratie* and *La morale universelle*—both published in 1776—derived moral rules from people's earthly happiness, needs, and interdependence. The state must provide security, well-being, and an education through which cooperation becomes a recognized interest.

Social explanation and institutional incentives take the place of religious rewards and punishments. D'Holbach's atheism did not, however, entail the rejection of every traditional virtue: he sought to preserve temperance and mutual assistance by placing them on new foundations. D'Holbach's moral and political theory.

For Rousseau, the moral formation of the citizen became a fundamental condition of political order. *The Social Contract* and *Emile*—1762—address two connected questions: how an autonomous person can be educated, and how a citizen willing to assume responsibility for the community can be formed.

The lawgiver's task extends beyond drafting laws. The lawgiver must create attachments and a shared identity through which people experience the common interest as their own concern. Civil religion, communal practices, and education serve this purpose. Rousseau nevertheless recognizes the circularity: good institutions require an appropriate civic character, while those very institutions are supposed to create it.

His programme brings into particularly sharp focus a problem that has recurred ever since: **on what authority may the leaders of a community reshape people's desires and moral identifications?** Rousseau's political and educational philosophy.

The French Revolution sought to turn moral regeneration into a practice of the state. Republican festivals, the new calendar, political symbols, and educational plans were also intended to reorganize people's everyday attachments. Civic equality, patriotic duty, and republican virtue were to take the place of monarchical dependence and courtly corruption.

Robespierre's speech of 5 February 1794 treated virtue as the foundation of revolutionary government and also used it to justify terror. The Festival of the Supreme Being in June 1794 sought to create a shared moral and religious framework. Within the Revolution, however, different moral and religious programmes also struggled against one another. This is one of the earliest modern examples of how the aspiration to moral renewal can become connected with the moral exclusion and coercion of political opponents. Robespierre's speech, the Festival of the Supreme Being.

Kant limited deliberate moral reform through the requirement of autonomy. According to the *Groundwork of the Metaphysics of Morals*—1785—habit or authority alone does not provide sufficient moral justification. Principles of action must be subjected to reason's requirement of universalizability.

Kantian self-legislation does not, however, mean that everyone chooses a morality according to personal preference. An autonomous person obeys a principle that they can rationally regard as universally valid. The task of education is to develop this capacity for judgment.

For Kant, the distinction between outward conformity to law and an inward moral motive is therefore decisive. The state may require appropriate external conduct, but it cannot produce a virtuous motive by mere command. **This is one of the strongest philosophical limits on the complete transformation of morality through something resembling legislation.** Kant's social and political philosophy.

Bentham brought the deliberate reform of law and morality into a common evaluative framework. *An Introduction to the Principles of Morals and Legislation*—1789—seeks to assess moral and legal rules according to utility. The rightness of an established prohibition must be examined in terms of its consequences for human beings.

Particularly important is his distinction between physical, political, moral or public-opinion, and religious sanctions. Social conduct is shaped by several kinds of reward and disapproval; the reformer must understand how they operate together. Legislators sometimes regulate directly and at other times can give direction to the judgments of public opinion.

Bentham also examines in detail when punishment is unjustified: when it is ineffective, excessively costly, or unnecessary. Deliberate moral reform therefore requires a differentiated range of instruments. **Bentham is central to your parallel because he treats legislation, education, and moral incentives as distinct components of a single task of shaping society that can be subjected to deliberate examination.** Bentham: the limits of ethics and legislation.

John Stuart Mill connected this programme with the protection of individual liberty. According to *Utilitarianism*—1861, published as a book in 1863—education and public opinion can connect a person's sense of their own happiness with the good of others. Conscience has emotional and socializing conditions; the social force of a new moral principle must be developed through these.

On Liberty—1859—warns, however, that social disapproval can become tyrannical even without state coercion. Mill supports the development of character and the prevention of harm while protecting individual experiments in living against the homogenizing pressure of prevailing social morality. Mill: *Utilitarianism*.

The emancipatory movements of the eighteenth and nineteenth centuries demonstrated the social methods of targeted moral norm change. Movements against slavery reclassified an institution that had previously been widely accepted as a fundamental moral wrong. Associations, sermons, petitions, personal testimonies, and consumer boycotts connected distant suffering with everyday responsibility.

The programme of the American Anti-Slavery Society, established in 1833, explicitly sought to change public opinion through a network of local organizations, speakers, and publications. Those directly affected, including formerly enslaved people, also played a decisive role in this process. Changing norms thus emerged as something undertaken not exclusively by philosophers or states. Documents on the American abolitionist movement.

Mary Wollstonecraft called for the deliberate transformation of the ideal of female character. *A Vindication of the Rights of Woman*—1792—criticized practices that educated women for dependence, the desire to please, and subordination. She advocated education founded on women's intellectual capacities, moral responsibility, and independence. The objects of reform were simultaneously education, family relationships, and society's understanding of desirable female conduct. Wollstonecraft's work.

The Seneca Falls Declaration of 1848 subsequently itemized the legal and social practices that maintained women's subordination. Documents of this kind represented the transition between criticism of norms and organized reform: they identified the order to be changed and placed a publicly articulated alternative alongside it. The Seneca Falls Declaration.

Robert Owen expected the formation of moral human beings to result from reorganizing the social environment. The central idea of *A New View of Society*—1813–1814—is that people's character is largely shaped by circumstances whose creation they did not choose. Moral condemnation is therefore insufficient: education, working conditions, housing, and the organization of cooperation must be changed.

The practices of New Lanark and Owen's subsequent communal experiments aimed to create a "new moral world." Moral reform thereby became experimental social organization: the reformer creates living conditions expected to generate more cooperative conduct. Owen: *A New View of Society*.

Saint-Simon and Auguste Comte sought to create new spiritual leadership for industrial society. Saint-Simon's *New Christianity*—1825—reformulated fraternity as the organization of society towards improving the condition of the poorest.

Comte developed a more detailed institutional plan. *Système de politique positive*—1851–1854—and *Catéchisme positiviste*—1852—sought to sustain altruism and a sense of duty through the "religion of humanity," education, ceremonies, and a new spiritual authority. The moral order was to have its own designated custodians.

Comte separated spiritual power from temporal government, but sought a broad, shared moral consensus. **His is one of the nineteenth century's most explicit plans for an institutionally produced and maintained social morality.** Comte's philosophical and social programme.

Marx linked moral change to the transformation of social relations. His criticism of bourgeois morality is connected with the critique of property, production, and class relations. Transforming relations between people can create the conditions for a different kind of moral conduct.

The third of the *Theses on Feuerbach*, however, also criticizes a one-sided educational construction: the educator must also be educated. The reforming group does not stand outside society. The transformation of circumstances and people's self-transformation are connected in shared practice. Marx therefore cannot simply be described as a thinker who wished to impose on society a moral code drawn up by an intellectual elite. Marx: *Theses on Feuerbach*.

Nietzsche made the value of moral standards itself an object of investigation. *Daybreak*—1881—*Beyond Good and Evil*—1886—and *On the Genealogy of Morality*—1887—ask which historical forces, human types, and power relations produced prevailing evaluations.

The “revaluation of values” is more radical than the more consistent application of existing norms: it seeks to change the very foundations of evaluation. The value of self-denial, obedience, or equality is not self-evident to him either.

Nietzsche did not, however, develop a state procedure for promulgating a uniform new mass morality. Creative self-formation and the possibility of exceptional human achievements play a more important role. His historical significance lies in making moral standards themselves visible as products of human history and as evaluations capable of transformation. Nietzsche’s moral philosophy.

Jean-Marie Guyau sought a new foundation for moral motivation. *Esquisse d’une morale sans obligation ni sanction*—1885—sought to understand morality through the unfolding of life, its creative capacities, and its social extension, in place of absolute obligation and punitive sanctions. He investigated the internal sources of social action and thereby formulated an alternative to a purely commanding moral order. The critical edition of Guyau’s work.

Lévy-Bruhl is one of the most direct intellectual-historical antecedents to your question. *La morale et la science des mœurs*—1903—distinguished three levels: actually existing moralities; their empirical science; and a rational moral practice drawing on that science, the *art moral rationnel*.

In his view, traditional moral philosophies often retrospectively justify norms that already exist. A thorough knowledge of social morality, by contrast, can reveal which factors sustain a norm and how it might be changed.

Here, deliberate moral reform appears as the practical application of scientific knowledge. A fundamental problem nevertheless remains: science can reveal the conditions of change and its probable consequences, while selecting the goal to be pursued also requires a value judgment. Lévy-Bruhl’s preface and responses to criticism.

Durkheim developed an institutional programme of secular moral education. *Moral Education*, published in 1925 from lectures delivered in 1902–1903, argues that after religious justification weakens, the social authority of morality must be grounded and made experientially meaningful in a new way.

Discipline, attachment to social groups, and autonomous understanding together shape moral personality. The school is the central institution of this programme. Durkheim supported the development of a morality appropriate to industrial society and recognizing the dignity of the human person, but sought to proceed from society’s actual structure. A philosophical conception alone is insufficient to create a functioning morality. Durkheim’s theory of morality.

John Dewey made the continuous, experience-based revision of morality into a programme. According to *Ethics*—with Tufts, 1908; revised in 1932—and *Human Nature and Conduct*—1922—inherited habits are indispensable foundations of coexistence, but may become obstacles in new situations.

Reflective morality proceeds from the problems that arise. It investigates consequences, listens to those affected, and, where necessary, modifies the goals themselves. On this approach, rules are human solutions capable of improvement.

Dewey comes particularly close to the idea of deliberately revisable morality: **moral reflection is an ongoing process of learning** that requires public discussion, democratic participation, and appropriate education. He does not, however, designate an exclusive moral legislator. Dewey's moral philosophy, *Human Nature and Conduct*.

Gramsci made "intellectual and moral reform" part of the organization of social power. In the *Prison Notebooks*, written between 1929 and 1935, the durability of political leadership also depends on shaping everyday common sense, shared interpretations, and moral identifications.

The task of the party, intellectuals, the press, education, and cultural organizations is to form a new collective will. Reform is connected with economic and social transformation and must also take root in everyday thought. This became an influential theory of the relationship between morality, culture, and political leadership. It does not follow, however, that every subsequent cultural change was an implementation of Gramsci's plan. Gramsci's theory.

Eugenics around the turn of the twentieth century illustrates a discriminatory form of deliberate moral transformation. According to Francis Galton's programme of 1904, eugenics must first secure scientific acceptance, then become a practical programme, and finally become part of the national conscience. He sought to realign social judgments concerning marriage and childbearing with new considerations of heredity.

This is a particularly clear example of a movement seeking to create new moral obligations by appealing to scientific modernity. Subsequent discriminatory and coercive practices also demonstrated how serious the consequences of connecting scientific authority with moral classification could be. Eugenics spread in several Western countries; Nazi racial policy became associated with a particularly radical and murderous version of it. Galton's 1904 text, the history of eugenics.

Communist and fascist systems made the re-education of the whole person a task of the state and party organization. Lenin's 1920 speech on the tasks of youth organizations connected communist morality with the aims of the proletarian class struggle. Education, work, discipline, and political participation were jointly to form the new human being. The Soviet experience belongs here as a distinctive Eastern institutionalization of ideas originating in the West. Lenin's speech.

Italian fascism emphasized the ethical character of the state, sacrifice, and the subordination of individual life to the political community. Nazi education connected moral duty with racial exclusion, the leadership principle, and readiness for war. Party-states combined the instruments of instruction, role models, social rewards, and coercion. Despite differences in ideological content, a common structural feature was political power's claim to be the authoritative interpreter of moral truth as well. The doctrine of fascism, the educational role of the Hitler Youth.

Psychological approaches raised the question of changing morality at a new level. Freud examined the psychological consequences of inherited sexual morality: certain demands can cause suffering and illness. Their effects on the human personality thus also become a criterion of reform. At the same time, Freud acknowledged the restraints on instinct necessary for coexistence. Freud's 1908 essay.

B. F. Skinner advocated the deliberate design of the behavioural environment. The experimental community of *Walden Two*—1948—and subsequently *Beyond Freedom and Dignity*—1971—sought to produce more cooperative behaviour by shaping reinforcements and social conditions. Here the immediate object of reform is the environment that generates behaviour.

This brings the shaping of morality close to behavioural technology. The decisive question becomes who selects the desired behaviour and how the designer is held accountable. Producing appropriate behaviour does not automatically demonstrate the development of an independent capacity for moral judgment. Skinner: *Beyond Freedom and Dignity*.

In the second half of the twentieth century, a substantial part of moral reform sought to change established exclusions and relations of subordination. The American civil rights movement challenged both discriminatory law and social customs. Martin Luther King's 1963 letter from Birmingham jail justified organized nonviolent action aimed at awakening the community's conscience.

Here the mechanism of change connects making injustice visible, exposing moral inconsistency, public conflict, and negotiation. Moral reform was nourished both by religious conviction and by the demand for constitutional equality. King: *Letter from Birmingham Jail*.

Feminist consciousness-raising movements transformed the moral interpretation of private life. The collective examination of personal experiences revealed that situations experienced as individual failures could reflect social role expectations and patterns of power. Established evaluations of housework, sexuality, family authority, and self-sacrifice thereby became subjects of public debate.

Carol Hanisch's *The Personal Is Political*, written in 1969 and published in 1970, is a characteristic document of this method. Here moral transformation begins with the comparison and collective reinterpretation of the experiences of those affected. Hanisch's original essay.

Carol Gilligan also reconsidered the standard by which the capacity for moral judgment is assessed. *In a Different Voice*—1982—emphasized the importance of care, relational responsibility, and concrete vulnerability alongside abstract considerations of justice. She thereby questioned whether people's moral development could be ranked exclusively through a single moral language. Later approaches to the ethics of care also encouraged the reassessment of familial and institutional responsibility. Gilligan's work.

Herbert Marcuse regarded the transformation of human needs and sensibilities as part of liberation. According to *Eros and Civilization*—1955—and *An Essay on Liberation*—1969—an oppressive social order becomes embedded in desires and everyday evaluations. Lasting change therefore requires a new sensibility that relates differently to aggression, competition, work, and pleasure.

Counterculture, art, and social movements acquire an active role in shaping morality. This programme goes well beyond extending particular rights: it seeks to re-evaluate the desirable life as a whole. Marcuse: *An Essay on Liberation*.

Foucault approached the transformation of our relationship to morality from another direction. Investigating the historical emergence of norms can weaken the belief that the

present order is natural and inevitable. His later works emphasized deliberate work on oneself and the shaping of one's way of life.

Foucault is therefore important in the history of moral self-formation, but he did not propose the creation of a new universal code. Rather, he examined how we might become moral subjects differently and modify our relationship to the norms that shape us. Foucault's ethics.

Rawls placed the method of revising norms in the foreground. The reflective equilibrium of *A Theory of Justice*—1971—means the mutual correction of considered judgments about particular cases and general principles. Inherited moral judgments provide a starting point but are open to revision; at the same time, the theory itself may require modification if it leads to unacceptable consequences.

This is a deliberate method of moral learning. The later *Political Liberalism*—1993—recognizes that several comprehensive religious and moral doctrines may coexist in a free society. Rawls therefore does not seek to replace the entirety of prevailing social morality with a single philosophical system. The theory of reflective equilibrium.

Apel and Habermas developed the procedural conditions for justifying norms. According to discourse ethics, an inherited norm remains open to revision even if it has long been accepted. We must examine whether those affected could accept its consequences in sufficiently free and informed discussion.

In Habermas's *Moralbewußtsein und kommunikatives Handeln*—1983—post-conventional moral consciousness signifies this reflective relationship. A norm's validity is not guaranteed simply by tradition or authority.

Here we approach the analogy with legislation through procedure: who may participate, which arguments count, and how can distortions caused by power be avoided? **An actual majority decision, however, is not in itself identical with moral rightness.** Habermas's work.

Environmental and animal ethics sought deliberately to extend the scope of moral concern. Aldo Leopold's land ethic of 1949 extended the idea of membership in a human community to soils, waters, plants, and animals. Nature thereby becomes a community possessing moral significance, to which human beings also belong. Leopold's land ethic.

Peter Singer's *Animal Liberation*—1975—and *The Expanding Circle*—1981—called for the consideration of animal suffering and the interests of those living outside our immediate community. His later works also subjected the conventional morality of charitable giving to more demanding scrutiny: how much do we help, and with what results?

The instruments of change are argument, making suffering visible, modifying consumption habits, and organized assistance. Singer's utilitarian programme is an influential but contested strand of animal advocacy and altruism. Singer's works and their publication dates.

Conservative and religious movements have likewise engaged in deliberate moral formation. The establishment of the American Moral Majority in 1979 aimed to strengthen family and religious norms in politics, education, and the media. Here the programme of preserving tradition was combined with modern instruments of organization, communication, and mobilization. Such endeavours demonstrate that deliberate moral transformation has no

exclusively progressive political orientation. A collection of primary documents from the Moral Majority.

Explicit revision of norms can also occur within religious teaching. **The 2018 amendment of paragraph 2267 of the Catholic Catechism, concerning the death penalty,** constituted an institutional modification of a specific text. The new wording declares the death penalty inadmissible and supports its abolition.

From a formal perspective, this closely resembles the amendment of a norm: there is a competent authority, an identifiable text, and an official new version. Its own theological interpretation, however, speaks of the development of doctrine, justifying the change through a deeper understanding of fundamental principles. The amended text, the official explanation of the amendment.

Contemporary social science also examines the agents of norm change and the processes through which norms spread as distinct subjects of inquiry. In *Outsiders*—1963—Howard S. Becker called those who initiate new rules, or are concerned with enforcing them, "moral entrepreneurs." The concrete organizers, experts, movements, and enforcers behind the moral order thereby become visible. Excerpts from Becker's work.

Martha Finnemore and Kathryn Sikkink's 1998 study distinguished the emergence, accelerating diffusion, and internalization of international norms. The initiators of a new norm compete with existing norms, seek organizational support, and attempt to change what states or communities regard as acceptable conduct. This is a descriptive theory of the process; successful diffusion does not automatically establish moral justification. Finnemore–Sikkink: *International Norm Dynamics and Political Change*.

Philip Kitcher offers one of the most direct contemporary theories of deliberate moral improvement. *The Ethical Project*—2011—and *Moral Progress*—2021—interpret morality as an improvable human enterprise developed to address the difficulties of living together.

The criterion of improvement is overcoming concrete problems, exclusions, and suffering. Kitcher asks how this process might be made more systematic, more reliable, and less conflict-ridden. He assigns a central role to collective inquiry that is adequately informed and receptive to the perspectives of those affected.

He regards as particularly difficult cases in which subordinated people have already accepted their disadvantaged position and consequently do not even formulate complaints. Reform must also examine such internalized evaluations. At this point, however, the question arises: who is entitled to determine that others' acceptance is distorted, and how can paternalism be avoided? **Kitcher's programme makes the deliberate improvement of morality a task inseparable from criticism of the procedure of improvement itself.** Kitcher: *Moral Progress*.

Buchanan and Powell emphasize maintaining the social conditions of moral progress. According to *The Evolution of Moral Progress*—2018—the biological origins of human moral capacities do not permanently confine moral concern within narrow group boundaries. The cultural and institutional environment substantially shapes how inclusively we judge.

Progress can therefore be reversed, and maintaining it is a distinct task. It is also noteworthy that improvement sometimes consists in reducing moralization: removing a way-of-life choice

previously condemned as sinful from the sphere of communal stigmatization. Deliberate moral reform can thus mean not only creating new obligations but also lifting earlier moral prohibitions. Buchanan and Powell's work.

Kwame Anthony Appiah identifies one of the keys to moral revolutions in the reorganization of recognition and shame. According to *The Honor Code*—2010—counterarguments have often long been known while the objectionable practice persists. Change can accelerate when the system of honour and social esteem also changes.

Through examples such as British duelling and slavery, he therefore examines how previously accepted conduct can become shameful. Alongside argument, reformers also shape the community's self-image. Appiah's discussion of moral revolutions.

Markus Gabriel connects deliberate moral progress with moral realism. *Moralischer Fortschritt in dunklen Zeiten*—2020—calls for social renewal founded on universal values. This is an important difference from programmes emphasizing the human creation of morality: according to Gabriel, the object of correct moral understanding is not merely the outcome of social decisions. Institutions and shared convictions must be modified on the basis of a better recognition of moral truths. Gabriel's work.

Cristina Bicchieri develops empirically investigable methods of norm change. *Norms in the Wild*—2017—distinguishes people's own opinions from what they believe about others' behaviour and expectations.

A practice may persist even when many people privately reject it, because they believe that others support it and expect compliance. In such cases, the key to change lies in exposing mistaken mutual assumptions, identifying the relevant community, establishing collective commitments, and making new behaviour visible.

This can also be used in the practical planning of moral reform, although Bicchieri's theory concerns the broader category of social norms. Here too, knowledge of how to change a norm differs from determining which change is right. Bicchieri: *Norms in the Wild*.

Sally Haslanger emphasizes the deliberate transformation of concepts and institutional practices. "Ameliorative" conceptual analysis asks which concepts would better assist our understanding of social relations and our response to injustices. Moral judgment also depends on whether we describe a phenomenon as a natural difference, a personal failing, or an institutional disadvantage.

According to Haslanger, argument alone is often insufficient: institutional habits, expectations, and opportunities must also be changed. On this approach, moral reform is connected with reforming shared interpretive resources and everyday practices. Haslanger: *Changing the Ideology and Culture of Philosophy: Not by Reason (Alone)*.

Behavioural science and biotechnological moral enhancement have further expanded the scope of intervention. Richard Thaler and Cass Sunstein's *Nudge*—2008—seeks to encourage more desirable decisions by shaping choice environments, defaults, and feedback. This is a borderline case of moral formation, since behaviour may change while a person's moral convictions remain unchanged. Thaler–Sunstein: *Nudge*.

In *Unfit for the Future*—2012—Ingmar Persson and Julian Savulescu argue that our technological capacity to cause harm has grown faster than our moral self-restraint. Alongside traditional education, the possibility of biologically influencing moral motivation must therefore also be investigated.

This is a contested research programme, not an available, generally effective "morality-enhancing" technology. Moreover, increasing empathy or the propensity to cooperate does not in itself determine how we should choose between conflicting values. Persson and Savulescu's own account.

The contemporary form closest to legislation is the institutional creation and revision of professional ethical norms. Since the American Medical Association's ethical code of 1847, the development through which moral expectations become written rules, professional interpretations, and regularly revised guidance has been particularly well documented. A designated body maintains the code and develops its provisions in response to new problems. The AMA Code of Medical Ethics and its institutions.

The **Declaration of Helsinki** is an even more vivid example. The World Medical Association adopted it in 1964 and subsequently amended it on several occasions; its most recent comprehensive revision took place in 2024. It is associated with a designated body, documented versions of the text, research ethics committees, and requirements of oversight. This does indeed resemble procedurally regulated norm-making. Nevertheless, an ethical declaration, an institutional obligation, and a rule of state law remain different sources of norms. The Declaration of Helsinki and the history of its amendments.

International ethical standard-setting extends the same degree of organization into new fields. UNESCO's 2005 declaration on bioethics connects principles, ethics committees, and educational responsibilities. Its 2021 Recommendation on the Ethics of Artificial Intelligence, followed by the Recommendation on the Ethics of Neurotechnology adopted in November 2025, formulates jointly developed norms for new technologies.

In these documents, expert preparation, consultation, intergovernmental adoption, and recommendations for implementation constitute a clearly organized process of norm formation. The adoption of a recommendation does not, however, in itself signify either social moral agreement or the directly legally binding character of every provision. UNESCO: bioethics, 2005, artificial intelligence, 2021, neurotechnology, 2025.

The historical material can be organized as follows according to the mode of change. This is my own analytical classification; the individual authors did not necessarily use these designations.

Mode of shaping morality	Representative thinkers and examples	Principal instrument of change
Restoring an order regarded as right	Augustus, religious reforms	Authority, teaching, discipline, example
Deliberate education of character	Plato, Aristotle, Durkheim	Schooling, habituation, communal practice
Rational revision of norms	Kant, Bentham, Mill	Universalizability, consequences, critical argument

Mode of shaping morality	Representative thinkers and examples	Principal instrument of change
Transforming the conditions of life	Owen, Marx, Dewey	Work, institutions, relations of cooperation
Organizing a new shared moral order	Comte, Gramsci, revolutionary programmes	Spiritual or intellectual leadership, culture, political organization
Reform based on the experiences of those affected	Abolitionism, civil rights and feminist movements	Testimony, public discussion, organized action
Systematically developing the methods of norm change	Habermas, Kitcher, Bicchieri	Discussion, inquiry, modification of social expectations
Institutional amendment of written ethical norms	Professional codes, the Declaration of Helsinki	Competent body, procedure, revised text

Hart formulated the limit of the analogy with legislation with particular precision. In Chapter VIII of *The Concept of Law*, he treats morality as resistant to direct change through an act of enactment. This does not mean that morality cannot be deliberately influenced. The difference is that a valid legislative act directly changes the law, whereas its influence on morality operates through social processes and has uncertain results.

A parliament may repeal a statute, but it cannot eliminate a widely internalized moral prohibition merely through an equivalent act. Law, education, and public discussion can change evaluations over time. Leslie Green's 2013 study, *Should Law Improve Morality?*, specifically examines the possibilities and limits of such influence. Green's study.

An independent critical tradition concerning planned moral reform has also developed. Popper supported limited, testable, and corrigible reforms against programmes seeking to remake society as a whole according to a definitive plan. Oakeshott's critique of rationalism drew attention to the practical knowledge preserved in customs that is difficult to articulate explicitly. From these perspectives, before abolishing an old norm, we must understand what function it performs in social coexistence and what might take its place. Popper's social philosophy, Oakeshott's critique of rationalism.

Luhmann identified a further problem in the moralization of social conflicts. If a professional or political disagreement is transformed into a question of moral esteem or disesteem for persons, a difference of opinion can easily become a rejection of the person as a whole. The task of ethical reflection therefore also includes examining the consequences of moral communication. This is particularly important for understanding contemporary campaigns to change norms: promoting a new standard may be accompanied by the moral devaluation of those who follow the old one. Luhmann: *Paradigm Lost*, *The Sociology of the Moral and Ethics*.

I see the most fruitful further development of your proposition in the observation that **in the history of morality, the shaping of primary norms has gradually been supplemented by the development of procedures for changing those norms**. The question is no longer only what we regard as right, but also who may initiate a new norm, who must be heard, what experience justifies revision, and how a mistaken reform can be corrected.

This development has advanced furthest in professional and organizational ethics. In the morality of society as a whole, several competing centres of norm formation continue to

operate: families, religious communities, movements, schools, professional organizations, states, and institutions of communication. **The deliberate shaping of morality is a longstanding historical reality; a generally accepted moral "rule of change," however, remains an unresolved question.** Here lies both the most important promise of the parallel with the positivization of law and the deepest difference between the two.

Jeremy Bentham

Jeremy Bentham developed a systematic programme for deliberately reforming moral judgment, moral behaviour, and the institutions that sustain them. His central idea was that inherited rules should be examined according to their consequences for happiness and suffering, and that the circumstances shaping conduct could then be consciously reorganized. Legislation was one instrument in this programme; education, public discussion, reputation, and the design of institutions were others.

"Conscious moral change" is a useful description of this project, although it is not the title of a single Benthamian doctrine. The theory must be reconstructed from several bodies of writing. These address three distinguishable questions: **Which moral judgments deserve revision? How can people be motivated to behave differently? Who should possess the power to direct that transformation?** Bentham's answers form a coherent reforming enterprise, but they also contain tensions—especially between individual freedom and organized supervision, and between educating public opinion and manipulating it.

His intellectual development provides the necessary context:

Period and principal writings	Contribution to deliberate moral reform
A Fragment on Government (1776)	Separates explaining existing institutions from judging whether they ought to exist.
An Introduction to the Principles of Morals and Legislation, printed in 1780 and published in 1789	Establishes the utility principle, analyses motivation and sanctions, and distinguishes private ethics from legislation.
Writings on indirect legislation and institutional reform, beginning in the 1780s	Investigate how changing circumstances, opportunities, information, and incentives can prevent harmful conduct.
Deontology and *A Table of the Springs of Action*, largely composed in 1813–1815	Examine practical moral guidance, motives, and the evaluative language through which conduct receives approval or condemnation.
Chrestomathia (1816–1817)	Develops an educational project connecting useful knowledge with the improvement of individual capacities.
Later constitutional writings and *The Book of Fallacies*	Connect reform with public scrutiny, representative institutions, and criticism of arguments protecting established interests.

The publication history requires care. Bentham left extensive manuscripts, and the posthumous *Deontology* edited by John Bowring in 1834 is not a wholly reliable presentation of his own completed position. For precise interpretation, Amnon Goldworth's manuscript-based 1983 edition is preferable. The UCL Bentham Project's description of that edition explains the relationship among these ethical writings.

Bentham's first decisive move was to make inherited authority answerable to an independent standard of criticism. In *A Fragment on Government*, his attack on William Blackstone turns partly on the distinction between the *expositor*, who explains the law as it exists, and the *censor*, who evaluates what it ought to be. "Censor" here means a critical evaluator, rather than an official suppressing expression. The existence, antiquity, or prestige of an institution cannot establish its justification. Otherwise, describing the established order becomes a disguised defence of it. Bentham, *A Fragment on Government*, Preface.

Applied to morality, the implication is substantial. A society's condemnation of some behaviour is a fact to investigate. Whether that condemnation is justified is a further question. Historical persistence explains neither its correctness nor its entitlement to continue.

This makes possible a particular form of moral change: people can retain a commitment to reasoned moral evaluation while rejecting particular judgments previously regarded as unquestionable. Reform need not present itself as moral decline or the abandonment of morality. It can claim that established morality has failed its own proper purpose.

Bentham's programme emerged from the Enlightenment, with Helvétius especially important. Helvétius had connected conduct, interests, education, and political arrangements. Bentham developed this approach into more detailed proposals concerning the mechanisms through which institutions could bring personal interests closer to general welfare. Matthias Hoesch emphasizes both this intellectual debt and Bentham's greater attention to feasibility and to the active role of philosophers and citizens in reform. Hoesch, "From Theory to Practice: Bentham's Reception of Helvétius".

The difference matters for your question. A theory explaining why moral beliefs vary with social circumstances becomes a programme of conscious transformation once it asks which circumstances should be changed, by whom, and through what means.

The utility principle supplies the criterion by which Bentham judges both existing morality and proposed replacements. Conduct and governmental measures should be evaluated by their tendency to increase happiness or reduce suffering among those affected. The community has no separate happiness existing independently of its members: its interest must be understood through the interests of actual individuals. Utility, in this sense, means well-being; it cannot be reduced to administrative efficiency, profitability, or whatever benefits the state. Bentham, *Introduction*, chapter I.

This establishes an important limit on conscious moral reconstruction. A ruler, philosopher, or majority cannot make a norm morally justified merely by choosing or promulgating it. Their decision remains subject to the same evaluative standard as the custom they propose to replace.

Bentham therefore combines a relatively stable ultimate criterion with potentially extensive revision of particular rules. Moral change may become necessary because people discover previously neglected suffering, abandon a mistaken causal belief, recognize interests formerly

excluded, or encounter changed circumstances. The underlying commitment to happiness need not change whenever an application changes.

An analytical distinction follows: **Bentham advocates the deliberate reconstruction of moral rules, while denying that moral correctness is created simply by an authoritative act of will.** This will become crucial when considering the analogy between moral reform and legislation.

His criticism of established moral reasoning is particularly explicit in his treatment of *asceticism* and *sympathy and antipathy*. Asceticism, as he characterizes it, attaches positive value to deprivation or suffering as such. Sympathy and antipathy become defective principles when approval or aversion is treated as a sufficient reason for judgment, without examining consequences. His target is the elevation of a reaction into an unquestionable standard. He does not thereby deny that sympathy for others can motivate beneficial action. Bentham, *Introduction*, chapter II.

This criticism changes the structure of moral argument. A person saying that a practice is disgusting, impious, degrading, or contrary to custom must explain what follows from that description. Who suffers? What damage occurs? How probable is it? Does prohibition prevent the damage, or produce additional suffering?

The reformer's task consequently includes examining the emotional foundations of condemnation. A feeling may be sincere and socially widespread while remaining a poor guide to what ought to be prohibited.

Bentham also tries to discipline the assessment of consequences. His account of pleasure and pain considers intensity, duration, probability, temporal proximity, subsequent effects, and the number affected. The familiar calculation therefore extends beyond the immediate satisfaction of the person acting. It includes later consequences and their distribution across other people. Bentham explicitly acknowledges that this procedure cannot be carried out exactly before every moral or legislative judgment; it provides a standard toward which practical reasoning should approximate. Bentham, *Introduction*, chapter IV.

For deliberate moral reform, this functions as a demand for accountable comparison. A prohibition must be compared with toleration, education, compensation, or some alternative arrangement. Its claimed benefits must be considered alongside enforcement costs, fear, humiliation, and opportunities for abuse.

The philosophical ambition is clear: moral disputes should become more susceptible to investigation and argument. Nevertheless, specifying the dimensions of a calculation does not establish that all relevant experiences can be measured accurately or compared without controversy. Bentham offers a framework for disciplined judgment, rather than a completed empirical science capable of producing an indisputable answer to every moral question.

The motivational side of his theory is equally important. A proposal for better norms remains ineffective unless it explains why people would follow them. Bentham distinguishes motives from the consequences of the conduct they produce. The same broad motive—financial gain, reputation, affection, or religious commitment—can support beneficial or harmful action in different circumstances. Moral education should therefore avoid assuming that naming a motive settles the evaluation of an action. Bentham, *Introduction*, chapter X.

A desire for reputation, for example, can encourage honesty where honesty earns respect. It can encourage aggression where aggression earns admiration. This illustrates a central implication of Bentham's approach: reform can change the social direction of a motive without eliminating the motive itself.

His psychology should not be reduced to the assertion that everyone is incapable of caring about anyone else. Sympathy and benevolence have a place in his account. At the same time, his institutional proposals take the strength of self-preference seriously. The resulting problem is how to connect the interests people actually pursue with the interests that ought to count in moral evaluation. Philip Schofield's analysis of Bentham and Sidgwick shows why an automatic harmony between personal advantage and general happiness cannot simply be assumed. Schofield, "Sidgwick on Bentham: The 'Double Aspect' of Utilitarianism".

One can reconstruct the practical problem as follows. Reasoning may reveal that cooperation, honesty, or generosity already serves someone's longer-term interests. Where this connection is insufficient, institutions can provide additional incentives. Where the institutions reward harmful conduct, exhortation alone is unlikely to achieve much.

The theory of sanctions explains how moral expectations acquire practical force. Bentham identifies four principal sources of the pleasures and pains influencing conduct:

Sanction	**Source of its motivational force**	**Illustrative operation**
Physical	Ordinary consequences of actions and natural events	Pain resulting from injury or imprudence
Political	Persons exercising organized governmental authority	An officially imposed penalty
Moral or popular	The responses of other people in ordinary social relations	Esteem, disapproval, cooperation, or exclusion
Religious	Anticipated action of a supernatural power	Expected divine reward or punishment

A sanction, in this terminology, includes sources of reward as well as punishment. "Moral sanction" describes a social mechanism; it does not certify that the judgments enforced through it are correct. Bentham, *Introduction*, chapter III.

The last point is fundamental. Public opinion can reward beneficial behaviour, but it can also enforce prejudice. It can shame a person for causing suffering, or shame a person whose conduct is harmless. Consequently, a reformer must examine both the conduct being condemned and the social practice of condemnation itself.

This yields a second level of moral criticism. The person administering blame is also acting and producing consequences. Their conduct requires evaluation just as the original conduct does.

Consider a hypothetical community in which refusing a violent confrontation attracts disgrace. Simply advising individuals to avoid violence leaves a powerful reputational incentive untouched. Effective reform would also have to change what earns respect, how disputes are resolved, and whether refusal still carries social penalties. This example illustrates the logic of Bentham's approach: moral change involves reorganizing the relationship between rules, expectations, and rewards.

Bentham's distinction between private ethics and legislation prevents this programme from becoming a demand to enforce every moral judgment by law. In chapter XVII of the *Introduction*, private ethics means directing one's own conduct. It includes conduct affecting other people; "private" does not mean exclusively self-interested.

His classification there is:

Ethical concern	**Bentham's category**	**Main orientation**
One's own happiness	Prudence	Conducting one's life advantageously
Avoiding injury to others' happiness	Probity	Restraint from harmful conduct
Increasing others' happiness	Beneficence	Positive assistance

Private ethics and legislation share happiness as their objective, but their reach and instruments differ. An action can be morally appropriate without its legal compulsion being beneficial. The distinction also reflects limits to governmental knowledge of individual circumstances. Bentham, *Introduction*, chapter XVII.

The later *Deontology* develops practical ethical guidance within this broader project. Bentham's term should not be confused with the contemporary use of "deontology" for ethical theories opposing consequentialism. His deontology remains connected to utility. Its moral adviser investigates conduct, motives, and circumstances rather than merely repeating inherited commands.

Nor does improving another person automatically justify overriding their wishes. Bentham could be sharply critical of purported beneficence that substitutes someone else's conception of happiness for that of the intended beneficiary. This strand of his thought places an epistemic and practical restraint on moral paternalism: reformers cannot simply assume that they understand other people's interests better than those people do. Mario Rizzo, "The Problem of Moral Dirigisme," discussing Bentham's *Deontology*.

The limits of punishment express the same concern in a more systematic form. Punishment itself creates suffering and therefore needs justification. Bentham identifies four grounds for withholding it:

- It is **groundless** because there is no relevant harm to prevent.
- It is **ineffective** because it cannot prevent the harm.
- It is **too costly** because the suffering it creates exceeds the benefit.
- It is **unnecessary** because a less costly means can achieve the purpose.

These are limits derived from utility, rather than exceptions reluctantly conceded to it. Bentham, *Introduction*, chapter XIII.

For moral reform, this produces two different possibilities. Sometimes a received prohibition lacks justification and should be abandoned. In other cases, criticism of the conduct remains justified, but criminal punishment does not. Education, advice, voluntary cooperation, or other responses may be preferable.

Legal decriminalization and moral approval therefore require separate arguments. Equally, legal permission does not by itself establish that something is beneficial or worthy of praise.

Indirect legislation brings Bentham closest to a general theory of consciously altering the conditions that produce conduct. A direct penal rule threatens a consequence for a prohibited act. Indirect intervention can instead alter the circumstances that make harmful action attractive, possible, or likely.

In his writings on this subject, conduct depends on factors involving inclination, knowledge, and power or opportunity. Preventive intervention may therefore improve information, reduce opportunities for misconduct, remove incentives, or change the surrounding arrangements. The category has shifting boundaries across his writings and should not simply be equated with contemporary "nudging." Michael Quinn, "Jeremy Bentham, Choice Architect".

Angela Marciniak's reconstruction shows how this approach extends prevention beyond the threat of punishment. The relevant question becomes how circumstances generating harmful conduct can be changed before the harm occurs. Marciniak, "'Prevention of Evil, Production of Good'".

A hypothetical administrative example makes the distinction concrete. Suppose an institution rewards personal loyalty, conceals decisions, and imposes few consequences for favouritism. A campaign urging impartiality addresses only part of the problem. Reforming appointments, recording reasons, exposing decisions to scrutiny, and changing career incentives could give impartial conduct a more secure practical basis.

This is an analytical reconstruction of Bentham's method, rather than a specific programme quoted from him. Its significance is that moral improvement becomes partly an institutional achievement. The organization of situations affects whether better conduct is difficult, costly, attractive, or routine.

His codification project adds another element: rules should be made intelligible and accompanied by reasons. Bentham envisaged a comprehensive body of law, the *pannomion*, whose provisions would be systematically justified. The explanation would allow those governed to understand the purposes of the rules and assess their rationale. UCL Bentham Project, *"Legislator of the World": Writings on Codification, Law, and Education*.

In the *Introduction*'s Preface, he describes connected systems of legislative provisions and supporting reasons, capable of correcting and supporting one another. Bentham, *Introduction*, Preface.

The implication for moral change is that explicit justification can weaken dependence on inherited authority. A rule whose purpose has been stated can be challenged when it fails to achieve that purpose. Public explanation makes revision intellectually possible, even though it does not guarantee agreement.

Education operates on the capacities and interests through which people conduct their lives. In *Chrestomathia*, Bentham proposed a school emphasizing useful learning, science, and organized instruction. The project connected knowledge with the enlargement of capacities and sources of satisfaction. It was directed towards the middling and higher ranks and should not be described as an implemented system of universal moral schooling. The proposed school was

never built. Bentham, *Chrestomathia*, in *Works*, volume VIII; Catherine Fuller's study of the project.

Its relevance goes beyond teaching a list of duties. Someone with better knowledge can recognize consequences previously misunderstood. Someone with wider capacities may acquire satisfying activities that alter the attractions of harmful alternatives. Education can therefore change both judgment and the practical environment of choice.

Bentham also treats language as an instrument through which moral judgments become entrenched. His analysis of approving, condemning, and comparatively neutral terms exposes the way evaluation can be built into a description before an argument begins. Calling an action "depraved," for example, may seem to supply a reason when it merely repeats the condemnation requiring justification.

A Table of the Springs of Action develops this concern, while *The Book of Fallacies* examines arguments that obstruct reform through deference, fear, and misleading characterization. Its political relevance lies in showing how apparently persuasive language can protect established interests from examination. Bentham, *The Book of Fallacies*, critical edition edited by Philip Schofield.

From this perspective, deliberate moral change includes reforming the language of moral dispute. A productive description should make the conduct and its consequences available for assessment. Otherwise, the inherited vocabulary repeatedly reproduces the inherited verdict.

The later constitutional theory turns the problem of moral motivation back upon those who govern. Officials cannot be assumed to pursue general happiness simply because they occupy positions of authority. Their interests must be subjected to institutional checks.

Bentham's *Public Opinion Tribunal* names a dispersed body of public judgment rather than an actual state court. Publicity, newspapers, and access to evidence allow official conduct to be examined; reputational consequences can then influence those exercising power. His later defence of public scrutiny thus connects the moral sanction with political accountability. Publicity is not absolute, however: he allows exceptions where disclosure would cause greater harm, including protection for secret voting. Schofield, "Jeremy Bentham on Freedom of the Press, Public Opinion, and Good Government".

This creates a reciprocal structure. Public institutions influence conduct, while public judgment evaluates the institutions and their personnel. Moral reformers themselves become objects of moral and political supervision.

The development was not uniform. In his police writings of 1798–1799, Bentham envisaged publications that would help authorities inform and deliberately mould opinion. Michael Quinn documents the connection between these proposals, preventive policing, and anxiety about revolutionary contagion. Quinn, "Bentham on Preventive Police".

That tension deserves emphasis. The public may be treated as a participant capable of criticizing government, or as an audience whose beliefs administrators seek to manage. Bentham's writings contain resources for both approaches. A sound interpretation must examine their changing relationship rather than assigning his entire career a single democratic or authoritarian character.

His concrete reform proposals reveal how far this framework could carry him beyond established moral judgments.

Sexual morality provides an especially clear example of revising the grounds of condemnation. In the manuscript now known as "Offences Against One's Self," written around 1785, Bentham examines claims that same-sex conduct harms participants, population, marriage, or society. He distinguishes willing participation from coercion and examines the injuries generated by prosecution, including false accusation and extortion. His argument demands evidence of harmful consequences where prevailing opinion had relied heavily on disgust and religious condemnation. The essay remained unpublished until 1978. Bentham, "Offences Against One's Self," edited by Louis Crompton.

The conceptual change is deeper than reducing a punishment. The classification of the conduct as morally wrong itself becomes open to reconsideration.

His later sexual writings extend the analysis to the combined effects of law, religious fear, and public opinion. Removing legal penalties would leave substantial suffering intact if social condemnation and theological terror persisted. These manuscripts therefore illuminate a programme of changing the surrounding moral judgments as well as the law. They were not published in his lifetime, and their historical terminology and scope should not be treated as identical to a contemporary sexual ethic. Schofield, "Gloomy Terrors or the Most Intense Pleasure?".

The distinction between theoretical originality and historical influence is essential here. An unpublished argument can show how radical Bentham's thinking became without establishing that it directly caused a contemporary transformation in public morality.

His treatment of animals illustrates a different mechanism: expanding the range of interests receiving moral consideration. The capacity to suffer becomes a reason to include beings whose interests existing practices neglect. Yet this did not lead him to prohibit every human use of animals. His 1825 discussion of experimentation permits it where sufficiently weighty prospective benefits justify the suffering. Bentham, letter to the *Morning Chronicle*, 4 March 1825; Johannes Kniess, "Bentham on Animal Welfare".

The reforming logic is that familiar exclusions require justification. A society may have failed to apply its professed concern for suffering consistently because it overlooked some of those capable of experiencing it.

Bentham's religious criticism addresses another source of moral authority. In his later writings, he attacks religious arrangements that, in his judgment, sustain needless deprivation, fear, and hostility to pleasure. He also attempts to challenge received Christianity through distinctions between Jesus and subsequent religious authority. These are Bentham's polemical interpretations, rather than conclusions that should be accepted without separate historical examination. Schofield, "Jeremy Bentham: Prophet of Secularism".

For his theory of moral change, the point is institutional: where norms derive force from religious authority, philosophical criticism of individual rules may be insufficient. The authority maintaining them also becomes an object of criticism.

Defence of Usury supplies a published example of challenging an established moralized prohibition. Bentham examines restrictions on lending arrangements through their consequences, including differences in risk and the circumstances of borrowers. The argument asks whether prohibition improves the position of those supposedly protected. Bentham, *Defence of Usury*.

One need not accept his policy conclusions to recognize the method. A practice traditionally grouped under a condemnatory category is divided into its concrete forms, and the effects of regulation are examined alongside the effects of the practice itself.

The Panopticon displays the most forceful institutional side of his reforming ambition. Anticipated observation was intended to alter prisoners' conduct, promote orderly activity, and reduce the need for punishment. Here, reform works through the arrangement of the environment and repeated behaviour. Bentham's own prison scheme was never realized, so it cannot be treated as a successful empirical demonstration of his theory. UCL Bentham Project, "The Panopticon".

The scheme also included scrutiny of the managers. Bentham envisaged public inspection as a safeguard against mistreatment, recognizing that those administering reform needed supervision themselves. Bentham, *Panopticon; or, The Inspection-House*, especially Letter VI.

Nevertheless, it exposes a central philosophical difficulty. Changed behaviour under observation does not necessarily establish changed moral understanding. A person may comply because detection seems likely while remaining unconvinced that the required conduct is right. Institutional discipline, habituation, and reflective moral improvement are different achievements, even where they sometimes reinforce one another.

Bentham's project bears a qualified resemblance to the transition from customary law to deliberate legislation. To assess that resemblance, three meanings of "positivizing" morality should be separated:

Possible meaning	Relationship to Bentham's programme
Making moral rules explicit and giving reasons for them	Strongly present
Deliberately organizing education, incentives, and social responses to establish better conduct	Strongly present
Making a norm morally correct simply through an authoritative decision	Incompatible with the continuing authority of utility as a critical standard

This is an analytical reconstruction of the implications of his writings. It identifies both the strength and the limit of the legislative analogy.

The resemblance concerns conscious intervention. Norms need not remain an opaque inheritance reproduced without examination. They can be articulated, criticized, revised, and supported through planned arrangements.

The difference concerns authority and the mechanics of replacement. A legislature can repeal a statute through a recognized procedure. Popular moral condemnation ordinarily has no equivalent central authority or definitive date of repeal. Its force changes through many

judgments and relationships. Legal reform may assist that process, but cannot by itself ensure that former condemnation disappears.

A further distinction separates the social existence of a norm from its justification. A new moral expectation may become dominant through successful campaigning and still be objectionable on Bentham's principles. Conversely, a well-justified criticism may remain socially marginal.

The theory therefore requires three separate assessments: whether the proposed norm is beneficial, whether it has become accepted, and whether the means used to establish it are themselves beneficial.

Several unresolved problems emerge when the project is examined closely.

First, **the standard of utility cannot be established merely by showing that people pursue pleasure and avoid pain**. A description of motivation does not by itself demonstrate why everyone's welfare ought to count impartially. Bentham recognizes that a first principle cannot receive an ordinary proof from some more fundamental principle, but his confidence that utility supplies the appropriate starting point remains a philosophical commitment. A critic may dispute that commitment even while accepting much of his institutional analysis.

Second, moral reform encounters uncertainty about consequences. Effects are unevenly distributed, experiences differ, and long-term changes are difficult to forecast. An apparently precise vocabulary may conceal assumptions about whose suffering matters, which consequences receive attention, and what counts as an acceptable risk. Deliberate reform therefore needs mechanisms for correcting its own errors.

Third, existing preferences and aversions can create a problem for the reformer. If people suffer because they strongly disapprove of a harmless practice, their distress is a real experience. Yet treating that distress as decisive would allow entrenched hostility to justify continuing the conditions that sustain it. Bentham's criticism of antipathy points towards examining and changing such reactions. The deeper difficulty remains: reform must evaluate some preferences while also treating experienced welfare as its standard.

Fourth, aggregation creates questions about the protection of individuals and minorities. Bentham himself came to criticize the "greatest number" qualification when it encouraged exclusion of a minority's interests from consideration. J. H. Burns shows how the later discussion moves towards the greatest happiness principle understood across the community as a whole. Burns, "Happiness and Utility: Jeremy Bentham's Equation".

Counting minority suffering is essential, but it does not automatically resolve every conflict. An aggregative theory still faces the question of whether sufficiently large benefits elsewhere can justify severe losses imposed on particular people. Bentham's answer must be developed through consequences, security, and institutional protections; it does not begin from absolute rights immune to consequential assessment.

Fifth, the reformer's power requires scrutiny. Who selects the evidence, defines the alternatives, controls publication, and determines which conduct attracts disgrace? Bentham's later public-opinion theory provides an important answer by subjecting governors to criticism. But public opinion itself can be poorly informed, unevenly organized, or dominated by powerful speakers.

The problem of controlling the reformer consequently extends beyond the state to those capable of directing reputational rewards and punishments.

Finally, changing moral belief, producing beneficial conduct, and securing obedience should remain distinct standards of success. A successful prohibition may establish compliance while preserving resentment. An educational intervention may improve judgment without immediately changing institutions. A reform of incentives may reduce harmful conduct without producing deeper concern for others. Bentham's work is particularly valuable because it supplies tools for analysing all three processes; it is also open to criticism where institutional success is allowed to stand too readily for moral development.

For close study of this particular question, the most productive reading sequence is the *Introduction*'s chapters II, III, XIII, and XVII; the manuscript-based *Deontology* and *Table of the Springs of Action*; the writings on indirect legislation; and the later texts on public opinion and constitutional accountability. Read together, they reveal the enduring question at the centre of Bentham's enterprise: **how can human beings deliberately improve the norms governing their lives while also making the authority, methods, and consequences of the reform itself subject to criticism?**

****

Robert Owen

Robert Owen (1771–1858) developed a comprehensive theory of deliberate moral transformation: society could improve human conduct by consciously reorganizing the conditions through which people acquire their character, convictions, habits, and desires. Education, working conditions, economic security, family arrangements, and public institutions were therefore instruments of moral change. His distinctive contribution was to connect an account of character formation with practical experiments and, eventually, a proposal for reconstructing the whole social order.

The expression *conscious moral change* describes his project well, although it was not his standard technical term. Owen generally spoke of the formation of character, the rational organization of society, and the creation of a *new moral world*. His project combined three questions that should be distinguished analytically: **What makes people behave as they do? What kind of conduct promotes human happiness? How can institutions reliably produce that conduct?**

These questions remained connected throughout his career, but his answers became more radical. The enlightened manufacturer who initially improved conditions within an existing enterprise gradually became an advocate of cooperative communities and a fundamental revision of property, religion, and marriage. Reading the later programme back into his earliest experiments would obscure this development. Ophélie Siméon, "Robert Owen: The Father of British Socialism?"

His principal writings provide the following framework:

Work or experiment	Place in the development of his theory
Management of New Lanark, beginning in 1800	Practical investigation of how living conditions, education, and management affect conduct.
A New View of Society, essays published in 1813–1814, with a collected edition in 1816	Formulation of character formation as the foundation of moral and governmental reform.
Observations on the Effect of the Manufacturing System, 1815	Analysis of industrial organization as a producer of harmful habits and relationships.
Report to the County of Lanark, submitted in 1820 and published in 1821	Development of cooperative settlements as instruments of economic and moral reconstruction.

Work or experiment	Place in the development of his theory
New Harmony, 1825–1827	Attempt to establish a cooperative community, exposing difficulties in the transition to the proposed order.
The Book of the New Moral World, seven parts, 1836–1844	Systematic presentation of his mature account of human nature and rational society.
Outline of the Rational System of Society, 1841	Concise integration of psychology, morality, education, religion, and constitutional organization.

The dates of individual editions sometimes explain apparent discrepancies in bibliographies—for example, the Lanark report was presented in 1820 but published in 1821. Bibliography of Owen's writings

Owen's starting point was the experience of industrial society. He observed that growing productive power could coexist with exhausted workers, neglected children, insecurity, drunkenness, deception, and antagonism between employers and employees. He interpreted these conditions as evidence that an economic system also operates as a system of moral education.

In *Observations on the Effect of the Manufacturing System*, he argued that competitive commercial life trained people to pursue immediate monetary gain. Employers learned to treat workers as instruments of profit; workers experienced relationships increasingly stripped of durable mutual obligations. Children subjected to prolonged factory work acquired habits under conditions that damaged their physical and intellectual development. Society subsequently condemned conduct that its own arrangements had helped generate. Owen, *Observations*, second edition, 1817

The theoretical importance of this argument extends beyond criticism of harsh employment. **For Owen, morality is continuously being produced through ordinary social organization.** A factory, household, marketplace, or school teaches people what to expect from others and how to secure their own well-being. These lessons operate even when nobody formally announces an ethical doctrine.

His proposal for conscious change can therefore be reconstructed as a shift from largely unexamined character formation to deliberate, informed character formation. The reformer must identify the practices that generate harmful dispositions and establish practices capable of producing better ones. This is why moral preaching alone could never accomplish his objective: daily experience might continually teach the opposite of the sermon.

The foundation of this approach was Owen's doctrine that character is formed through causes that individuals did not originally choose. A child does not choose its bodily constitution, parents, language, religion, social position, or earliest experiences. Yet those circumstances deeply shape the beliefs and habits through which it later interprets the world.

Owen's mature formulation is more nuanced than the claim that human beings are completely blank slates. The five fundamental facts presented in *The Book of the New Moral World* can be paraphrased as follows:

1. Character develops through the interaction of a person's original constitution and the circumstances encountered throughout life.
2. Feelings and convictions arise independently of a direct act of will.
3. Feelings and convictions generate the will that leads to action.
4. Individuals differ in their original constitutions, and education cannot make them precisely identical.
5. Nevertheless, the circumstances surrounding infancy can produce very different degrees of human development.

His accompanying twenty laws elaborate the formation of beliefs and habits, individual differences, pleasure and pain, and the balanced exercise of human capacities. Owen, *The Book of the New Moral World*, Part I, 1836, especially pp. 1–2

Two propositions must consequently be kept separate. Owen recognizes variation between individuals, while attributing enormous transformative power to their environments. His theory is deterministic without requiring that everybody have the same abilities or personality.

The strongest version of his argument, however, goes considerably beyond acknowledging social influence. He repeatedly treats correct character formation as something that can be achieved with near certainty. This creates a substantial distance between a plausible observation—that circumstances profoundly influence development—and his much stronger expectation that rational arrangements can reliably eliminate the principal sources of moral evil.

The doctrine also changes the meaning of moral responsibility. Owen attacks the assumption that individuals are the ultimate authors of the character for which society rewards or punishes them. His striking example in the second essay of *A New View of Society* asks readers to exchange the upbringing of judges and criminals. Had their formative circumstances been reversed, some people occupying the judicial bench might have appeared before it as defendants, and some condemned defendants might have become judges. Owen, *A New View of Society*, Second Essay

The force of this example lies in its challenge to moral self-satisfaction. Successful and respectable people commonly interpret their position as evidence of personal superiority. Owen asks how much of that superiority depends on opportunities and training unavailable to those they condemn.

His position does not erase the distinction between beneficial and harmful behaviour. It changes the explanation of harmful behaviour and the justification for responding to it. We can distinguish three questions:

- Did an action cause harm?
- What conditions and dispositions produced it?
- Does its author deserve suffering because they ultimately created those dispositions?

Owen can answer the first affirmatively while directing attention to the second and rejecting the underlying assumption of the third. His practical orientation is preventive and corrective: remove causes of injury, develop better habits, and protect people against immediate harm.

There is nevertheless a tension between his sweeping rejection of punishment and the restraints he accepted during reform. His early writings allowed measures to prevent immediate injury,

and New Lanark continued to employ discipline. The mature ideal of dispensing with individual rewards and punishments depended on the successful prior formation of people within the new institutions. It should not be interpreted as an instruction to remove every restraint immediately from existing society.

A particularly original feature of Owen's theory is that **understanding the causes of character is itself supposed to change moral feeling**. Knowledge of how another person became what they are should weaken hatred, resentment, and vindictiveness. The offender, religious opponent, or foreign enemy becomes intelligible as a human being formed under different circumstances.

Thus, conscious moral change operates through two connected processes. One changes the environment that forms conduct. The other changes how people understand and respond to conduct. A society may improve materially while retaining an intensely punitive moral culture; conversely, individuals may acquire compassionate convictions while continuing to live under institutions that generate insecurity and conflict. Owen's programme seeks to transform both.

His determinism also applies to the powerful. Priests, rulers, and wealthy proprietors have themselves been educated into mistaken beliefs about their interests. In his *Catechism*, Owen maintains that even apparently privileged groups lose the fuller benefits of a rational social order. This helps explain his hope that different classes could be persuaded to support a common transformation. Owen, *Catechism of the New Moral World*, especially pp. 5–7

This extension gives his theory a consistent universalism: nobody is excluded from causal explanation or prospective improvement. It also produces a weakness. Those enjoying power may understand its immediate advantages accurately and defend them deliberately. Their opposition need not result simply from misunderstanding the superior happiness available to everyone.

Owen nevertheless requires a standard by which one kind of character counts as better than another. His central criterion is happiness, understood through physical well-being, intellectual development, agreeable social relations, and the exercise of human capacities. His conception of happiness includes temperance: satisfying human needs appropriately, rather than pursuing unlimited stimulation. The *Catechism* makes this relationship between needs, balanced activity, and happiness particularly explicit. Owen, *Catechism*, pp. 2–4

His argument also links individual happiness to the happiness of others. In the first essay of *A New View of Society*, he presents enlightened self-interest as a basis for conduct that advances the community's well-being. Cooperation should become something people experience as beneficial, rather than a duty continually imposed against their interests. Owen, *A New View of Society*, First Essay

Analytically, however, three different claims are involved: that people desire happiness; that their happiness often depends on others; and that a rationally organized society can reconcile everybody's interests. The first two do not establish the third. Disputes concerning authority, affection, recognition, scarce opportunities, and preferred ways of living can persist even when material insecurity is reduced.

Education supplied Owen's most developed practical answer to the problem of forming cooperative dispositions. He understood education broadly, encompassing early relationships, bodily activity, emotional experience, recreation, and habitual participation in common life.

In the third essay of *A New View of Society*, he describes the playground as an important setting for moral formation. Children should learn to avoid injuring their companions and actively contribute to their happiness. Older children, already accustomed to this practice, would help establish it among newcomers. Repeated experience would make cooperative behaviour familiar and eventually habitual. Owen, *A New View of Society*, Third Essay

The proposed mechanism combines instruction, example, practice, and experienced satisfaction. A child hears a simple rule, observes it in others, acts upon it, and experiences the advantages of reciprocal kindness. Moral learning becomes part of ordinary interaction.

This approach also explains Owen's concern with pleasant learning and recreation. Where education is associated with fear and humiliation, children may learn avoidance, concealment, or submission. Where learning is associated with curiosity and enjoyment, they have reasons to participate willingly. Music, dancing, conversation, objects, and the observation of nature could therefore contribute to moral development.

Ian Donnachie's study of the New Lanark schools shows both the breadth of these practices and their disciplinary aspects. Education encouraged sociability and tolerance, while also valuing order, uniformity, and tractable behaviour. The historical evidence supports substantial educational innovation; it does not establish that Owen had discovered an infallible method of producing moral character. Donnachie, "Education in Robert Owen's New Society"

The Institute for the Formation of Character, opened in 1816, embodied this enlarged educational ambition. At New Lanark, schooling and recreation formed part of a wider programme that included improved living conditions and restrictions on young children's factory work. Children under ten were excluded from mill employment, and education continued through evening provision. New Lanark Trust, account of Owen's reforms

Yet New Lanark also demonstrates why moral improvement and freedom cannot simply be treated as equivalent. The establishment remained under managerial authority. Workers were supervised, resident behaviour was regulated, and cleanliness and temperance were matters of organized intervention.

The famous *Silent Monitor* makes the tension especially visible. A coloured block displayed an assessment of each worker's conduct, and assessments were recorded in books of character. Even when harsh punishment was avoided, conduct remained publicly classified and subjected to reputational pressure. The Open University's account also cautions that relatively little direct testimony from the workforce survives. Open University, *Robert Owen and New Lanark*

A sympathetic interpretation is that Owen was employing comparatively mild methods within a transitional setting. A critical interpretation is that a benevolent manager's conception of good character could include precisely those qualities that made workers easier to govern. Both interpretations deserve consideration.

The resulting question is fundamental to conscious moral reform: **Who decides which dispositions should be cultivated, and how can those undergoing formation challenge that**

decision? A programme can improve welfare while leaving its recipients with limited authority over the definition of improvement.

Owen's movement towards cooperative organization followed from his belief that education could not be isolated from economic life. If people were continually exposed to insecurity and antagonistic interests, schooling would struggle to establish durable cooperation.

The *Report to the County of Lanark* therefore connects employment, production, distribution, and character. Owen proposed settlements combining agricultural and other productive activities with shared arrangements and education. His economic diagnosis emphasized the failure to organize and distribute the benefits of expanding productive power. Labour-based exchange and cooperation were intended to reduce the scope for deceptive bargaining and conflicting interests. Owen, *Report to the County of Lanark*

The moral logic can be reconstructed as follows:

Institutional condition	**Dispositions it tends to encourage in Owen's account**	**Proposed direction of change**
Insecurity over basic provision	Anxiety, defensive accumulation, mistrust	Reliable provision and useful employment
Rivalry for economic advantage	Concealment, deception, antagonism	Cooperative production and shared benefit
Exhausting work	Irritability, narrowed development, harmful forms of escape	Work combined with education and recreation
Sectarian upbringing	Hostility towards people educated differently	Knowledge of how beliefs are acquired
Punitive treatment	Fear, resentment, concealment	Prevention and corrective care
Repeated cooperative activity	Trust and concern for others	Make cooperation a normal feature of daily life

This table expresses the causal structure of his programme; it should not be read as a set of universally verified empirical laws.

Owen's socialism consequently has a distinctive moral purpose. Collective arrangements are supposed to change the relationship between personal advantage and social obligation. People should cease to encounter kindness as a recurrent sacrifice demanded by an otherwise competitive world.

This also explains his attraction to relatively comprehensive communities. A cooperative school surrounded by radically different practices in employment, domestic life, and commerce would face persistent counterinfluences. The community offered a setting in which several formative institutions could reinforce one another.

Such comprehensiveness has an obvious cost. As more of life becomes part of an integrated educational project, fewer spaces remain outside the community's authoritative conception of good development. Owen's effort to remove contradictory influences therefore raises questions about intellectual independence and the legitimacy of alternative ways of life.

Religion constituted another major field of revision. Owen increasingly attacked established religious systems because they divided people, encouraged doctrines of deserved punishment, and perpetuated beliefs acquired before critical reflection was possible. His public attack on religion in 1817 marked an important stage in his radicalization. Siméon, "Robert Owen: The Father of British Socialism?"

His mature position nevertheless included a *Rational Religion*. It acknowledged an incomprehensible originating power while locating religious practice in kindness and the promotion of human happiness. A simple description of his programme as atheistic would therefore be inadequate. The central transformation concerned the authority of sectarian doctrines and the moral consequences drawn from them. Owen, *Outline of the Rational System of Society*, pp. 7–9

Marriage provides an especially concrete example of Owen's attempt to replace established moral judgments. His argument begins with the involuntary character of affection. People cannot simply command themselves to love a particular person permanently. Institutional arrangements that demand the continuation of affection regardless of experience may generate dishonesty and suffering.

In his marriage lectures, Owen defended lasting affectionate unions and challenged arrangements sustained by wealth, dependence, or compulsory permanence. The appendix to the 1840 edition proposed procedures for marriage and separation, accessible divorce, protection against the disgrace attached to remarriage, and arrangements for children. His programme therefore cannot adequately be described as indiscriminate opposition to marriage. Owen, *Lectures on the Marriages of the Priesthood of the Old Immoral World*, especially appendix, pp. 86–90

The theoretical significance is substantial. A practice previously condemned—ending an unhappy union—could become legitimate under a different moral criterion. Conversely, conduct previously praised—maintaining a marriage despite destroyed affection—could become open to criticism.

Here Owen was changing more than the means of obtaining familiar virtues. He was revising the content of moral approval and disapproval. Affection, sincerity, well-being, and care for those affected supplied grounds for reconsidering inherited rules. Legal procedures and collective provision would then help make the revised morality practicable.

His theory also required a strategy of transition. People educated under existing arrangements could not be expected immediately to possess the dispositions required by the new society. Nor could a reformer simply assume that an institutional blueprint would implement itself.

Owen relied on persuasion, practical demonstration, gradual institutional change, and the support of people controlling resources. In the fourth essay of *A New View of Society*, he advocated revising laws that encouraged harmful conduct and developing a national system of character formation. Government should address the conditions producing crime. Owen, *A New View of Society*, Fourth Essay

His strategy can be understood as an attempt to initiate a cumulative process: establish workable arrangements, show their advantages, gain confidence, extend their scope, and educate

succeeding generations within them. Transitional compromises were therefore part of the project.

Chris Rogers has emphasized this distinction between the arrangements needed to begin transformation and those Owen expected in a fully developed cooperative society. The presence of investors, temporary hierarchies, or other inherited practices does not by itself prove that Owen had abandoned his eventual objective. Rogers, "Robert Owen, Utopian Socialism and Social Transformation"

This is an important qualification, but it cannot settle every objection. A transitional authority may perpetuate itself. A reformer must explain how those initially subjected to guidance acquire the capacity and institutional opportunity to govern the process themselves.

New Harmony exposed the practical difficulty. Owen and William Maclure acquired the Indiana settlement in 1825; Owen's communal experiment dissolved in 1827. Educational and scientific activity continued afterwards, so the failure of the cooperative organization should not be confused with the disappearance of the settlement or all its achievements. Indiana Archives and Records Administration, New Harmony

Personal conflicts and difficulties organizing productive labour and agriculture contributed to the breakdown. These problems resist explanation through a single cause. University of Southern Indiana, historical account of New Harmony

For the theory of moral change, the experiment exposes a circularity: the community required cooperation among adults before its institutions had produced the cooperative dispositions on which success supposedly depended.

This observation does not refute the proposition that environments influence character. It challenges the expectation that a new setting can rapidly overcome entrenched habits while simultaneously solving problems of production, authority, and disagreement. Moral development and institutional viability have different time scales.

Owen's mature constitutional proposals attempted to give his programme an explicit organizational form. The 1841 *Outline* envisaged councils organized by age, with domestic administration assigned to members aged thirty to forty and external affairs to those aged forty to sixty. It also proclaimed freedom of conscience, equal rights and education for both sexes, and procedures for dealing with failures by governing councils. His constitutional imagination consequently included both shared participation and extensive administration of social life. Owen, *Outline*, pp. 10–14

A troubling provision treated people who became physically, intellectually, or morally diseased as requiring hospital care until recovery, using the mildest effective treatment. Owen did not simply equate every disagreement with illness, and his protection of conscience must be taken seriously. Nevertheless, the provision raises a clear danger: an authority convinced that rational institutions necessarily produce rational conduct may interpret persistent deviation as pathology.

This illustrates a broader difficulty. Replacing punitive language with therapeutic language does not automatically reduce the power exercised over individuals. Much depends on who

determines what counts as illness, what treatment is justified, and how the person concerned can contest those judgments.

For the broader question of whether morality can be consciously reconstructed in a manner comparable to legislation, Owen is especially revealing. His programme contains several features resembling deliberate norm creation:

- Existing customs are subjected to explicit criticism.
- Alternative principles are formulated in writing.
- Schools and associations transmit those principles.
- Institutional arrangements support the expected conduct.
- Public judgments concerning blame, respectability, marriage, and cooperation are deliberately revised.

Yet calling this the *positivization of morality* requires qualification. Owen did not regard moral principles as valid merely because an authorized body enacted them. He claimed to discover enduring truths about human nature and happiness, from which appropriate social arrangements could be derived.

There are therefore two levels in his theory. Human institutions are consciously made and can be changed. The ultimate justification for changing them lies in supposedly objective laws of human nature. His position combines deliberate institutional construction with confidence in a universally valid moral foundation.

This also distinguishes his project from an open-ended system of moral legislation. A legislature can acknowledge that it has selected one arrangement among several and may later replace it. Owen often writes as though the fundamental principles of the rational order have already been conclusively established. Implementation can improve, but the foundations are presented with much greater finality.

The strength of this position is that it denies inherited customs an automatic claim to obedience. Its weakness is that the reformer's preferred norms can acquire the authority of scientific necessity. The question then becomes whether opponents are mistaken about facts, defending another legitimate value, or merely resisting the reformer's project.

Comparison with Bentham helps locate Owen's distinctive emphasis. Both assessed institutions by their consequences for happiness, and Bentham became associated with the New Lanark enterprise. Owen gave exceptional prominence to the formation of the dispositions with which people entered social relationships. His concern extended to the production of motives themselves. The difference is one of emphasis: both thinkers were interested in institutions, incentives, and human development. Open University, discussion of Owen and Bentham

John Stuart Mill identifies an important corrective to the strongest interpretation of Owen's doctrine. In his *Autobiography*, Mill explains that accepting the formation of character through circumstances need not imply helplessness. A person's desires and decisions can themselves alter circumstances and influence future character. Causal explanation remains compatible with meaningful self-development. Mill, *Autobiography*, Chapter V

Applied to Owen, this suggests that conscious moral change should include people's reflective participation in forming themselves. They can learn to examine and revise the habits they have

acquired. The practical question becomes how education develops that capacity, including the capacity to question the educator.

Marx's third thesis on Feuerbach raises the related problem that the educator must also be educated. A theory that assigns enlightened reformers the task of changing everybody else risks placing one group conceptually above the society whose members are all supposedly shaped by circumstances. Marx, *Theses on Feuerbach*, Thesis III

Owen can answer part of this objection: reformers themselves may acquire insight through distinctive experiences, just as anybody acquires other convictions. Determinism does not make discovery impossible. The more difficult problem concerns authority. Acquiring an insight does not establish a right to organize other people's lives comprehensively.

Several further limitations follow from the structure of his argument.

First, explaining the origin of a moral conviction does not establish its truth or falsity. An inherited belief can be reasonable, while a consciously invented principle can be mistaken. Owen's account of social formation therefore requires an independent justification of the values by which he evaluates its results.

Second, eliminating some causes of wrongdoing does not establish that all significant conflict can disappear. Economic security and humane education can improve relationships while leaving disagreements about distribution, recognition, intimacy, religion, and authority. A durable moral order needs procedures for handling disagreement as well as institutions for encouraging cooperation.

Third, observable conformity is not sufficient evidence of moral conviction. A worker may behave orderly because of supervision; a child may repeat a rule because the teacher expects it. Assessing conscious moral change requires distinguishing public compliance, habitual disposition, reflective acceptance, and freedom to dissent.

Fourth, Owen's experiments were not controlled demonstrations of his universal theory. New Lanark involved an unusually capable manager, a particular business, inherited institutions, and a specific population. Its successes provide evidence about possibilities under those conditions. They cannot establish that an entire society could be reconstructed with predictable results. Conversely, New Harmony's failure cannot establish that every cooperative arrangement must fail.

Modern historical work has consequently become more attentive to the people and institutions surrounding Owen. Siméon's research emphasizes New Lanark as a community with an earlier history, inherited practices, and residents capable of resisting managerial intervention. This complicates the image of a reformer shaping an otherwise passive population. Mark Allison's review of Siméon's *Robert Owen's Experiment at New Lanark*

Owen's most durable contribution is his insistence that moral demands must be examined together with the conditions under which people are expected to satisfy them. A society that requires honesty while rewarding deception, demands responsible parenting while exhausting parents, or preaches cooperation while intensifying insecurity creates obstacles to its own professed morality.

His more ambitious claim—that a correctly organized environment could establish universal moral harmony—remains much less persuasive. The enduring research problem opened by his work is therefore how to make the institutions that form character themselves answerable to those whose character they shape. That problem connects education, economic organization, moral philosophy, and constitutional limits in a way that remains central to any serious theory of conscious moral change.

****

Auguste Comte

Auguste Comte developed one of the nineteenth century's most comprehensive programs for the deliberate transformation of morality. His ambition was to reconstruct the social processes through which people acquire their convictions, cultivate their feelings, recognize their obligations, and judge one another. Moral philosophy, education, economic organization, public opinion, historical memory, and religious practice therefore became interconnected parts of his project.

The central problem, as he understood it, was how to make concern for others prevail over the stronger impulses of self-interest. His answer combined a scientific understanding of human beings with the systematic cultivation of attachment, gratitude, reverence, and benevolence. The eventual institutional expression of this program was his **Religion of Humanity**.

This makes Comte especially significant for the history of conscious moral change: he attempted to explain both **what a new morality should contain** and **how an entire society could be brought to live by it**.

The project developed across his career, with a substantial change in emphasis during the 1840s. The principal stages are these:

Period and principal work	Contribution to deliberate moral transformation
Plan of the Scientific Work Necessary for the Reorganization of Society (1822)	Establishes the need for intellectual reconstruction before durable social reorganization.
Course of Positive Philosophy (1830–1842)	Develops the scientific and sociological framework for understanding social order and historical change.
Discourse on the Positive Spirit (1844)	Explicitly argues for morality grounded in knowledge of humanity rather than theology.
A General View of Positivism (1848)	Connects moral regeneration with education, women, workers, art, and a new spiritual authority.
System of Positive Polity (1851–1854)	Constructs the mature social and religious system.
Catechism of Positive Religion (1852)	Presents the beliefs, worship, obligations, and practices of the Religion of Humanity.

These works belong to an enduring project of social reconstruction, although the later writings give affection and religion a prominence that disturbed several early admirers. Michel Bourdeau, "Auguste Comte."

Comte associated his emotional transformation with his relationship with Clotilde de Vaux, whom he met in 1844 and who died in 1846. Her memory became central to his religious imagination. Mary Pickering nevertheless cautions against accepting Comte's autobiographical explanation uncritically: the intellectual development of his later system cannot simply be reduced to this relationship. His longstanding concerns about social cohesion and moral authority also matter. Pickering, "Clotilde de Vaux and the Initial Encounter with Comte."

The interpretive consequence is important. Treating his later religion as merely an eccentric personal addition prevents us from seeing how it addresses a problem left unresolved by scientific explanation: **how knowledge becomes an enduring motive for conduct.**

Comte's starting point was the diagnosis of a crisis in the foundations of social authority. In his 1822 *Plan*, the declining theological and feudal order coexists with an emerging order that has not yet acquired a coherent intellectual organization. Destruction and reconstruction proceed together, but criticism has advanced further than the formation of shared principles. Continued revolutionary negation consequently becomes an obstacle to consolidation. Comte, *Plan*, introduction.

His concern was therefore deeper than disagreement over particular moral rules. A society might still possess many decent individuals while losing agreement about why obligations bind, which authorities deserve confidence, and what purposes social cooperation should serve.

This diagnosis distinguishes three questions that are easily confused:

- What moral convictions do individuals hold?
- What institutions reproduce those convictions?
- What common understanding makes those institutions authoritative?

A reform that addresses only the first question may produce persuasive books without changing ordinary conduct. A reform that addresses only the second may preserve institutions whose teachings have lost credibility. Comte wanted a reconstruction encompassing all three.

Mary Pickering describes the intended sequence as intellectual revolution, moral renewal, and subsequent social and political restructuring. Its purposes included moderating class conflict and establishing a more peaceful and just social order. Pickering, "Comte, Auguste."

This sequence should be understood as a claim about the foundations of lasting change, rather than a claim that institutions never influence beliefs. A constitution cannot, by itself, supply the trust and commitment needed for its operation.

His historical theory supplied an argument for changing the foundations of morality while preserving some inherited moral achievements. Comte's law of three stages distinguishes theological explanations, metaphysical explanations through abstract entities, and positive explanations through observable relations and regularities. The stages concern the development of thought; they should not be treated as three perfectly synchronized historical blocks.

In the *Discourse on the Positive Spirit*, he argues that Christianity—especially medieval Catholicism—helped establish morality's independence from political expediency. But morality's continued dependence on increasingly disputed theology now threatened its

authority. He also recognizes that the discrediting of a doctrine can lead people to reject valuable precepts associated with it. Comte, *Discourse on the Positive Spirit*.

The resulting reform principle is selective. The collapse of a religious justification does not establish that every practice once supported by it is worthless. Equally, the historical usefulness of an institution does not establish its permanent validity.

For example, gratitude, fidelity, and assistance to the vulnerable can be reconsidered in terms of their contribution to human relationships. Their moral significance need not disappear when their theological explanation is abandoned.

Comte's position thus combines historical continuity with deliberate reconstruction. He wanted to identify achievements embedded in earlier moral systems, supply them with a new justification, and integrate them into a more comprehensive order.

The difficulty is that his philosophy of history already identifies the direction in which reconstruction ought to proceed. This creates a recurring danger: his preferred moral future can appear to possess the authority of a discovered historical necessity.

His moral psychology explains why an intellectual revolution would be insufficient on its own. Comte distinguishes affective impulses, intellectual functions, and practical capacities. Within the affective domain, he contrasts self-regarding tendencies with three principal forms of altruistic feeling: attachment, veneration, and benevolence. His detailed classification drew on nineteenth-century cerebral physiology and phrenology. These biological claims cannot be accepted as established neuroscience. Garry W. Trompf, "Comte's Classification of the Sciences."

The underlying philosophical question remains intelligible even when that physiology is rejected: **what makes a person want to act morally?**

A person may understand that cooperation benefits society while continuing to prioritize personal advancement. Knowledge of social interdependence does not automatically generate gratitude, restraint, or concern. Moral change must therefore alter the relative strength and organization of motivations.

This also clarifies the role of reason. Reason can disclose consequences, correct mistaken beliefs, compare courses of action, and stabilize commitments. But reasoning requires ends toward which it works. Comte's mature position assigns the direction of life to social affection and gives intellect the task of making that direction informed and effective.

There are two distinct propositions here:

1. Human action ordinarily depends upon motives that are not themselves products of detached reasoning.
2. Benevolent motives ought to govern the use of intelligence.

The first is a psychological proposition; the second is a normative commitment. Comte's system often brings them close together, but their justification requires different arguments.

Altruism is the organizing principle of his intended moral transformation. Comte introduced the term to designate other-directed feeling and conduct. Its original meaning encompassed both sympathetic emotional responses and action benefiting others; it was broader than an occasional act of charitable self-sacrifice. Suzanne Keen emphasizes this connection between feeling and action, which later discussions of empathy and altruism often separate. Keen, "'Altruism' Makes a Space for Empathy, 1852."

His maxim *vivre pour autrui*—live for others—therefore concerns the orientation of a life.

Analytically, this can be understood at three levels:

- **Attention:** whose needs and experiences enter a person's awareness?
- **Motivation:** how much weight do those needs receive?
- **Practice:** what sustained conduct follows?

A reform can succeed at one level and fail at another. Someone may become more aware of suffering without helping; another may comply with charitable expectations mainly to preserve reputation. A serious account of moral transformation must explain the connections between awareness, motivation, and action.

Comte's strength is that he does not regard moral instruction as complete once a person can repeat correct principles. His vulnerability is the tendency to assume that the coordinated cultivation of sympathy will resolve conflicts among obligations more decisively than it actually can.

Concern for one's family, fellow citizens, humanity at large, and future generations can pull in different directions. Altruism identifies an orientation without automatically providing an adequate method for resolving every conflict within it.

Comte explicitly describes moral development as susceptible to organized intervention. In *A General View of Positivism*, he argues that social life already restrains selfish impulses and exercises sympathetic ones. Deliberate individual and collective action can strengthen favorable influences and weaken unfavorable ones. He calls this the art of morals, acknowledges its limits, and distinguishes it from merely spontaneous development. Comte, *A General View of Positivism*, chapter II.

This is the clearest answer to the question of whether he possessed a theory of **conscious** moral change: he expressly proposed converting an existing developmental tendency into the object of systematic practical effort.

His reasoning can be reconstructed as follows:

1. Human beings possess several competing dispositions.
2. Social circumstances repeatedly activate some dispositions more than others.
3. Repeated activation can establish durable habits.
4. Institutions can organize those circumstances.
5. Consequently, moral dispositions can be cultivated through deliberate institutional arrangements.

Consider an illustrative application. A school might praise individual success while teaching cooperation in its ethics lessons. From a Comtean perspective, the ordinary distribution of recognition could undermine the formal lesson. Moral reform would require examining the school's routines, examples, honors, and relationships alongside its explicit instruction.

The same reasoning applies to workplaces and public life. A society cannot confidently expect solidarity if its most visible rewards consistently celebrate domination, accumulation, and competitive victory.

The empirical question, however, remains open in each case: which arrangements produce durable concern, which generate outward conformity, and which provoke resistance? Comte offered a general program rather than experimental proof of every mechanism.

The family matters because moral attachments must develop through concrete relationships. The intended movement toward humanity-wide concern begins with relationships in which another person's welfare becomes emotionally significant.

This yields a plausible developmental insight: it is difficult to build stable concern for an immense abstraction without experiences of caring, dependence, gratitude, and responsibility.

But the transition is not automatic. Family solidarity can become hostility toward outsiders, favoritism, or the protection of inherited advantage. Philippe Steiner shows that Comte recognized domestic egoism, including the moral distortions associated with inheritance and parental ambition. Steiner, *Comte, Altruism and the Critique of Political Economy*.

The educational task is therefore more demanding than strengthening attachment. It must also **extend and discipline attachment**.

A useful analytical distinction is between learning that *someone matters to me* and learning that *other people can matter even when they are not mine*. The first supplies emotional force; the second requires reflection, imagination, encounters, and institutions capable of challenging partiality.

This is one point at which Comte's project can be assessed independently of his religious conclusions. His attention to the development of motives remains valuable even if one rejects the idea that a single spiritual authority should supervise that development.

His treatment of women exposes a fundamental tension within the proposed reform. Comte assigned women a privileged role in moral cultivation while defending their exclusion from important forms of public authority. Vincent Guillin's study of the Comte–Mill debate shows how biological and phrenological assumptions supported this defense of sexual hierarchy. Guillin, "Le penchant biologique de la sociologie comtienne : La question de l'égalité des sexes."

The distinction between moral esteem and institutional equality is decisive. A group may be described as indispensable, admirable, or morally superior while being denied the power to determine the conditions under which it performs its assigned functions.

This creates a methodological problem for conscious moral reform. The reformer must decide which existing arrangements are contingent and revisable, and which express supposedly

permanent features of human nature. If inherited expectations are classified as biological necessities, scientific language can protect conventions from criticism.

Comte is consequently both an advocate of extensive moral reconstruction and an example of its selective limits. His willingness to reconsider religious authority did not produce an equivalent willingness to reconsider every established hierarchy.

The Religion of Humanity supplies the comprehensive institutional form of his mature program. Its structure joins doctrine, worship, and a regulated way of life. Andrew Wernick emphasizes that these form an interconnected religious organization rather than a collection of independent philosophical propositions. Wernick, *Auguste Comte and the Religion of Humanity*.

Their respective functions can be expressed as follows:

Component	Intended contribution to moral change
Doctrine	Gives people a common interpretation of their situation and obligations.
Worship	Repeatedly engages the feelings that sustain commitment.
Regulation of life	Connects convictions and feelings to ordinary conduct.

The conceptual innovation lies in their coordination. A convincing doctrine can remain emotionally weak; powerful collective emotion can lack direction; disciplined conduct can persist without understanding. Comte tries to make each component support the others.

Humanity becomes the common object of knowledge, love, and service. Its historical dimension places each individual within a succession of generations. Language, knowledge, skills, and institutions precede any particular person and make that person's achievements possible.

One analytical implication is a challenge to the image of the entirely self-created individual. Recognition of inherited benefits can support gratitude and responsibility toward successors.

Yet dependence alone does not settle what is owed, to whom, or under which conditions. The historical inheritance also includes oppression, error, and exclusion. Gratitude needs to coexist with the capacity to criticize what has been transmitted.

Moreover, Comte's Humanity is a morally selected collective rather than simply the biological totality of human beings. That raises an additional question: who determines which lives and contributions deserve incorporation into its authoritative memory?

Ritual is intended to make moral commitments recurrent and emotionally present. The *Catechism* sets out private and public worship and nine social sacraments marking stages from birth and education to death and posthumous incorporation. It also explicitly states that these sacraments should be legally optional, with parallel civil institutions available. Comte, *Catechism of Positive Religion*, especially the discussion of social sacraments.

The significance of this arrangement is clearer if we ask what rituals do. They allocate attention, connect personal events to public expectations, establish occasions for remembrance, and make commitments visible to others.

A ceremony associated with entering a profession, for example, can represent work as a responsibility toward people who depend on its competent performance. A commemoration can change which achievements a community admires.

This is a theory of moral change through the organization of time and recognition. Values become more durable when they recur in collective practices.

However, repeated participation cannot be equated with moral conviction. Ritual may deepen commitment, become routine, or teach people how to display the expected emotion. Its effects depend partly on whether participants can interpret, question, and reshape it.

Comte's provision for legal optionality is therefore substantive. It prevents a simple description of his project as compulsory religious observance. But legal freedom to abstain does not remove every source of social pressure.

Education is the principal connection between the proposed doctrine and successive generations. Comte envisaged early formation under maternal guidance, followed by systematic education; the *Catechism* places the transition at fourteen.

The broader purpose of this sequence is to coordinate emotional development, intellectual understanding, and preparation for social responsibilities. Its success would mean that people encounter moral obligations through relationships and practices before receiving a systematic explanation of them.

This conception asks more of education than transmitting information. It concerns the formation of attention, taste, aspiration, and judgment.

There is nevertheless an essential distinction between educating someone to understand an obligation and training someone to accept an assigned place. A common curriculum can enable cooperation, but it can also suppress disagreement about what cooperation should serve.

The strongest version of moral education would make people capable of recognizing injustice within the very institutions that educated them. Comte's system is much less secure at this point, because the anticipated positive consensus leaves limited conceptual space for enduring disagreement among adequately informed and morally serious adults.

The separation of spiritual and temporal power is meant to place moral limits on practical authority. Comte's spiritual power educates, advises, and judges conduct; temporal power directs practical affairs. In *A General View*, he explicitly distinguishes consultative theoretical authority from practical command. The arrangement is intended to prevent wealth and government from becoming their own final moral judges. Comte, *A General View of Positivism*, concluding discussion.

Its attraction is readily understandable. Those who control resources should remain answerable to standards they cannot simply purchase or decree.

But this transfers part of the problem to the institution interpreting those standards. Independence from government does not establish intellectual reliability, moral impartiality, or accountability.

Three questions therefore become unavoidable:

- How are moral authorities selected?
- How can their judgments be challenged?
- What happens when several credible authorities disagree?

An institution may restrain economic power while becoming oppressive in its own domain. The relevant safeguard is consequently more demanding than separating two powers: each must also be open to criticism and correction.

This is where Comte's confidence in a unified doctrine creates difficulty. Disagreement can too readily appear as incomplete education, selfish resistance, or historical backwardness rather than a possible indication that the doctrine itself requires revision.

His economic proposals seek to change the moral meaning of industrial interdependence. Steiner explains that, for Comte, people already work for one another through the division of labor while frequently understanding their activity primarily through self-interest. Altruism should make this social contribution conscious. Accumulated capital likewise connects generations; its administration and transmission must be assessed socially. Steiner, *Comte, Altruism and the Critique of Political Economy*.

This is a substantial extension of moral reform into economic life. The issue becomes how participants understand the purposes of their roles.

An employer may regard productive assets as instruments of personal enrichment or as resources carrying responsibilities. A professional may understand competence as a private advantage or as something owed to those relying on the profession. The formal activity can look similar while its governing motivation differs.

Nevertheless, moral reinterpretation is not an adequate substitute for examining material power. A worker's protection cannot safely depend entirely on an employer developing the appropriate conscience.

There are therefore two distinct questions:

1. How can economic actors become more responsible?
2. What arrangements protect people when those actors remain irresponsible?

Comte contributes substantially to the first. His reliance on moral authority is less convincing as a complete answer to the second.

His preference for duties over rights makes the stakes especially clear. The *Catechism* argues that human beings enter a world already laden with obligations to predecessors, contemporaries, and successors. Comte accordingly seeks to replace the governing language of individual rights with reciprocal social duties. Comte, *Catechism*, discussion of public life.

The underlying challenge is powerful: individuals receive benefits they did not create and cannot fully repay. Moral life cannot be reduced to defending personal claims.

But the proposed replacement does not follow necessarily from that observation. Rights can identify what others owe a person; they need not express hostility to obligation or social membership.

A duty imposed on the powerful and a right held by the vulnerable differ in their practical allocation of initiative. With a right, the vulnerable person can formulate a claim without first persuading the stronger party to become benevolent.

The crucial assessment is therefore whether Comte's moral system gives everyone an effective position from which to contest mistreatment. Reciprocal language alone cannot establish reciprocity in practice.

Comte's project is especially revealing when compared with the deliberate creation of law. There are clear similarities: norms are formulated explicitly, interpreted by specialists, taught systematically, and supported by organized responses to misconduct.

Yet "positive" in Comte's philosophy should not be confused with the idea that a norm is valid merely because an authorized institution has enacted it. His proposed authority rests on an asserted knowledge of humanity and its development.

The differences can be expressed precisely:

Question	**A model of deliberate lawmaking**	**Comte's proposed moral reconstruction**
Why should the norm bind?	Its legal authority depends on recognized institutions and procedures.	Its authority is claimed through knowledge of human and social order.
How does it become effective?	Promulgation, administration, adjudication, and enforcement.	Education, emotional cultivation, guidance, and public judgment.
What must change?	Legally relevant conduct and institutions.	Conduct, motives, habits, and the interpretation of social life.
How is revision authorized?	Through identifiable procedures for amendment, repeal, and interpretation.	Through a doctrine presenting itself as the mature framework of social reconstruction.

This comparison is an analytical reconstruction, rather than Comte's own table.

It reveals a central asymmetry. Comte provides an extensive justification for superseding an obsolete moral authority, but a much less adequate account of how his own eventual moral authority could legitimately be superseded.

The problem of conscious moral change therefore has two levels: how to organize a transition, and how to preserve the legitimacy of future transitions. A system can be highly innovative at the first level while restrictive at the second.

John Stuart Mill's criticism identifies the danger of comprehensive moral supervision. Mill objected to the extent of Comte's proposed direction of individual life and defended room for diverse pursuits and personal enjoyment. He also distinguished the authority arising from

evidence and agreement from the authority of an organized body prescribing conduct. His criticism recognized that unrestricted discussion could coexist with oppressive social opinion. Mill, *Auguste Comte and Positivism*.

The question is not whether public judgment should ever influence behavior. Every society relies on approval, criticism, and informal expectations. The question is how comprehensive and concentrated that influence should become.

A person may be legally permitted to dissent yet face a coordinated judgment that dissent proves moral deficiency. If education, commemoration, professional recognition, and religious standing all operate through one doctrine, the formal absence of criminal punishment may leave little practical independence.

This criticism also clarifies a demanding feature of Comte's altruism. If every activity must justify itself as service to the whole, ordinary personal pursuits become permanently answerable to an encompassing moral tribunal.

A defensible morality needs to explain both what people owe one another and what they may legitimately decide for themselves.

The most serious philosophical weakness is the insufficient separation of empirical explanation, moral justification, and institutional design. These are three different undertakings:

- An empirical theory asks what strengthens cooperation or changes dispositions.
- A normative theory asks which forms of cooperation and which dispositions deserve support.
- An institutional theory asks who may act on these judgments and under what constraints.

Success in one undertaking does not guarantee success in the others.

Suppose research established that a particular educational practice produces greater obedience and social stability. That would not yet establish that obedience is appropriate, that the stabilized arrangements are just, or that imposing the practice is legitimate.

Conversely, demonstrating that a value is defensible does not prove that a proposed method will successfully cultivate it.

Comte's system is most persuasive when it reveals connections among these questions. It is least persuasive when it treats those connections as if they eliminated the need to answer each question separately.

His historical importance lies in making the social production of morality itself an explicit object of theory. Several durable insights can be extracted without accepting his complete program: convictions require institutions of transmission; motives require cultivation; economic activity has moral meanings; collective memory distributes admiration; and the decline of theology does not abolish the problems previously addressed by religious organization.

His influence also traveled through selective appropriations. Keen traces the dissemination of altruism through British positivist and literary circles, illustrating how one element could circulate much more widely than adherence to the entire religion. Keen, "'Altruism' Makes a Space for Empathy, 1852."

Such influence is evidence of intellectual productivity, not proof that Comte's complete institutional design would accomplish its promised moral transformation.

The enduring issue his work leaves us is therefore precise: **how can a society consciously cultivate better moral dispositions while keeping the authority that defines and cultivates them open to criticism?** Comte explains with unusual breadth why moral change needs organized social support. His system also demonstrates why that support requires limits, accountability, and legitimate opportunities for disagreement.

****

Émile Durkheim

Émile Durkheim's central contribution to the theory of conscious moral change is the idea that a society can increasingly understand, criticize, and deliberately reshape its own moral life—but only by working through the social relationships and institutions that give morality its authority. Moral reform therefore requires more than formulating better principles. It requires creating the conditions under which people can recognize those principles as binding, become attached to them, and incorporate them into everyday conduct.

This concern runs through his sociology of the division of labour, his account of moral education, his proposals for occupational organization, his political sociology, and his later analysis of religion. These writings do not constitute a single finished treatise on moral transformation. Nevertheless, they support a coherent reconstruction of its conditions, mechanisms, and limits.

Durkheim's position is particularly significant because it addresses three different questions together: **How do new moral ideals emerge? How can people consciously guide their development? How do those ideals become an effective, enduring morality?** Keeping these questions distinct helps explain both the strengths and the unresolved tensions of his theory.

His project developed in response to a historically specific problem: the moral reconstruction of modern France. Industrialization, occupational specialization, secularization, and republican political development were weakening older sources of authority while creating new expectations of individuality and justice. Durkheim's career joined sociology to education from the outset: he began teaching social science and pedagogy at Bordeaux in 1887 and moved to the Sorbonne in 1902. Investigating society and preparing educators belonged to the same practical undertaking. Jean-Claude Filloux emphasizes this connection between sociological knowledge, educational institutions, and the formation of a society respectful of the individual. Filloux, "Émile Durkheim."

The principal writings relevant to this undertaking are:

Work	Date	Contribution to understanding moral change
The Division of Labour in Society	1893; significant second preface in 1902	Changing solidarity, economic injustice, and the need for occupational moral institutions

Work	Date	Contribution to understanding moral change
The Rules of Sociological Method	1895	Social diagnosis, normality, pathology, and the possible contribution of deviance
Suicide	1897	Deficient integration and regulation as problems requiring institutional remedies
"Individualism and the Intellectuals"	1898	Individual dignity as a collective ideal capable of authorizing criticism
Moral Education	Principally the 1902–1903 course; published 1925	Secular moral formation, discipline, attachment, and autonomy
"The Determination of Moral Facts," with subsequent discussions	1906	Moral authority, moral criticism, and the role of scientific reflection
"Value Judgments and Judgments of Reality"	1911	The social formation of ideals and valuations
The Elementary Forms of Religious Life	1912	Collective experiences, symbols, and rituals as sources of moral creativity
Professional Ethics and Civic Morals	Earlier lectures; French publication 1950, English translation 1957	Occupational regulation, democratic reflection, rights, property, and contractual justice

The posthumous publication dates matter: the educational and political lectures belong to Durkheim's intellectual development before his death in 1917. They should not be read as later additions to his thought. Bibliographical overview of Durkheim's writings.

The starting point is that morality already exists as a social reality before a philosopher begins to reconstruct it. People encounter obligations, prohibitions, standards of honour, expectations of fairness, and judgments of blame within social life. These are embodied in practices and institutions as well as explicit statements.

Durkheim consequently treats moral rules as objects of empirical investigation. Their history, authority, social distribution, and relationship to institutions must be examined. A reformer who begins with a supposedly self-evident definition of the good risks overlooking the conditions that make any moral order effective.

His account also distinguishes two dimensions of moral experience. A moral requirement presents itself as an **obligation**, carrying authority beyond immediate inclination. At the same time, moral action involves **attachment to something valued**: people can experience devotion to collective purposes as fulfilling and desirable. Fear of punishment alone cannot explain this combination. Durkheim's account of moral obligation and desirability.

The implication for deliberate reform is substantial. Changing what people say they approve of is only one part of changing their morality. A successful reform must also alter what they feel

responsible for, which violations provoke disapproval, which achievements deserve esteem, and which collective purposes attract their commitment.

For example, imagine an institution adopting a principle of equal treatment. If its everyday routines continue to reward favouritism, its members learn contradictory lessons. The official principle may exist as a declaration while the operative morality continues to be organized around personal allegiance. A Durkheimian analysis would therefore examine the institution's relationships and practices alongside its stated values.

Modern differentiation creates a need for moral change because the conditions of solidarity themselves change. In *The Division of Labour*, Durkheim distinguishes solidarity grounded predominantly in similarity—mechanical solidarity—from solidarity associated with differentiated but interdependent activities—organic solidarity.

Where collective life rests heavily on similarity, morality can demand extensive conformity. As specialization develops, people increasingly occupy different roles and lead different lives. Their cooperation requires obligations appropriate to those differences. General moral agreement remains necessary, but it cannot consist in everyone sharing an identical occupation, outlook, or way of life.

The resulting crisis arises when older regulations lose effectiveness while suitable new rules remain underdeveloped. Durkheim's response is to organize the obligations appropriate to interdependence, including fair relations between specialized functions. He accordingly connects moral reconstruction to greater justice and to the correction of inequalities that obstruct solidarity. Robert Alun Jones's analysis of *The Division of Labour*.

This supplies a first criterion for evaluating reform: **does the proposed morality enable people to cooperate under their actual social conditions?**

The question has two sides. A surviving rule may obstruct relationships that a differentiated society needs. Conversely, removing an inherited restraint may leave a vital relationship without adequate regulation. Moral modernization therefore cannot be measured simply by counting how many prohibitions disappear or how many new obligations are announced.

His analysis of anomie identifies what can happen when change dissolves regulation faster than new regulation develops. In *Suicide*, Durkheim argues that disturbances associated with both economic crisis and sudden prosperity can disrupt established expectations. The problem is not reducible to material deprivation: people can lose a socially supported sense of what they may reasonably expect and pursue.

He also distinguishes inadequate regulation from inadequate integration. The former concerns the ordering of aspirations; the latter concerns attachment to groups and purposes. Their practical remedies cannot be identical. More rules do not necessarily supply belonging, while stronger belonging does not automatically establish fair or intelligible limits.

Durkheim therefore looks beyond exhortations directed at isolated individuals toward the organization of social life. These are historical sociological arguments, rather than a complete contemporary explanation of suicide. Jones's analysis of *Suicide*.

For the theory of conscious moral change, the enduring issue is the management of transition. Reform can undermine the credibility of old expectations without immediately establishing credible replacements. A programme that declares familiar duties obsolete consequently assumes a further responsibility: explaining how cooperation, mutual expectations, and personal orientation will subsequently be sustained.

Scientific reflection is supposed to make moral transformation more deliberate. Durkheim's ambition is to move beyond both unexamined acceptance of tradition and speculative moral construction. The investigation of moral phenomena should help identify their causes, the conditions under which they function, and the changes needed when those conditions alter.

François-André Isambert aptly describes the connection between Durkheim's science of morality and his project of a secular morality. Scientific understanding was intended to provide a capacity for acting upon moral phenomena, just as knowledge of natural processes increases the possibility of acting upon them. The sociologist and the moral reformer are therefore connected roles, although explanation and practical intervention remain distinguishable. Isambert, "Durkheim: une science de la morale pour une morale laïque" (1990).

Reconstructed as a practical method, this means asking:

- What social relationships originally supported a rule?
- Which of those relationships still exist?
- What effects does the rule now produce?
- Which emerging expectations does it fail to recognize?
- What institutions could support a more appropriate arrangement?

These questions are an analytical reconstruction of his approach, rather than a procedural checklist he formally established.

An especially important consequence follows: the reformer's conviction that a principle is admirable cannot substitute for investigation into its possible social realization. Yet evidence that a principle is difficult to realize does not settle whether it is worth pursuing. Durkheim's theory is strongest when it disciplines reforming ambition through social knowledge; it becomes more controversial when scientific diagnosis appears to determine the moral goal itself.

His distinction between the normal and the pathological is intended to support criticism, but it also creates a major difficulty. In *The Rules of Sociological Method*, normality is assessed relative to a social type and a stage of development. Mere occurrence within one society does not establish that a practice is appropriate to its current conditions.

Durkheim also allows deviance a role in moral development. His example of Socrates shows that conduct condemned under existing rules may anticipate a morality that is becoming possible. Some flexibility of collective judgment is necessary if moral innovation is to occur. This does not make every offence progressive; it denies that conformity is an adequate measure of moral worth. Jones's analysis of *The Rules of Sociological Method*.

The difficulty is deciding, prospectively, which departures express a developing moral possibility and which damage indispensable relationships. Historical success cannot supply a

sufficient answer: a movement may prevail through domination. Nor does the persistence of an institution prove its justice.

Durkheim thus opens a space for informed moral criticism without providing an uncontested method for adjudicating every conflict within that space.

His most explicit defence of moral reform distinguishes society from its prevailing opinion about itself. In the 1906 discussions accompanying "The Determination of Moral Facts," he argues that public moral opinion can contain survivals, misunderstandings, and inadequately articulated aspirations.

The moral critic may therefore appeal from established opinion to a better understanding of collective life. Durkheim presents Socrates as expressing needs and tendencies that existed beyond his individual consciousness, even though his judges failed to recognize them.

Scientific reflection, on this account, can correct and direct morality. Its task is to articulate a more adequate collective ideal. Durkheim nevertheless doubts the practical value of a proposed morality entirely disconnected from any existing social aspiration. Durkheim, "The Determination of Moral Facts" and replies to objections.

This position is best understood as a theory of **criticism from within historically developing social life**. The critic can expose tensions between a society's institutions and its own emerging commitments.

Consider a society that recognizes equal civic standing while tolerating inherited exclusion from important opportunities. A reformer can make the contradiction explicit and organize support for institutional change. The new demand has a social history, yet its implications exceed what the majority currently accepts.

There remains a danger: someone may claim to know society's "true" needs better than its members. Durkheim's argument establishes the possibility of justified dissent, but the language of deeper collective reality can also become a resource for paternalism.

Moral individualism supplies the substantive direction of much of Durkheim's reforming thought. His 1898 essay "Individualism and the Intellectuals," written during the Dreyfus affair, distinguishes devotion to human dignity from the pursuit of private advantage.

The human person has acquired a standing that Durkheim describes through the language of sacredness. Respect for the individual is itself a collective moral commitment. An injustice against a person can consequently offend society's fundamental morality, even when powerful institutions defend that injustice.

The essay also insists that inherited declarations of liberty require further development. Freedom must be made usable, capacities must be cultivated, and contractual relationships must become more just. Intellectuals possess no exclusive political jurisdiction merely because they are scholars; ordinary practical moral judgment can recognize the need for a fair hearing. Durkheim, "Individualism and the Intellectuals."

This gives his theory a distinctive structure. Collective authority can support individual independence, and criticism of an institution can defend collective morality.

Hans-Peter Müller accordingly interprets Durkheim's project as a moral politics of individualism: the central problem is how freedom and personal dignity acquire effective social, institutional, and cultural support. Müller, "Emile Durkheims Moralpolitik des Individualismus" (2009).

The reforming implications extend beyond toleration. If individuality is valuable, institutions must enable people to develop and exercise it. Education, occupational relations, and political arrangements become matters of moral concern because they can either support or obstruct that development.

Secular education is Durkheim's most sustained practical programme for consciously changing morality. The opening lecture of *Moral Education* argues that removing theological language cannot by itself produce an adequate secular education. Religious forms had carried real moral authority and motivation; their replacement requires understanding those functions.

He also explicitly calls for substantive innovation. Increasing sensitivity to individual dignity makes previously accepted relationships appear unjust. Educators must help younger generations understand emerging ideals and acquire the capacity to develop them. The inherited moral code must remain open to further achievement.

At the same time, he describes pedagogy as practical theory: it draws on available knowledge to guide action under conditions that cannot wait for a completed science. It requires informed judgment rather than mechanically applicable recipes. Durkheim, *Moral Education*, Lecture 1.

Two dimensions of change are joined here. One concerns the **grounds of obligation**: why a rule deserves recognition. The other concerns the **content of obligation**: what people owe one another. A secular reform that changes only the vocabulary may leave the second dimension untouched.

Durkheim's educational proposal can therefore be evaluated by asking whether it equips pupils merely to repeat socially approved judgments or also to recognize inadequacies in those judgments.

The three elements of moral education are discipline, attachment to social groups, and autonomy. Discipline develops regularity, self-command, and respect for rules. Attachment gives conduct purposes beyond immediate private inclination. Autonomy adds understanding of the rules and conditions under which one acts.

In Lecture 7, Durkheim gives particular weight to the active and inventive conscience during periods of transition. When new morality is still developing, stronger insistence on inherited discipline is insufficient. In Lecture 8, knowledge of morality's social conditions can support informed acceptance and the correction of abnormal elements.

His autonomy is consequently a historically acquired capacity for understanding and directing socially formed conduct. It does not mean that each person can create an entire moral order according to preference. Durkheim, *Moral Education*, especially Lectures 2–8.

Each element addresses a different problem in reform. Discipline makes sustained conduct possible. Attachment supplies motivation to pursue collective ends. Autonomy allows inherited conduct to become an object of examination.

Their relationship is difficult. A pupil may learn self-control while also learning excessive deference. Attachment may encourage generosity, but it may also produce hostility toward outsiders. Understanding why a rule exists can illuminate its purpose without proving its legitimacy.

Domingo Fernández Agis identifies the balance between individual autonomy and obligatory collective rules as a central problem in Durkheim's educational thought. Fernández Agis, "La educación moral" (2011).

A defensible development of the programme would therefore require education to make criticism practicable. Pupils would need opportunities to distinguish a rule's purpose from the interests of the person enforcing it, and to examine whether its application remains consistent with its justification. This is a reconstruction of what the autonomy requirement demands, rather than a claim that Durkheim supplied a complete contemporary pedagogy of classroom deliberation.

Schools matter because they organize experiences through which moral dispositions develop. In *Education and Sociology*, Durkheim treats education as the formation of the social being. A society transmits common expectations while also preparing people for differentiated social environments. Educational aims emerge historically from collective life and its organization. Durkheim, *Education and Sociology*, selected text.

The practical implication is that a school's conduct communicates moral lessons alongside its curriculum. How it distributes attention, applies rules, recognizes achievement, handles disagreement, and responds to mistakes becomes part of moral education.

An institution teaching impartiality through arbitrary decisions weakens its own lesson. One teaching responsibility while withholding every opportunity to exercise judgment creates a different inconsistency. These examples show why deliberate moral change requires attention to the organization of experience.

Jeffrey Dill's comparison with Dewey helps identify a remaining tension: Durkheim gives substantial weight to an authoritative moral order confronting the child, while Dewey gives greater prominence to inquiry and social experience in moral development. Dill also asks how such educational projects can operate under conditions of cultural pluralism. Dill, "Durkheim and Dewey and the Challenge of Contemporary Moral Education" (2007).

Occupational organization supplies the corresponding programme for adult life. Durkheim's 1902 preface to *The Division of Labour* makes a decisive clarification: habitual adjustment between specialized activities does not fully explain how binding rules emerge. A constituted group must invest rules with authority, and it can actively formulate them and arbitrate conflicting interests.

He therefore advocates occupational bodies capable of developing obligations appropriate to economic practices. These "corporations" are organized occupational communities, rather than companies in the ordinary contemporary sense. Their significance lies in combining practical knowledge, continuity, collective authority, and regular relationships.

He also emphasizes the moral consequences of everyday economic conduct. An economic sphere organized around insufficiently restrained competition cannot be isolated from the rest of moral life. Durkheim, preface to the second edition of *The Division of Labour* (1902).

This is one of his most useful insights into deliberate reform: **a society needs institutions capable of producing and maintaining specific obligations.**

General appeals to fairness leave many practical questions unsettled. Who owes information to whom? What constitutes an abuse of professional discretion? How should competing responsibilities be ordered? Which obligations continue when fulfilling them becomes costly?

A functioning moral institution can translate general commitments into determinate expectations, discuss difficult cases, and revise standards in light of experience. Its effectiveness depends on participants treating it as more than an external enforcement mechanism.

The proposal also has obvious vulnerabilities. Occupational solidarity can protect insiders, conceal misconduct, or exclude competitors. Practical expertise can be used to justify privileges that outsiders cannot effectively challenge. Creating a norm-producing institution consequently raises a further question: how will its authority remain answerable to the wider community?

Durkheim's political sociology addresses this through the relationship between individuals, intermediate groups, and the state. The state can protect individuals from oppressive family, religious, local, or occupational powers. Intermediate groups, in turn, are necessary to prevent individuals from confronting an excessively dominant state in isolation.

Mark Cladis develops this relationship as central to Durkheim's account of pluralism and solidarity. Collective institutions can sustain individual freedom when their powers constrain one another and when occupational life remains accountable to a wider political community. Cladis, "Beyond Solidarity? Durkheim and Twenty-first Century Democracy in a Global Age."

The institutional lesson is that no single source of moral authority should absorb the whole person. An occupational body may be competent to judge professional responsibilities without being competent to prescribe every aspect of its members' lives. Equally, the state's responsibility for common conditions does not establish its competence to direct every relationship.

This architecture provides a way of thinking about moral reform as distributed work. Different institutions contribute distinct knowledge and forms of authority. Their interaction can expose defects that each institution would otherwise leave unquestioned.

Democracy contributes by making collective arrangements available to conscious examination. In the political lectures, Durkheim connects democracy with the widening scope of reflective collective consciousness and closer communication between governing institutions and the population. Matters governed by inherited routine become subjects of discussion, criticism, and deliberate modification.

Deliberative institutions thereby increase society's capacity to anticipate and direct change. Government, however, must do more than reproduce immediate opinion: reflection requires

some distance from its fluctuations. Durkheim, *Professional Ethics and Civic Morals*, especially Lectures VII–IX.

This is a direct contribution to the theory of conscious moral change. The transformation concerns not only particular rules but also **the way rules are formed**. Practices that once persisted through unexamined inheritance can become revisable through continuing public attention.

Yet reflection depends on motivation as well as argument. Atle Møen's comparison with Habermas emphasizes Durkheim's view that public communication draws strength from collective sentiments and ceremonies. People need commitments that make participation in public reasoning matter to them. Møen, "Democracy and Public Communication: A Durkheimian Lens on Habermas."

The unresolved problem is how to preserve that collective commitment while leaving substantive disagreement open. An institution that identifies criticism with disloyalty destroys part of the reflective capacity that democracy is supposed to develop.

Justice in economic relations is a substantive test of the proposed moral reconstruction. In his account of the forced division of labour, Durkheim argues that solidarity is impaired when inherited position or external inequalities prevent people from occupying appropriate functions.

His notion of spontaneity therefore presupposes conditions under which participation is not distorted by entrenched advantage. Formally available opportunities are insufficient if inherited constraints continue to determine access. The correction of these conditions belongs to the moral organization of the division of labour. Durkheim, *The Division of Labour*, Book III, Chapter II.

The relevant analytical distinction is between a rule's formal symmetry and its operation within unequal relationships. Two parties may possess the same formal permission to accept or refuse an agreement while facing radically different consequences from refusal.

Moral reform must consequently examine the conditions under which agreement occurs. Otherwise, it can present dependence as choice and inherited advantage as achievement.

Durkheim's criticism of inheritance forms part of this argument. Alvin Gouldner's introduction to *Socialism and Saint-Simon* draws attention to the relationship between inherited wealth, unequal contractual power, and Durkheim's proposal that occupational institutions could assume economic functions previously performed by the family. Gouldner's introduction to Durkheim's *Socialism and Saint-Simon*.

Durkheim also leaves the moral horizon open. At the end of the political lectures, he considers how sympathy can extend beyond proportional reward for merit, since natural talents are themselves unchosen. Duties previously understood as charity can become stricter obligations and acquire institutional expression. Durkheim, *Professional Ethics and Civic Morals*, Lecture XVIII.

This suggests a particularly important form of moral change: an action can move from being admired generosity to being an expected duty. Conscious reform then involves recognizing that altered expectation, specifying its limits, and creating institutions capable of fulfilling it.

His later theory of religion explains why moral innovation cannot be reduced to rational planning. In *The Elementary Forms*, intensified collective experience can generate new ideals and invest symbols with exceptional significance. Rituals and commemorations subsequently renew the commitments associated with those experiences.

Durkheim treats a society's ideals as part of its reality. Conflict can occur between inherited ideals and emerging ones, rather than simply between reality and an ideal external to it. He also explicitly criticizes Comte's attempt to revive historical materials artificially into a living cult. Future forms of moral enthusiasm cannot be fully specified in advance by an intellectual designer. Their emergence depends on collective life. Durkheim, *The Elementary Forms of Religious Life*, Conclusion.

His essay on value judgments adds that valuations cannot be explained solely by the intrinsic properties or practical usefulness of objects. Collective ideals contribute to the significance objects and actions acquire. Durkheim, "Value Judgments and Judgments of Reality."

The practical consequence is a limit on moral administration. Officials can organize occasions for participation, educators can clarify emerging aspirations, and associations can develop standards. None can guarantee that a prescribed ideal will acquire living authority.

The distinction between generating commitment and sustaining it is also useful. A powerful collective experience may produce enthusiasm without durable practices. Conversely, an elaborate institutional framework may preserve forms whose motivational force has disappeared. Lasting reform requires a relationship between these two dimensions.

The different strands can be combined into a model of deliberate moral transformation. The following is a reconstruction across Durkheim's works, rather than a sequence he presented in this exact form:

Moment in moral transformation	Contribution of conscious action	Limit confronting reformers
Existing rules become inadequate	Identify the relationships and expectations that have changed	Discontent alone does not determine an appropriate replacement
New aspirations appear	Clarify experiences of injustice and articulate possible ideals	Intellectual formulation cannot ensure collective commitment
Ideals gain public significance	Connect reflection with participation and shared experience	Enthusiasm may be unstable or exclusionary
Duties become more definite	Establish bodies capable of discussing and authorizing standards	Organized groups may defend their own interests
Institutions and practices change	Alter the conditions under which people act and cooperate	Formal change may leave established dispositions intact

Moment in moral transformation	Contribution of conscious action	Limit confronting reformers
New dispositions develop	Educate through practice, attachment, and understanding	Socialization can become conformity
The new morality is examined again	Maintain criticism and institutional revision	Reformers may treat their own settlement as final

This model highlights the reciprocal relationship between morality and institutions. Institutions help form dispositions, but morally motivated people also criticize and alter institutions. Neither side can be treated as a passive reflection of the other.

It also explains why reform can proceed unevenly. Public language may change before everyday expectations; organizational procedures may change before emotional attachments; younger generations may interpret principles differently from those who introduced them.

Durkheim's theory illuminates the comparison between conscious moral change and conscious lawmaking, while also exposing its limits. Both involve making inherited norms explicit, subjecting them to criticism, and establishing organized capacities for revision.

Morality, however, does not ordinarily possess a single authoritative procedure through which an obligation can be made socially ineffective at a specified moment. A legal enactment can have a determinate formal status. Moral expectations persist through relationships, conscience, esteem, disapproval, and institutional habits. Their transformation has no equivalent general mechanism of promulgation and repeal. A Durkheimian account therefore directs attention to the social processes surrounding codification. A professional code might express an already developing commitment, help crystallize a partially formed one, or remain largely ceremonial. Which of these occurs depends on the community that receives and practices it.

Likewise, abandoning an official standard need not eliminate the expectations attached to it. People may continue to feel bound by an inherited duty after institutions cease to require it. Others may recognize a new duty before any authoritative body acknowledges it.

Steven Lukes and Devyani Prabhat's analysis is useful here because it challenges the interpretation that Durkheim simply favours legal enforcement of whatever morality currently prevails. They emphasize the role of individual dignity in distinguishing criticism that develops modern morality from conduct that damages its central commitments. Lukes and Prabhat, "Durkheim on Law and Morality: The Disintegration Thesis."

The resulting distinction is between **deliberately organizing moral formation** and **assuming that moral authority follows automatically from an authoritative decision**. Durkheim offers extensive resources for the former. His account gives strong reasons to doubt the latter.

The deepest philosophical objection concerns the passage from explaining morality to justifying it. Showing that a norm sustains social relations does not, by itself, establish that those relations deserve preservation. Showing that people experience an obligation as authoritative does not settle whether they ought to do so.

Frederick Neuhouser formulates this difficulty precisely: Durkheim does not adequately explain how historically specific moral systems possess an authority irreducible to their

contribution to social reproduction. His scientific account can identify problems within an order while leaving the ultimate authority of its fundamental commitments insufficiently justified. Neuhouser, "Durkheim: A Science of Morality" (2022). Durkheim's commitment to human dignity provides a substantial normative resource. Nevertheless, its relationship to his sociological method remains contested. Is dignity justified because it has become a central collective ideal, because it supports differentiated social life, or because persons deserve respect independently of those facts? These are different arguments.

Several further limitations follow.

First, **cohesion and justice can diverge**. A group can sustain strong solidarity through exclusion. The capacity to create moral commitment is therefore not sufficient evidence that the commitment is defensible.

Second, **claims to expertise require institutional checks**. A sociologist, educator, or administrator may identify a real discrepancy between prevailing opinion and social conditions. That possibility does not confer unlimited authority to decide which way society should develop.

Third, **moral pluralism may be enduring**. Some disagreements concern competing interpretations of common ideals; others concern which ideals should govern a relationship. Occupational mediation and democratic discussion cannot guarantee a final reconciliation.

Fourth, **collective emotional intensity has no automatic progressive direction**. The processes that strengthen devotion to dignity and solidarity can also strengthen hostility and domination. A theory of the production of moral authority must remain distinct from an assessment of its content.

Fifth, **the endpoint of economic justice is uncertain**. Roger Cotterrell's survey shows how differently Durkheim has been interpreted: as seeking fairer operation of differentiated relations, as requiring substantial reconstruction of those relations, and as negotiating tensions between equality of opportunity and equality of treatment. These disagreements reflect real ambiguities in the reform programme. Cotterrell, "Durkheim on Justice, Morals and Politics" (2010).

For further study specifically focused on your question, Hans Joas's essay **"Durkheim's Intellectual Development: The Problem of the Emergence of New Morality and New Institutions as a Leitmotif in Durkheim's Oeuvre"** is an especially relevant starting point. Its very subject brings together the two problems that must remain connected: the emergence of new moral commitments and the construction of institutions capable of sustaining them. It appeared in Stephen P. Turner's edited volume *Émile Durkheim: Sociologist and Moralist* (1993), pp. 229–245. Verified bibliographical entry. The question Durkheim leaves for any conscious moral reformer is consequently demanding: **what makes a proposed principle worthy of commitment, through which social relationships can it become effective, and how will those relationships preserve the ability to criticize the principle once it has acquired authority?**

****

Lucien Lévy Bruhl

Lucien Lévy-Bruhl's central contribution to the theory of conscious moral change is his proposal for a rational art of moral reform, grounded in the scientific study of existing moral life. His argument is that societies can become more capable of deliberately transforming their moral rules if they discover how those rules arise, acquire authority, persist, and interact with other institutions. Moral reform consequently requires knowledge of social reality as well as moral conviction.

He developed this programme principally in *La morale et la science des mœurs* (1903), translated by Elizabeth Lee as *Ethics and Moral Science* (1905). His later anthropological studies illuminate some of its underlying problems, particularly the relations between collective beliefs, emotions, and reasoning. They should, however, be distinguished from the explicit programme of moral reform presented in the earlier book. Lévy-Bruhl, *Ethics and Moral Science*.

The distinctive feature of his position is the combination of considerable ambition with methodological restraint. He envisages humanity acquiring some deliberate control over its moral development. At the same time, he rejects the assumption that a philosopher can first design a complete, universally valid morality and then introduce it into society through persuasion or legislation. Existing morality is too deeply embedded in collective life for that procedure to work as its proponents imagine.

To understand both the importance and the difficulties of this position, we must examine his criticism of moral philosophy, his conception of moral reality, the mechanisms of change he identifies, and the unresolved question of the standards by which deliberate changes should be judged.

The starting point is a distinction between existing morality, knowledge of morality, and intervention in morality. Lévy-Bruhl believes that discussions become confused because the same vocabulary is used for all three.

Dimension	What it comprises	Its relation to moral change
Existing moral reality	Obligations, prohibitions, ideals, judgments, feelings, practices, and sanctions operative in a society	The reality to be understood and potentially modified

Dimension	What it comprises	Its relation to moral change
The science of morals—*science des mœurs*	Historical, comparative, and causal investigation of that reality	Knowledge of how morality works and under what conditions it changes
The rational moral art—*art moral rationnel*	Practical intervention informed by that knowledge	Deliberate modification of particular practices, institutions, and obligations

Here, *art* means a practical capacity or technique, comparable to medicine. And *mœurs* encompasses socially established ways of living and judging, extending beyond customs understood as mere outward habits.

The distinction immediately changes the meaning of moral expertise. A person may feel duties intensely, comply with them conscientiously, and explain their conventional justification without knowing their historical origins or the conditions that sustain them. Familiarity with morality does not amount to scientific knowledge of it.

Conversely, a researcher may understand how a prohibition developed without thereby either endorsing or rejecting it. The existence of an obligation as a social fact and the justification of that obligation are different questions. Much of Lévy-Bruhl's argument concerns the need to keep these questions distinguishable while establishing a more fruitful connection between knowledge and action.

His first major target is what he calls **theoretical ethics**: the attempt to combine an explanatory science with a system that prescribes what everyone ought to do. The provocative title of his first chapter announces that theoretical ethics, in this particular sense, neither exists nor can exist.

His reasoning is that a science investigates a given reality and discovers relations within it. Traditional moral philosophy, by contrast, frequently begins with a conception of the supreme good, duty, happiness, perfection, or human nature and uses it to formulate prescriptions. It presents those prescriptions as the practical consequences of its theoretical discoveries. Lévy-Bruhl questions whether this is a genuine account of how such systems are produced.

He observes that philosophical schools often disagree fundamentally about the foundations of morality while recommending remarkably similar everyday conduct. A religious thinker, a rationalist, and an empiricist may offer incompatible accounts of moral obligation yet converge on the accepted duties of their society.

His explanation is sociological: their practical conclusions are already substantially supplied by the surrounding moral culture. Philosophical argument subsequently organizes and rationalizes them. The professed sequence runs from principle to obligation; the actual historical sequence often runs from established obligation to philosophical justification.

This criticism has a direct bearing on deliberate reform. If the philosopher unknowingly incorporates prevailing morality into the premises of a supposedly universal system, that system may reinforce inherited norms while appearing to subject them to independent examination. A society's particular arrangements concerning family, property, authority, or social rank can acquire the appearance of requirements of reason itself.

Lévy-Bruhl nevertheless recognizes that philosophical reflection has had real effects. Bringing scattered obligations into a more coherent formulation can expose tensions, generalize principles, and give morality some independence from religious or legal institutions. His argument is therefore strongest when understood as a criticism of philosophy's exaggerated account of its independence from social life. It becomes less convincing if interpreted as a claim that moral argument never produces innovation. These qualifications are already relevant to his own historical discussion. *Ethics and Moral Science*, chapters I–III and IX.

A second target is **the assumption that an adequate account of human nature can be obtained by reflecting on the individuals immediately around us**. A philosopher studies familiar motives, sentiments, and relationships, abstracts from them, and then speaks about humanity in general.

For Lévy-Bruhl, historical and anthropological diversity makes this procedure unreliable. People's experiences of obligation depend partly on forms of kinship, religion, economic organization, political membership, and language that vary considerably. Even apparently elementary notions—what counts as a person's property, whose interests command concern, or what someone owes to relatives—cannot simply be detached from those contexts.

He is not obliged by this argument to deny every common feature of human beings. His methodological point is that common features must be established through inquiry. They cannot be assumed in a form that makes comparative investigation unnecessary.

A third target is **the supposed unity of moral consciousness**. Philosophical systems frequently treat the conscience as though it expressed an underlying harmonious order that a sufficiently penetrating analysis could reconstruct.

Lévy-Bruhl instead interprets moral consciousness as a historical accumulation. Obligations inherited from different periods survive alongside one another. Religious duties, economic expectations, family loyalties, political commitments, and newer ideals may have different origins and may not fit together.

Conflicts of duty can therefore reveal real tensions in the social formation of morality. They need not be mere errors in an individual's reasoning. Someone may experience two obligations as equally compelling because both have acquired authority through different, still active parts of collective life.

This is one of his most productive insights for a theory of reform: **moral change can become necessary because inherited obligations cease to fit either one another or the changing organization of society.** The task is then more substantial than persuading individuals to obey their existing duties more consistently.

The corresponding positive proposal is to investigate morality as a **social reality with discoverable conditions of existence**. Lévy-Bruhl sometimes calls this a moral nature, drawing an analogy with physical nature.

The analogy does not mean that obligations are material objects or that consciousness is irrelevant. It means that moral practices have a reality that exceeds individual preference. They precede particular individuals, become embodied in institutions and expectations, and confront

resistance with consequences. A person cannot simply decide that the surrounding system of duties has ceased to exist.

It also means that these practices have causes and relationships that ordinary moral awareness does not reveal. Feeling that a duty is binding tells us something about our experience. It does not by itself explain why that particular duty acquired its form, why it is stronger in one group than another, or why its force changes over time.

The appropriate inquiry therefore combines several kinds of investigation:

- Historical reconstruction of how obligations and institutions developed.
- Comparison between societies and between periods within a society.
- Examination of the relations among morality, law, religion, economic life, and political organization.
- Study of actual practices and sanctions alongside publicly professed principles.
- Investigation of the collective representations and feelings through which obligations become personally compelling.

These inquiries must inform one another. A purely psychological account risks treating socially produced experiences as timeless features of the individual. An account confined to formal institutions risks overlooking the emotions and expectations through which institutions operate.

The result is a conception of moral reform in which **knowledge of interdependence is essential**. Changing a legal arrangement may alter family expectations; changing economic opportunities may affect judgments about responsibility; changing religious beliefs may gradually weaken certain prohibitions. The practical consequences of an intervention cannot reliably be inferred from its immediate target alone.

Lévy-Bruhl's discussion of **moral feeling** makes this argument considerably more subtle. He does not conceive morality as a collection of propositions that people accept after examining their reasons. Obligations are bound up with respect, shame, remorse, reverence, indignation, and attachments to others.

Moreover, beliefs, practices, and feelings do not necessarily change at the same speed. A person may abandon the explicit belief formerly used to justify an obligation while continuing to experience the obligation's emotional force. Conversely, someone may sincerely accept a new principle without yet feeling or acting in accordance with it.

This produces a discrepancy between an individual's intellectual self-description and the forces actually governing conduct. People can think themselves emancipated from an old outlook while continuing to respond through dispositions formed within it.

For conscious reform, the consequence is substantial. Refuting a belief may be insufficient to remove the moral reaction associated with it. An objection can be intellectually accepted while disgust, shame, loyalty, or reverence persists.

The reverse is equally significant. A new idea may need to become connected with collective feelings before it acquires practical authority. Its formulation by a philosopher is only one moment in its possible social career.

Lévy-Bruhl consequently emphasizes the exceptional resistance of collective moral feelings to direct alteration. Their transformation often follows changes in other areas of life. This suggests a predominantly indirect route to moral change: alter the institutions, relationships, and experiences through which moral dispositions are reproduced, and the dispositions themselves may gradually change. That is a reconstruction of the practical implications of his analysis; he does not offer a complete set of procedures for producing these effects. *Ethics and Moral Science*, chapter VIII, especially pp. 179–202.

His account also explains why moral criticism provokes unusually strong resistance. A proposal to alter a particular rule may be experienced as an attack on the integrity of the whole moral order. The collective conscience tends to protect inherited obligations together, even when their histories and present functions differ considerably.

From the reformer's perspective, the issue may concern one obsolete prohibition. From the defender's perspective, accepting its revision may seem to weaken the authority of morality altogether. This helps explain why presenting additional arguments does not always resolve the conflict: the participants are responding to different perceived stakes.

The rational moral art is intended to make intervention more effective under these conditions. Its governing analogy is **medicine informed by scientific knowledge**.

A physician does not need to invent the organism being treated. The organism already exists, has its own structure, and responds to intervention according to conditions the physician must investigate. Likewise, moral reform operates on an existing social world whose organization cannot be disregarded.

The analogy introduces several restrictions. Knowledge is incomplete; treatment must be appropriate to the case; interventions can have adverse consequences; some desired results may be unattainable; and an effective remedy in one context may fail in another.

It also changes the explanation of failure. Traditional moral exhortation can attribute persistent wrongdoing to deficient will: people know their duties but refuse to perform them. Lévy-Bruhl asks whether the failure also exposes ignorance about the causes of conduct or inadequacies in the duties and institutions themselves.

His crucial distinction is between improving conformity to an existing morality and improving that morality. Moral education commonly pursues the first objective. A rational moral art must also be capable of pursuing the second.

For example, more effective techniques for teaching obedience could strengthen a harmful obligation. Success in producing conformity would then coexist with failure to improve moral life. A science of morals must make the content and social operation of the obligation available for examination.

Lévy-Bruhl does not claim that this rational art already exists in mature form. He presents it as a programme whose development depends on the progress of several sciences. Its initial achievements would be partial and uneven. It would address determinate problems before becoming capable of coordinating more extensive transformations. *Ethics and Moral Science*, chapter IX, pp. 203–211.

His concrete examples show what he has in mind.

Penal institutions provide a particularly clear case. Existing punishment combines inherited beliefs about responsibility, retribution, intimidation, correction, and social protection. These components may coexist without forming a consistent system.

Lévy-Bruhl proposes investigating the physiological, psychological, and social conditions associated with different offences. Such knowledge could support more effective preventive and punitive arrangements. The desired advance would concern the actual organization of responses to wrongdoing, extending beyond renewed abstract debate about the essence of punishment.

A contemporary reader should distinguish the two claims involved. Empirical knowledge can help determine whether a measure reduces offending or produces harmful consequences. Whether that measure is proportionate, respects persons, or distributes burdens justly remains a further question. His example therefore displays both the promise of his programme and its normative difficulty.

Prostitution and its regulation supply another example. Lévy-Bruhl observes that debates rapidly connect public health, individual liberty, sexual morality, economic circumstances, and the wider organization of social relations. The common conscience may condemn a practice in principle while society accommodates it in fact.

His point is that an intervention cannot be assessed adequately by isolating one moral injunction. The conditions producing the practice, the interests maintaining it, and the repercussions of regulation must be investigated together. The example illustrates why moral, legal, and economic reform frequently overlap.

Education and social assistance appear as areas where initial advances are already conceivable. Knowledge of children's development can correct inherited educational practices. Investigation of poverty and unemployment can improve assistance by revealing causes that moral judgments about individuals may overlook.

These examples share a common structure:

Conventional formulation	Question introduced by Lévy-Bruhl's approach
People should obey the moral rule	What sustains or obstructs compliance, and what effects does the rule produce?
Wrongdoers deserve punishment	What produces different forms of wrongdoing, and how do particular responses affect them?
The unfortunate should receive charity	What generates their circumstances, and which arrangements effectively address them?
Children should acquire proper conduct	How do dispositions actually develop, and what do educational practices accomplish?
An inherited institution deserves respect	What role does it now perform, and how has that role changed?

The examples and their interconnections are developed especially in chapters IV and IX. *Ethics and Moral Science*, particularly pp. 217–220.

One might suppose that this dependence on developing knowledge condemns society to wait indefinitely before acting. Lévy-Bruhl explicitly resists that conclusion.

Life presents urgent decisions before adequate science is available. People must therefore act using experience, incomplete knowledge, and the most reasonable judgments available to them. The appropriate response to uncertainty is to acknowledge the provisional character of a decision and improve its grounds where possible.

His medical analogy applies here too: the incompleteness of medicine does not imply that every patient must go untreated. Equally, the incompleteness of social science does not require moral or political abstention.

This aspect of his position matters because his caution can otherwise be mistaken for blanket conservatism. He certainly opposes unfounded confidence in comprehensive schemes of reconstruction. But he also considers existing morality capable of preserving obsolete institutions and obstructing improvement.

His discussion of **social hypocrisy** is especially relevant to the conscious revision of norms. He distinguishes a duty that people recognize but violate from a supposed duty that they continue publicly to acknowledge after it has ceased to command their conscience.

In the first case, misconduct occurs within a still effective normative order. In the second, the professed order and the operative order have diverged. The same outward failure to comply can therefore have very different meanings.

This distinction challenges the familiar diagnosis that every departure from inherited rules represents moral weakness. Sometimes it indicates that the rules themselves have lost their social foundation.

Lévy-Bruhl argues that continuing to teach such rules as universally binding can corrupt moral education. Children learn that the official language of duty differs from the conduct and commitments of the adults teaching it. A practice intended to cultivate sincerity then trains people in insincere professions.

Scientific inquiry could help distinguish obligations that remain active from those maintained primarily through convention and formula. Reform could thereby make public moral language more truthful about the society's actual condition.

Nevertheless, identifying such a discrepancy does not by itself determine how to resolve it. Society might abandon the professed rule, renew commitment to it, or revise both the rule and the relevant institutions. That further choice brings us to the deepest difficulty in Lévy-Bruhl's programme. *Ethics and Moral Science*, chapter IX, pp. 220–226.

How can a science of existing morality determine what counts as moral improvement?

Lévy-Bruhl's answer is more complex than the proposition that whatever exists is right. He expressly allows criticism of existing obligations. Nor does he simply prescribe conformity to the majority's present convictions.

His approach begins from broad practical interests, such as the desire of individuals and societies to live and to live better. It then asks what more precise improvements become possible as knowledge develops. An institution may be shown to obstruct purposes that its members already acknowledge, to retain harmful features, or to rest on mistaken beliefs.

He also argues that knowledge can transform the ends people conceive. Science does more than supply means to goals already completely specified by philosophy. By revealing previously unknown possibilities, it can suggest new practical objectives. In the third-edition preface, he uses technological development to illustrate how expanding capacities can make previously unimagined ends available. Lévy-Bruhl, third-edition preface.

This is a significant contribution to the theory of moral innovation. Ends and means need not stand in a rigid sequence in which the former are settled once and for all. Discovering that a form of suffering is preventable, or that an apparently indispensable institution can be replaced, may change what people consider acceptable.

Consider an illustrative application of his reasoning: if a social arrangement appears necessary for collective survival, criticism of its harms may seem impractical. Once an alternative becomes demonstrably feasible, the moral defence of those harms changes. The empirical discovery does not independently create the value of avoiding suffering, but it changes the implications of that value.

Lévy-Bruhl's position can therefore support **criticism from within a society's existing commitments**, informed by a more adequate understanding of causes and possibilities. This is a stronger interpretation than the claim that sociology mechanically generates moral commands.

It nevertheless leaves unresolved cases in which the commitments themselves conflict. Greater security may require restrictions on freedom. A more cohesive community may impose burdens on dissenters. A measure that increases aggregate welfare may disadvantage a minority. Establishing the consequences does not automatically establish their proper ranking.

The medical analogy partly conceals this problem. Even medicine encounters disagreement about acceptable risks, quality of life, and consent. In the moral and political sphere, disagreement over the ends of intervention is still more pervasive.

Lévy-Bruhl's programme thus supplies a powerful argument that reform requires empirical knowledge. It does not conclusively establish that empirical knowledge makes independent normative reasoning unnecessary.

This difficulty was recognized by contemporaries. In his 1905 criticism, **Alfred Fouillée** challenged the reduction of morality to sociological explanation and defended the significance of individual thought, valuation, and moral invention. He argued that explaining the conditions under which ideas arise does not exhaust their meaning or eliminate their independent contribution to historical change. His criticism also presses the distinction between the existence of social pressure and the moral authority of what it demands. Fouillée, "La Science des mœurs remplacera-t-elle la morale?".

For conscious moral change, this debate exposes two questions that must remain distinguishable: how can people effectively transform a moral order, and why should they

choose one transformation over another? Lévy-Bruhl greatly strengthens the first inquiry while offering a contested reconstruction of the second.

His position also raises **the question of who is to direct reform**. The importance he assigns to specialized knowledge might suggest government by scientific experts. Yet his book does not provide a developed institutional scheme establishing sociologists as authorized moral legislators.

He does not specify a complete procedure for choosing reformers, settling disagreement among experts, representing affected groups, or reviewing the use of scientific authority. His rational moral art is primarily a methodological programme. Its political constitution remains underdeveloped.

This gap becomes especially important when his theory is considered in relation to the deliberate creation of norms. Knowing how an institution functions does not by itself confer authority to alter it. Scientific competence, moral judgment, and legitimate decision-making are distinct capacities, even when reform requires all three.

The same distinction clarifies the relationship between his programme and **the positivization of morality**.

“Positive” in his conception of science principally concerns the investigation of facts through appropriate methods. It should not be equated with “positive” in the sense of norms enacted by an authorized lawmaker.

His project makes inherited morality increasingly available for conscious examination and revision. In that broad sense, it contributes to a movement from received obligation toward deliberate normative reconstruction. But it does not reproduce the institutional structure of legislation.

A legislature can repeal a legal provision through a recognized procedure. No corresponding declaration automatically removes a moral obligation from collective consciousness. Feelings, reputations, informal sanctions, and habitual judgments may persist after legal reform. Conversely, a moral expectation can weaken long before the legal order acknowledges the change.

The legal analogy therefore identifies a real development but also a limit. Conscious moral revision is possible; centralized enactment and repeal do not by themselves describe how it occurs. From a Lévy-Bruhlian perspective, even a formally proclaimed new moral code would remain an intervention whose actual effects required investigation.

His treatment of morality’s historical development reinforces this point. At the end of the book, he sketches **three broad forms of the relationship between moral practice and reflection**.

In the first, people inherit obligations as part of collective life without systematically investigating their origins or justification. In the second, philosophical and religious reflection organizes and legitimizes those obligations through more general principles. In the third, emerging scientific inquiry treats moral life as an object of investigation, making more informed practical intervention possible.

This is a provisional reconstruction, not a demonstrated universal sequence through which every society must pass. Lévy-Bruhl expressly declines to assert that all societies necessarily traverse the same stages.

The proposed transition also concerns the place of reflection within morality. Reflection begins by articulating and rationalizing the inherited order. It later becomes capable of investigating the order's formation and questioning particular elements through comparative knowledge.

This is the point at which moral change could become more consciously directed. But the resulting capacity would remain limited, historically situated, and dependent on continuing inquiry. *Ethics and Moral Science*, chapter IX, pp. 226–233.

His intellectual relationship with **Comte and Durkheim** helps locate this programme. From Comte he takes the ambition to study social life through discoverable relations and the insistence that social phenomena must be understood in their interdependence. He also shares the expectation that knowledge can expand humanity's practical capacities. Yet he examines critically the coexistence in Comte of scientific sociology and the religious veneration of Humanity. For Lévy-Bruhl, treating society as an object of dispassionate knowledge requires a difficult separation from the sentiments through which it is revered.

His proximity to Durkheim is evident in the treatment of moral obligations as social facts, the importance of collective representations, and the use of historical and comparative explanation. Their projects overlap substantially. Lévy-Bruhl's distinctive emphasis in this book is the separation of existing morality, its scientific investigation, and the future practical arts that investigation may support.

Finally, his **later anthropology** adds another dimension, although it must be used carefully. In works beginning with *Les fonctions mentales dans les sociétés inférieures* (1910) and *La mentalité primitive* (1922), he examines forms of collective representation in which emotional and mystical participation shapes experience.

The underlying relevance to moral change is that people's judgments operate within socially organized ways of experiencing relationships and events. A belief may be inseparable from affiliations, symbols, and practices through which a community understands itself.

His early terminology of a "prelogical" mentality is problematic, and its opposition to modern thought requires substantial qualification. In his late notebooks, published posthumously as *Carnets* in 1949, he expressly revises that opposition. He abandons the conception of a distinct primitive mentality characterized by prelogical and mystical features and recognizes mystical participation as present in human minds generally. These revisions preserve his interest in participation while altering the conceptual division through which he had earlier studied it. Lévy-Bruhl, *Carnets*, particularly the entries of June–August 1938.

An implication for moral reform can be drawn from this later position: modern populations also combine explicit reasoning with powerful symbolic and emotional attachments. Increasing education or scientific information need not eliminate those attachments. Reformers must understand how they operate.

That is an interpretive extension of his anthropology, rather than a fully elaborated later programme for creating morality. It also creates a productive tension within his work. The

earlier programme anticipates the emancipating effects of scientific understanding; the later investigations help explain why collective life may never become wholly transparent to, or governed by, detached reasoning.

Several critical judgments follow from the whole argument:

- **His strongest achievement is to make the conditions of moral reform an independent object of inquiry.** A defensible principle, sincere intention, and effective intervention are different things. Reform can fail because its proponents misunderstand how obligations and conduct are sustained.
- **His account of historical layering explains why conscience contains conflict.** Moral disagreements can arise within a person or community because obligations inherited from different social circumstances coexist. This directs attention to institutional change as well as conceptual clarification.
- **His distinction between conformity and improvement is indispensable.** A technique that increases obedience may strengthen either a beneficial or a harmful norm. The success of moral training cannot be assessed solely by compliance.
- **His criticism of philosophical abstraction sometimes goes too far.** Demonstrating that an argument has social origins does not establish that its reasoning is invalid. General principles can also reveal contradictions in prevailing practices and provide resources for criticizing them.
- **His confidence in causal social science exceeds what his own evidence establishes.** He recognizes complexity, uncertainty, and unintended effects, but the expectation of discovering sufficiently dependable laws to guide reform remains partly a scientific hope.
- **His account of moral authority is incomplete.** Empirical competence cannot resolve every conflict of ends or determine who should decide for others. The political and institutional organization of reform requires additional theory.
- **His comparative ambition is constrained by the hierarchical assumptions of his period.** He challenges the automatic universalization of European moral consciousness while also employing classifications of societies that require critical scrutiny. The methodological demand to investigate one's own assumptions must therefore be applied to his work itself.

Lévy-Bruhl's enduring contribution is the proposition that **a society's morality can become an object of deliberate, informed transformation, while remaining a reality that no individual thinker or governing body creates at will**. His rational moral art directs attention to the historical formation of obligations, the institutions supporting them, the emotions investing them with authority, and the consequences of attempts to change them.

The most fruitful development of his programme would preserve that empirical demand while making the normative and political dimensions more explicit: which changes are justified, whose interests count, who participates in deciding, and how the results are assessed. His work provides a substantial theory of why conscious moral change requires social knowledge; it leaves the full justification and legitimate governance of that change open.

****

John Dewey

John Dewey's central contribution to the theory of conscious moral change is his account of moral reconstruction: people can deliberately improve their moral judgments, habits, and institutions through inquiry into the problems and consequences of their shared lives. This makes moral change an ongoing educational and democratic activity. It also makes the reformer's own purposes subject to criticism: discovering how to achieve a proposed moral ideal can reveal that the ideal itself needs revision.

For your question, Dewey is especially significant because he connects three problems often treated separately: **how inherited morality becomes open to criticism; how criticism can actually change conduct; and how deliberate reform can acquire justified authority over collective life.** His answers draw together moral philosophy, social psychology, education, and democratic theory.

The account below reconstructs this position across his works. "Conscious moral change" is a useful description of their common concern, rather than the title of a single Deweyan doctrine.

His principal writings relevant to this question form the following sequence:

Work	Date	Contribution to the problem
My Pedagogic Creed	1897	Education as a deliberate agency of social reconstruction
Ethics, with James Hayden Tufts	1908; substantially revised 1932	Customary and reflective morality; judgment, obligation, character, and social life
Moral Principles in Education	1909	Moral formation through the organization of school experience
How We Think	1910; revised 1933	The development of disciplined reflective inquiry
Democracy and Education	1916	Democratic association and the capacity to readjust institutions
Reconstruction in Philosophy	1920	Reorienting philosophy toward practical human problems
Human Nature and Conduct	1922	Habit, impulse, deliberation, and changes in character

Work	Date	Contribution to the problem
The Public and Its Problems	1927	Public inquiry, expertise, communication, and collective action
"Three Independent Factors in Morals"	1930	The irreducible claims of good, right, and virtue
Liberalism and Social Action	1935	Economic power and the institutional conditions of freedom
Experience and Education	1938	Distinguishing educative growth from experiences that obstruct development
Theory of Valuation	1939	Critical appraisal and the reconstruction of desires and ends

Dewey's position developed from an early engagement with Hegelian idealism toward an experimental naturalism influenced by evolutionary thought and psychology. He retained a concern with the formation of individuality through social relationships, while increasingly explaining that formation through historically changing activities and institutions. Moral development consequently became a practical question about the conditions under which people acquire and revise their ways of living. It no longer required an account of progress toward a completed metaphysical order. Internet Encyclopedia of Philosophy, "John Dewey."

This development matters because **Dewey does not identify moral progress with the necessary course of history**. Increasing scientific knowledge, industrial production, or administrative coordination does not by itself establish that human relationships are improving. Such changes create possibilities and problems whose moral significance must be investigated.

For example, a society may become more productive while work becomes less meaningful; it may communicate more rapidly while public discussion becomes less intelligent. A Deweyan assessment therefore asks what changes do to people's capacities, relationships, and opportunities to participate in directing their lives. "Modernization" is a description of historical change; its moral justification remains a further task.

The initial distinction is between *customary morality* and *reflective morality*. In the jointly authored *Ethics*, customary morality receives authority from established ways of living, inherited expectations, and collective approval. Reflective morality arises when those expectations become objects of examination. Contact between different customs, changing circumstances, and conflicts within a community can make previously unquestioned practices problematic.

The historical treatment in the 1908 book was principally Tufts's contribution, so it should be attributed to Dewey and Tufts rather than exclusively to Dewey. Their account also recognizes losses and dangers in the transition to reflection: questioning inherited restraints can produce self-exemption and moral disintegration as well as improvement. Critical consciousness is therefore a possibility of progress, not proof of it. Dewey and Tufts, *Ethics*, preface and chapter IX.

This distinction gives a precise answer to the question of when moral change becomes conscious. A community may alter its practices through migration, economic pressure, imitation, or generational replacement without explicitly considering whether those changes are

desirable. Deliberate moral reconstruction begins when people investigate the authority, effects, and alternatives of a practice.

Consider an illustrative workplace norm: loyalty requires concealing colleagues' mistakes. The norm may initially appear morally admirable because it expresses solidarity. Reflection begins when someone recognizes that concealment exposes other people to harm. The problem then concerns the meaning of loyalty itself. Does loyalty require protecting colleagues from accountability, helping them correct mistakes, or preserving the trustworthiness of the institution?

The emerging moral change involves reorganizing a familiar value. Some inherited meanings are retained; others are rejected; new responsibilities become visible. This explains why deliberate reform often proceeds through reinterpretation before it produces an explicitly new rule.

Dewey's social psychology explains why changing an accepted rule is only one part of changing morality. In *Human Nature and Conduct*, habits are acquired dispositions that organize perception, emotion, judgment, and action. They include intelligent skills as well as rigid routines. Character consists substantially in their organization, and their operation depends on surrounding conditions. Consequently, appeals to willpower cannot reliably transform conduct while leaving the conditions sustaining it intact. Deliberation becomes possible when established activities encounter obstruction or conflict; Dewey describes it as a "dramatic rehearsal" of possible actions. Dewey, *Human Nature and Conduct*, especially Parts I and III.

An institution that officially praises honesty but penalizes anyone who reports failure illustrates the practical significance. Its moral declarations and its formative effects diverge. Employees repeatedly learn that concealment is safer, advancement depends on pleasing superiors, and responsibility means avoiding blame.

A campaign announcing a new ethical code might change the vocabulary without changing this practical education. Effective reform would also have to examine reporting procedures, promotion criteria, responses to mistakes, and the conduct of supervisors. These arrangements influence what people learn to notice, fear, admire, and do.

The example brings out a distinctive feature of Dewey's approach: **moral reconstruction operates through the relationship between persons and their environments**. It cannot be completed entirely within either one. Institutional reform without changes in conduct may remain formal; personal commitment without enabling arrangements may remain fragile.

Habits also explain why reflection cannot become an uninterrupted examination of everything. People need established capacities to act coherently. The aim is to cultivate habits sufficiently flexible to respond to unfamiliar circumstances, including a habit of reconsideration when difficulties arise. Dewey's psychology leaves room for novelty because inherited patterns do not exhaust the possibilities of impulse and activity. Education can cultivate curiosity, imagination, and responsiveness before existing routines become rigid. Elizabeth Anderson emphasizes this connection between the plasticity of conduct and Dewey's account of ethical inquiry. Anderson, "Dewey's Moral Philosophy," §§1–2.

This produces an apparent paradox with considerable importance for conscious moral change: a society requires stable habits that enable the revision of its habits. Examples would include

listening before condemning, investigating a disputed claim, acknowledging an error, and considering the experience of affected persons. Such dispositions make change possible without requiring every controversy to begin from the destruction of social trust.

The imaginative dimension deserves particular attention. Deliberation involves anticipating what different courses of action would be like to undertake and undergo. It can reveal consequences previously ignored and disclose possibilities beyond the alternatives initially presented. Moral imagination also concerns the agent's prospective character: what kind of person would I become through repeatedly acting in this way? Steven Fesmire's interpretation makes this creative and affective dimension central to Dewey's ethics. Fesmire, *John Dewey and Moral Imagination: Pragmatism in Ethics*.

Returning to workplace loyalty, a person may initially imagine only two possibilities: conceal wrongdoing or betray colleagues. More adequate imagination could identify confidential correction, collective acknowledgment, independent review, or support for the harmed party. The moral achievement is partly the discovery of a previously unavailable way of responding.

This also explains why sympathy has a cognitive function. Attending to another person's experience can disclose relevant features of the situation that an apparently coherent argument has omitted. Emotional responsiveness can improve judgment, although it too requires criticism: sympathy may be selective, sentimental, or captured by the most immediately visible suffering.

Dewey's mature ethics gives this complexity a more explicit structure. In "Three Independent Factors in Morals," he distinguishes three sources of moral consideration:

Factor	Experiential source	Question for moral reflection
The good	Desires, purposes, aspirations	Which ends deserve pursuit?
The right	Claims, demands, obligations	What may others legitimately require of us?
Virtue	Approval, admiration, blame	Which dispositions and forms of character deserve cultivation?

These considerations may support one another, but they can also conflict. A desirable result does not automatically settle a question of obligation; an established demand does not automatically deserve recognition; admired conduct may sustain damaging institutions. Moral difficulty can arise among serious claims rather than between an obvious duty and selfish temptation. Dewey therefore resists giving one factor an automatic priority that resolves every case. Dewey, "Three Independent Factors in Morals."

For deliberate reform, this suggests three interconnected objects of change. Reformers may need to reconsider what society pursues, what it requires, and what it admires. A community could formally prohibit an abusive practice while continuing to celebrate the character traits that reproduce it. Conversely, it could praise compassion while preserving rules that make compassionate action exceptionally difficult.

Dewey's pluralism is especially useful for analyzing such inconsistencies. It directs attention to the relationships between aspirations, obligations, and social approval. It also explains why

describing him simply as a utilitarian obscures his mature position. Consequences matter deeply, but his account does not reduce every morally relevant consideration to an aggregate quantity of pleasure. Fesmire traces the emergence of this pluralism in the writings surrounding the revised *Ethics*. Fesmire, "Dewey's Independent Factors in Moral Action."

The resulting method can be reconstructed as a sequence of activities. This is an analytical synthesis, rather than a fixed protocol Dewey prescribed for every reform:

1. **Recognize a disturbance.** Some established practice produces conflict, suffering, exclusion, or an inability to act coherently.
2. **Investigate the situation.** Determine whose experience matters, what conditions sustain the problem, and which assumptions have gone unexamined.
3. **Form possible responses.** Develop alternatives, including changes to the purposes initially taken for granted.
4. **Anticipate their consequences.** Examine effects on relationships, obligations, character, and future possibilities.
5. **Act on a considered judgment.** Give the proposed response a practical form appropriate to the circumstances.
6. **Examine what happens.** Compare actual consequences with expectations and revise the response or the original diagnosis.

The general structure draws on *How We Think*: reflection begins in difficulty, develops suggested explanations or solutions, reasons through their implications, and seeks evidence through observation and action. Its discipline includes resisting premature closure. Dewey, *How We Think*, chapters I and VI.

An extended example helps distinguish this approach from a campaign that simply propagates a predetermined norm. Imagine a school where pupils are expected never to question a teacher publicly. Complaints arise because errors and unfair treatment remain uncorrected.

An initial diagnosis might be that teachers need to become more tolerant. Inquiry could reveal a more complicated situation: teachers fear losing the ability to maintain order; pupils lack acceptable procedures for raising objections; administrators interpret complaints as evidence of professional failure; and classroom assessments reward repetition more than reasoned disagreement.

Several possible reforms then become available. The school could establish appeal procedures, introduce structured discussion, change how administrators evaluate complaints, and teach pupils how to challenge a claim responsibly. These changes require different resources and may produce different effects.

The proposed moral ideal might itself develop during this process. An initial demand for unrestricted expression could become a more precise commitment to reasoned contestability combined with responsibility for others' learning. Equally, an initial defense of obedience could become an account of authority that includes an obligation to explain and correct decisions.

The effects would then need assessment. Are previously silenced pupils actually heard? Do confident pupils monopolize discussion? Can teachers acknowledge mistakes without humiliation? Does the new arrangement develop judgment, or merely create another formal procedure?

This example is a contemporary illustration of the method. It shows why **the definition of the moral problem is itself revisable**. Inquiry can transform both the proposed remedy and the understanding of what requires remedy.

The reconstruction of purposes receives its most explicit treatment in *Theory of Valuation*. Dewey distinguishes *prizing*—holding something dear—from *appraisal*, which examines its conditions, relations, and consequences. The fact that something is desired does not establish that it deserves pursuit. An *end-in-view* is an anticipated outcome that directs present activity. Investigating the means needed to achieve it can alter its attractiveness or reveal a better purpose. Appraisal can therefore change what people value, rather than merely identify efficient ways to satisfy existing preferences. Dewey, *Theory of Valuation*.

Suppose the school's declared objective is "respect." At first this may mean immediate compliance. Investigation could show that compulsory deference produces resentment, concealment, and dependence. Respect might then be reconstructed to include attentive listening, justified authority, reciprocal responsibility, and the capacity to acknowledge error.

Here moral reasoning creates a more adequate object of commitment. It does not simply compare means while leaving the content of respect unchanged.

Dewey's treatment of means and ends also limits what reformers can justify in the name of a desirable future. His 1938 response to Trotsky argues that the methods of political action require inquiry into their actual relationship to liberation. A doctrine about historical necessity cannot establish that a particular means will achieve an emancipatory end. The choice of means must remain open to evidence, and their consequences may require reconsidering the political program itself. Dewey, "Means and Ends."

The implication for moral transformation is substantial. Practices used to establish a new order help form the people and relationships within it. A movement seeking independent judgment must examine what happens when it systematically rewards unquestioning loyalty. A reform promising public responsibility must examine the formative effects of secrecy.

This does not mean that every disagreement can be settled without authoritative decisions or coercion. It means that those decisions belong within the moral assessment of the reform. They cannot be exempted simply because the announced destination is desirable.

Dewey's idea of *growth* supplies another criterion. In *Reconstruction in Philosophy*, he treats moral life as continuing development and asks institutions to support the development of their members. Improvement concerns an increasing capacity to learn from experience and to direct activity intelligently. The ethical significance of government, economic arrangements, and education includes their effects on this continuing process. Dewey, *Reconstruction in Philosophy*, chapters VII–VIII.

Growth nevertheless requires qualification. Dewey explicitly considers the objection that a burglar can grow more skillful. In *Experience and Education*, he asks whether development in a particular direction opens or closes possibilities of further development. An experience may strengthen a limited competence while producing rigidity or obstructing broader learning. He also rejects the assumption that abandoning traditional authority makes every experience educationally valuable; teachers still bear responsibility for arranging conditions that support learning. Dewey, *Experience and Education*, especially chapters II–III.

Applied to morality, this means that efficient conformity cannot itself be the standard of successful reform. A new moral order might produce disciplined conduct while diminishing people's ability to recognize problems or challenge abuse. Its apparent success would then conceal a loss of the capacities needed for future correction.

Education becomes central because it shapes these capacities before any particular controversy arises. In *My Pedagogic Creed*, Dewey already identified education as a fundamental agency of social reform and connected character formation with institutional life. His early formulation placed striking confidence in the school's ability to reconstruct society. Dewey, *My Pedagogic Creed*, Article V.

Moral Principles in Education develops a more concrete distinction between knowing about morality and possessing ideas that effectively guide action. The entire organization of school life has moral effects: relationships, teaching methods, opportunities for cooperation, and the uses of authority. Punctuality and obedience cannot exhaust moral education. Pupils also need judgment, initiative, and the capacity to influence social change. Character requires effective action, understanding of circumstances, and concern for others. Dewey, *Moral Principles in Education*.

This yields an institutional test that extends beyond schools. An organization's operative morality can be examined by asking what people must repeatedly do to succeed within it. Its practical lessons may differ sharply from its official statements.

For instance, teaching cooperation through consistently competitive rewards creates a tension that verbal instruction alone cannot resolve. Similarly, teaching independent judgment while making disagreement professionally dangerous gives people a practical reason to distrust the proclaimed ideal.

In *Democracy and Education*, Dewey broadens the argument. Democratic association involves shared interests and communication across group boundaries. Such interaction exposes habits to unfamiliar experiences and creates possibilities for readjustment. Education should enable people to participate in these processes rather than confine some to executing purposes determined by others. The separation between those trained to direct and those trained merely to obey is therefore a moral as well as an educational problem. Dewey, *Democracy and Education*, especially chapters VII, XXIII, and XXVI.

The resulting conception of conscious moral reform is intergenerational. A reform is incomplete if it teaches only the currently preferred answers. It must also develop the capacities through which later participants can discover the limitations of those answers.

Dewey's account additionally gives cultural experience a role in changing moral perception. His aesthetics associates art with the renewal of attention and the disruption of habitual ways of seeing. Literature and other arts can make unfamiliar lives and possibilities available to imagination. Their ethical contribution need not consist in delivering explicit moral instruction. David Hildebrand's account explains this connection between aesthetic experience, diminished prejudice, and enlarged perception. Hildebrand, "A Brief Account of John Dewey's Ethics, Political Theory, and Philosophy of Art and Aesthetics."

This helps explain how moral change can begin before a new principle is formulated. People may first perceive an established practice differently—its humiliation, waste, cruelty, or

unrealized possibilities—and only subsequently develop the language needed to criticize it. Conscious reform includes bringing these emerging perceptions into articulate public examination.

The question of who participates in that examination leads directly to Dewey's democratic theory. In *The Public and Its Problems*, expert knowledge is indispensable, but specialists cannot independently determine all relevant social needs. Those who experience consequences possess knowledge that experts may lack. Effective public judgment requires investigation, accessible communication, discussion, and institutional responsiveness. Dewey also distinguishes experimental political inquiry from literal laboratory experimentation: policies must remain answerable to their effects, without pretending that society can be placed under fully controlled conditions. Dewey, *The Public and Its Problems*, chapters V–VI.

Applied to conscious moral change, this provides a reason for participation beyond the procedural claim that everyone should have a vote. Excluding people can damage the knowledge on which reform rests. A practice experienced as efficient by administrators may be experienced as degrading by its recipients. Neither description becomes adequate through authority alone; the relationship between them requires inquiry.

Participation also affects moral agency. People asked to understand and assess a reform occupy a different position from people whose conduct is merely to be modified. The former can become contributors to the purposes governing their lives.

"Creative Democracy—The Task Before Us" extends this argument to everyday dispositions. Democratic life depends on confidence in people's possibilities, openness to their experience, and ways of handling disagreement that permit mutual learning. Formal liberties become ineffective when communication is governed by contempt, fear, or intolerance. Democratic habits therefore require deliberate cultivation. Dewey, "Creative Democracy—The Task Before Us."

The moral significance is that **a reform changes the treatment of disagreement as well as the content of agreement**. The ability to criticize the reform, identify its harms, and propose alternatives forms part of its democratic quality.

Dewey's concern with participation should not obscure his recognition of material power. In *Liberalism and Social Action*, he argues that educational transformation cannot occur wholly within individual attitudes while institutions remain untouched. Concentrated economic power shapes opportunities, purposes, and the effective meaning of freedom. He calls for organized social intelligence and substantial institutional reconstruction, and recognizes that existing arrangements already exercise coercive power. His objection to dogmatic revolutionary violence does not amount to treating the status quo as innocent or noncoercive. Dewey, *Liberalism and Social Action*.

An analytical implication is that access to moral discussion has material conditions. Time, security, knowledge, organizational support, and protection against retaliation influence who can participate effectively. Calling for open discussion while ignoring these conditions may preserve the very exclusions that inquiry should expose.

For the analogy between conscious moral change and conscious legislation, Dewey's contribution can now be stated more precisely. The following comparison is a reconstruction of the implications of his position:

Question	Deliberate legal change	Deweyan moral reconstruction
What is changed?	Authoritative rules and institutional arrangements	Judgments, habits, values, expectations, and arrangements
Who initiates change?	Recognized lawmaking authorities, among others	Educators, associations, affected publics, investigators, political actors
How does a proposal gain force?	Enactment and institutional application	Persuasion, participation, altered practices, education, and institutional support
How is error addressed?	Amendment, repeal, review, and further interpretation	Criticism of consequences and renewed reconstruction
What establishes practical success?	Legal efficacy is one relevant dimension	Changes in lived conduct and relationships must also be examined

Dewey supports making inherited practices deliberately examinable and revisable. His framework does not establish a single moral authority whose declaration can make an obligation disappear in the way a competent legislature can repeal a statute. Moral authority remains connected to the reasons for a practice, its consequences, and its place in shared experience.

Law can nevertheless participate in moral reconstruction. Changes in legal arrangements can alter incentives, expectations, opportunities, and the conditions under which habits form. Dewey's "Logical Method and Law" itself treats legal reasoning as practical inquiry and criticizes reasoning that conceals consequential choices behind supposedly automatic deductions. Dewey, "Logical Method and Law."

A useful distinction follows. Legal enactment can settle what an institution will officially require, while leaving open whether people understand, accept, or habitually support that requirement. These are different dimensions of a reform. Their interaction is precisely what a comprehensive account of conscious moral change must investigate.

Several difficulties remain, and they should be treated as substantive problems in Dewey's theory.

The first concerns the standard of improvement. If moral inquiry evaluates consequences, it must still decide which consequences matter and how conflicting claims should be weighed. Dewey's commitments to growth, freedom, and participation give direction to inquiry, but they are themselves moral commitments. Experimental method does not generate them from value-free observations.

This need not invalidate his position. It does mean that his theory should be understood as a democratic ethical orientation joined to a method of learning. Describing it as an entirely neutral science of moral reform would conceal its normative basis.

Herbert Marcuse pressed a related objection in his review of *Theory of Valuation*. Socially produced desires can support authoritarian arrangements, and their coherence with available means does not establish their legitimacy. Marcuse questioned whether Dewey's treatment

sufficiently confronted the depth and social reach of such desires. This is a criticism of the adequacy of his justification, rather than evidence that Dewey endorsed authoritarian purposes. Marcuse, review of *Theory of Valuation*.

The second concerns power over the inquiry itself. Who can define the problem, select the evidence, and announce that a reform has succeeded? An employer may define success as reduced conflict while employees experience increased fear. A government may interpret silence as agreement. An educational authority may count obedience as moral development.

Dewey provides reasons to challenge these interpretations, especially through participation and attention to consequences. Yet the general philosophy does not supply a complete strategy for overcoming every concentration of power that blocks inquiry.

The third concerns the relationship between reform and its prerequisites. Democratic moral reconstruction needs some trust, freedom of communication, and willingness to reconsider. These may be precisely what the existing order lacks. A plausible Deweyan response is to pursue mutually supporting changes: limited institutional openings can develop capacities that support further openings. But this remains a difficult political process, not a solution guaranteed by the theory.

The fourth concerns irreversible harm. Learning from consequences is easier to defend when mistakes can be corrected. Some social interventions produce losses that later revision cannot repair. A responsible contemporary development of Dewey's approach therefore needs explicit attention to proportionality, protection of vulnerable persons, and the distribution of experimental risks. Calling a policy provisional does not make its consequences provisional.

The fifth concerns education and paternalism. Educators necessarily select experiences, establish expectations, and exercise authority. Their claim to cultivate freedom can conceal an attempt to reproduce their own preferred convictions. A demanding internal test would ask whether those educated are acquiring the capacity to scrutinize the educators' purposes as well as the purposes of their opponents.

The sixth concerns permanent disagreement. Better information and more imaginative alternatives will not necessarily reconcile all goods and obligations. Dewey's mature pluralism makes room for this fact, but it leaves judgment demanding. Some decisions remain contestable even after careful inquiry; institutional procedures must then make action possible while preserving avenues of challenge.

These difficulties help locate his originality. His theory explains moral change at several connected levels: the perception of a problem, the reconsideration of purposes, the formation of character, the organization of institutions, and the public assessment of consequences. It also places the reformer within the field of criticism. A proposal to improve others' morality must disclose its own assumptions and remain open to correction.

For further study, several interpretations illuminate different parts of this argument:

- **Jennifer Welchman, *Dewey's Ethical Thought* (1995):** the development of his ethics across his career. Cornell University Press.
- **Steven Fesmire, *John Dewey and Moral Imagination* (2003):** the creative, emotional, and imaginative dimensions of deliberation. Indiana University Press.

- **Gregory Fernando Pappas, *John Dewey's Ethics: Democracy as Experience* (2008):** the relationship between moral experience, character, and democratic life. Indiana University Press.
- **Kory Sorrell, "Pragmatism and Moral Progress: John Dewey's Theory of Social Inquiry" (2013):** an explicit defense of inquiry as a basis for moral criticism and progress across changing circumstances.

****

B. F. Skinner

B. F. Skinner's distinctive contribution to the theory of deliberate moral change is the proposal that societies can consciously redesign the conditions under which people acquire, practise, and transmit moral conduct. His central concern is how an ethical ideal becomes an effective way of life: how cooperation, honesty, restraint, and concern for others become dependable patterns of behaviour.

For Skinner, this requires examining the social consequences attached to conduct—approval, affection, prestige, economic advantages, criticism, exclusion, and punishment—and deliberately changing their arrangement. Moral reform consequently becomes part of the larger project he calls *cultural design*.

There is, however, an essential distinction within his position. Skinner offers both an explanation of how moral practices develop and a proposal concerning how they should be changed. His explanatory framework does not automatically establish the moral authority of his preferred reforms. Much of the philosophical importance—and controversy—of his work lies in this relationship between behavioural science and normative judgment.

His position develops across several works rather than appearing in a single systematic treatise on ethics. The principal texts are *Walden Two* (1948), *Science and Human Behavior* (1953), *Verbal Behavior* (1957), *The Technology of Teaching* (1968), *Contingencies of Reinforcement* (1969), *Beyond Freedom and Dignity* (1971), and *About Behaviorism* (1974). Later essays refine his treatment of democratic control, cultural survival, and alternatives to punishment. The B. F. Skinner Foundation's bibliography documents this development.

The word "conscious" needs careful interpretation in Skinner's theory. He accepts deliberate planning, self-observation, discussion, and the comparison of alternatives. What he rejects is the claim that these activities originate in an autonomous will operating independently of biological and environmental conditions.

His *radical behaviourism* does not deny thoughts, feelings, or private experience. It treats them as phenomena requiring explanation. Explaining a reform by saying that someone decided to improve society therefore leaves further questions: how did that person acquire the relevant concerns, knowledge, and capacities for intervention? Skinner's position is that the reformer belongs within the causal explanation, alongside the people whose behaviour the reformer hopes to change. Skinner, *About Behaviorism*, introductory excerpt.

This means that deliberate moral change is entirely possible within his framework. A teacher can investigate why children intimidate one another, consider alternative classroom arrangements, introduce changes, and examine their consequences. Those activities themselves have causes, but they also have effects.

The relevant distinction is between **unexamined social formation and deliberately investigated social formation**. Determinism, in this sense, does not imply that intervention is pointless. Intervention is one of the processes through which future behaviour is produced.

For moral theory, the consequence is considerable: Skinner relocates practical attention from the strength of a person's supposedly independent moral will to the conditions that make responsible conduct possible and sustainable.

The basic mechanism is operant learning: consequences alter the future probability of behaviour. Several terms are necessary to understand the proposal accurately:

Concept	Meaning	Illustrative moral relevance
Positive reinforcement	A consequence is added, increasing the behaviour that preceded it.	Constructive participation becomes more frequent when it receives effective recognition.
Negative reinforcement	Behaviour increases because it removes or avoids something aversive.	Someone complies to stop persistent criticism.
Punishment	A consequence is arranged to suppress behaviour.	A person is sanctioned for deception.
Extinction	Behaviour previously reinforced ceases to receive that reinforcement.	Provocation may diminish when it no longer produces its usual attention.
Shaping	Successive approximations to a more complex performance are reinforced.	A child gradually learns the component skills needed for cooperative disagreement.

"Positive" and "negative" here mean addition and removal; they do not mean morally good and morally bad. Negative reinforcement is also different from punishment: the former strengthens the behaviour that produces relief.

A further distinction concerns the difference between a reward and a reinforcer. An authority may intend a reward, but whether it reinforces a particular behaviour is an empirical question. Public praise, for example, may be welcome to one person and embarrassing to another.

Applied to morality, this framework directs attention to discrepancies between what institutions officially commend and what they actually make advantageous. An institution may proclaim honesty while promoting those who conceal difficulties. The proclamation and the effective consequences then pull conduct in different directions.

Skinner understands ordinary morality as a socially maintained pattern of control. In *Science and Human Behavior*, especially the chapter on group control, ethical classifications emerge through the reactions of other people. Conduct is encouraged or discouraged through diffuse social practices, often without any centralized authority or written code.

This does not mean that he regards prevailing morality as automatically justified. He explicitly recognizes that moral classifications may preserve superstition, outlive the circumstances in which they developed, or reflect unequal participation in the group. Existing approval therefore

cannot simply be equated with everyone's welfare. Skinner, *Science and Human Behavior*, chapter XXI.

The analytical importance of this account is that morality has more than a verbal existence. A prohibition can remain publicly accepted while becoming practically ineffective; conversely, a new expectation can become powerful before anyone formulates it as a general ethical principle.

Consider an organization that starts treating the admission of mistakes as responsible professional conduct. Its moral vocabulary may change first. But a deeper transformation occurs only when admitting mistakes no longer predictably brings humiliation, when corrective action receives support, and when concealment ceases to secure advancement. This is my illustrative application of Skinner's framework, rather than an example taken from his writings.

On this interpretation, studying moral change requires following both **changes in stated judgments and changes in the social consequences of action**.

Moral language itself also becomes an object of explanation. Skinner's treatment of verbal behaviour asks how a person's speaking is shaped by a community of listeners. Applied to ethics, this means investigating the circumstances in which people learn to call something good, wrong, admirable, or shameful.

Alexandre Dittrich's analysis clarifies that this should be understood as a functional investigation of evaluative language, rather than as a claim that every use of "good" transparently reports an immediate pleasure. People may repeat inherited judgments, adjust their language to an audience, or condemn conduct that they nevertheless find rewarding. Ethical speech is affected by several interacting conditions. Dittrich, "The True, the Good, and the Beautiful" (2023).

This permits a useful distinction between changing someone's moral declarations and changing the conduct those declarations ostensibly describe.

Suppose employees learn that a particular ethical vocabulary is necessary for professional acceptance. They may become highly proficient at using it. That result alone does not show that their treatment of colleagues has changed. An assessment of moral reform would therefore need to examine behaviour beyond statements of allegiance.

The distinction cuts both ways. A person might reliably help others without possessing an elaborate philosophical vocabulary for explaining that conduct. A verbal test of moral sophistication could underestimate such practical reliability.

Rules and reasoning still have an important place in Skinner's account. In *Contingencies of Reinforcement*, he distinguishes behaviour shaped through direct encounters with consequences from behaviour governed by rules. A rule can guide conduct before the individual personally experiences the consequences it describes. Instructions, warnings, and ethical precepts can therefore transmit the results of experience across people and generations.

Following a rule, however, has its own learning history. Moreover, conduct maintained by obedience to an authority may differ from conduct sensitive to the practical circumstances the rule concerns. Skinner, *Contingencies of Reinforcement*.

This distinction is particularly valuable for analysing the replacement of inherited norms. There are at least two possible failures.

First, a community may retain an old rule even when the circumstances that once supported it have changed. Second, reformers may introduce a new rule that wins verbal acceptance but has little influence on everyday action.

A successful reform would have to address both the intelligibility of the rule and the conditions under which people follow it. Merely substituting a new ethical formula for an old one leaves much of the practical problem unresolved.

There is also a philosophical question that Skinner's analysis does not settle simply by redescribing reasoning as behaviour: whether the validity of a reason supplies a distinctive basis for accepting a norm. A causal account of why an argument persuades someone and an assessment of whether that argument is sound remain distinguishable inquiries.

Education becomes the principal means of developing the capacities a changed morality requires. Skinner's educational programme emphasizes active performance, appropriately sequenced tasks, feedback, and the progressive acquisition of competence. *The Technology of Teaching* includes a specific discussion of discipline, ethical behaviour, and self-control. Its larger ambition includes teaching complex performances rather than merely requiring students to repeat correct statements. Skinner, *The Technology of Teaching*.

For moral reform, the significance is that ethical expectations often presuppose abilities that people have not been adequately taught.

"Resolve disagreements peacefully" requires a repertoire: listening, describing a grievance, tolerating interruption, proposing alternatives, and recognizing when negotiation has failed. "Take responsibility" may require learning how to report an error, estimate its consequences, and organize repair.

My reconstruction of the educational implication is that reformers should ask three separate questions:

1. What conduct are we expecting?
2. Do people possess the skills needed to perform it?
3. Does their environment support its continued use?

A moral programme that answers only the first question can mistake an absence of competence or opportunity for a defective character.

Self-control is relevant here as well. In a behavioural interpretation, people can learn to arrange their own circumstances: avoid a recurring temptation, seek advice before acting, create reminders, or establish procedures that make impulsive action less likely. Such capacities are forms of practical agency even though Skinner does not explain them through an uncaused inner will.

His attack on punitive morality concerns the behaviour of punishers as well as that of the punished. In "The Non-Punitive Society," Skinner argues that punishment persists partly

because its immediate results reinforce those administering it. A threat may rapidly stop an irritating action, while resentment, evasion, withdrawal, or retaliation appear later.

He consequently advocates developing constructive alternatives and attending to desirable conduct. He also acknowledges that behavioural techniques can serve exploitative management. His preference for positive reinforcement is therefore not a demonstration that every reward arrangement is humane. Skinner, "The Non-Punitive Society" (1979).

This supplies an unusually reflexive account of moral reform. A society wishing to reduce cruelty must investigate what makes cruelty effective or advantageous for those exercising authority. Condemning cruelty without altering those conditions may leave its reproduction largely intact.

For example, imagine that a supervisor obtains immediate compliance by publicly humiliating employees. Complaints about humiliation are then treated as disloyalty. In such a setting, asking the supervisor to become more considerate addresses only one part of the problem. The organization also needs to change how supervisory performance is evaluated and how employees can respond.

The point is not that punishment can never suppress conduct. It is that suppression alone does not establish the acquisition of an alternative, the legitimacy of the intervention, or its beneficial long-term consequences.

A practical account of deliberate moral change can be reconstructed from these ideas. Skinner does not present the following table as a single standardized procedure; it is an analytical synthesis of the approach.

Task	Question for the reformer
Specify the proposed change	What would people actually do differently?
Investigate existing practices	What currently makes the unwanted conduct effective, attractive, or difficult to avoid?
Examine the reforming institution	What encourages authorities to preserve the present arrangement?
Develop alternatives	What skills and opportunities make the desired conduct possible?
Change social consequences	How will cooperation, honesty, or restraint be supported in practice?
Examine transfer	Does the change persist outside the setting in which it was taught?
Evaluate wider effects	Who benefits, who bears costs, and what unintended conduct emerges?
Revise the intervention	What evidence would justify modifying or abandoning it?

Consider a hypothetical school seeking to replace a culture of ridicule with one of cooperative learning. A declaration against ridicule may help establish expectations, but the investigation cannot stop there. Ridicule may bring peer attention, adults may model it, and students may lack alternative ways of gaining recognition.

An intervention informed by Skinner would examine those relationships, teach alternative forms of participation, and make helpful contributions socially effective. Evaluation would then ask whether pupils behave differently when the teacher is absent, whether exclusion has taken less visible forms, and whether the apparent improvement depends entirely on temporary prizes.

These questions also expose a limitation of any narrowly measured intervention. A falling count of public insults could conceal increasing private intimidation. Behavioural precision is useful only if the chosen observations adequately represent the moral problem.

The difficulty becomes greater when Skinner moves from changing behaviour to identifying the ends of change. His discussions distinguish personal goods, the good of others, and the good of a culture. These interests can diverge. Conduct immediately advantageous to an individual may harm other people; practices beneficial to current members may undermine future conditions of collective life.

Camila Muchon de Melo, Marina Souto Lopes Bezerra de Castro, and Julio C. de Rose distinguish the descriptive and prescriptive elements of this position. Skinner's preferred cultural design seeks to combine individual well-being with collective survival, but that preference does not follow logically from the behavioural methods themselves. The same methods could be directed toward different ends. De Melo, de Castro, and de Rose, "Some Relations Between Culture, Ethics and Technology in B. F. Skinner" (2015).

This distinction prevents two misunderstandings. Skinner is not simply recommending the maximization of immediate rewards. Nor does the capacity to produce a particular behaviour establish that producing it is morally desirable.

My assessment is that his programme contains three different questions:

- **Explanatory:** Why do people acquire and maintain particular values?
- **Technical:** How can conduct and evaluative practices be changed?
- **Justificatory:** Which changes ought to be pursued, and under whose authority?

His strongest distinctive contribution concerns the first two. His answers to the third remain substantially more contested.

Cultural evolution explains why he considers deliberate intervention necessary. In "Selection by Consequences," Skinner distinguishes biological evolution, the selection of behaviour during an individual's lifetime, and the selection of cultural practices. These are connected processes, but they operate at different levels.

Practices can spread or persist because of their effects on groups, even when individuals neither understand nor intend those effects. Deliberate cultural design introduces and evaluates alternatives within this evolutionary process; it does not require a designer standing outside nature. Skinner, "Selection by Consequences" (1981).

The moral significance can be expressed through a problem of timing. Immediate advantages may support practices whose destructive effects emerge much later. Waiting for disastrous consequences to eliminate those practices may be an unacceptable method of learning.

A contemporary illustration would be an institution that obtains short-term benefits by exhausting a shared resource. Merely informing participants about eventual exhaustion may be insufficient if their present circumstances strongly favour continued exploitation. Deliberate reform would attempt to make longer-term consequences relevant to decisions now.

This illustration does not prove that a particular incentive scheme will work. It identifies what an adequate explanation must address: the relationship between present action, delayed consequences, and the institutions connecting them.

Skinner's treatment of freedom and dignity changes the moral meaning of reform itself. In *Beyond Freedom and Dignity*, he challenges explanations that assign ultimate authorship to an autonomous individual. He also questions practices of allocating credit and blame while overlooking the conditions that produced conduct.

He nevertheless recognizes the historical achievement of resistance to aversive control. His controversial claim is that the absence of felt coercion does not mean behaviour is uncontrolled. Positive reinforcement can influence people without producing the experience of being forced. Skinner, *Beyond Freedom and Dignity*.

For analysis, this suggests separating several meanings often combined under "freedom": freedom from threats, practical access to alternatives, participation in decisions, and metaphysical independence from causes. Rejecting the last does not resolve the status of the others.

Likewise, explaining wrongdoing causally need not eliminate every institution of accountability. It changes the questions asked of that institution. Does accountability prevent further harm, support repair, and improve future conduct? Or does it principally express deserved suffering? These possibilities require separate justification.

The most concrete representation of his project is the fictional community in *Walden Two*. Its moral education aims to cultivate cooperation and generalized self-control. Frazier, the community's principal spokesman, argues that no code can specify every future circumstance, so education must develop capacities extending beyond particular instructions.

The famous exercises involving temptation and delay illustrate this ambition. Some training also includes deliberately arranged aversive experiences, which complicates an entirely benign reading. Moreover, the novel's narrator questions whether Frazier has experimentally established everything he claims. The book is a proposed social experiment presented through fiction, not a report demonstrating its own success. Skinner, *Walden Two*, chapters 13–14.

The imagined community makes moral development inseparable from the organization of everyday life. Its significance is therefore broader than rewarding individual good deeds: economic arrangements, education, work, and interpersonal relations are considered together.

This raises a useful distinction between **reforming people within an institution and reforming the institution that continually produces the problem**. If an arrangement systematically encourages rivalry, moral instruction against envy faces a persistent obstacle. Changing the arrangement may alter the frequency and meaning of the conflict.

But the institutional design of *Walden Two* also contains a political difficulty. Roy Moxley documents how its revisable code coexists with restrictions on collective debate: members can pursue objections through administrative channels, while general campaigning about the code is constrained. Experimental revisability therefore does not automatically amount to democratic participation. Moxley, "B. F. Skinner's Other Positivistic Book: Walden Two" (2006).

The underlying issue is whether residents participate in determining what counts as improvement or principally supply evidence for judgments made by planners. Those are different distributions of moral authority.

Skinner's later position cannot be reduced to the institutions of his novel. In the 1976 preface, "Walden Two Revisited," he acknowledges the speculative character of practical behavioural engineering when the novel was written. He advocates small experimental communities and practical changes in consumption, work, education, and childcare. His proposed route includes cultural experimentation outside the pursuit of national political office. Skinner, "Walden Two Revisited".

In "Human Behavior and Democracy," he criticizes concentrated institutional power and emphasizes direct reciprocal influence between people. The behavioural expert need not remain permanently in command: a successful intervention can establish practices that continue without the expert's supervision. Skinner, "Human Behavior and Democracy," in *Reflections on Behaviorism and Society*.

This is relevant to his concept of *countercontrol*: the responses through which those subject to control resist, escape, or modify it. Yet recognizing reciprocal influence is only the beginning of an institutional solution. An employee and an employer may influence one another while possessing radically unequal capacities to make their objections effective.

A more adequate assessment therefore asks whether criticism can be expressed without retaliation, whether information about failure reaches decision-makers, and whether those affected can alter the programme's goals. These are my proposed criteria for evaluating the political adequacy of a Skinnerian intervention; they should not be mistaken for a constitutional scheme fully supplied by Skinner.

James Dinsmoor made a related criticism from within behaviourism: positive reinforcement does not require government by behavioural technicians, and behavioural knowledge can coexist with democratic oversight. Dinsmoor, "Walden Two: A 1950 Review," published in 1991.

The deepest philosophical objection concerns the difference between explaining a value and justifying it. Carl Rogers pressed this issue in his published exchange with Skinner. Scientific assessment needs a criterion of success; choosing that criterion raises another evaluative question. Rogers emphasized personal responsibility, creativity, and openness to changing goals. Skinner replied that choices and values themselves belong within the field of causal explanation. Rogers and Skinner, "Some Issues Concerning the Control of Human Behavior" (1956).

The disagreement is not resolved merely by conceding that all valuation has causes. Consider two programmes that successfully teach different moral priorities. One cultivates obedience to

established authority; the other cultivates willingness to criticize it. Demonstrating that both programmes work does not identify which should be adopted.

A causal history can explain why a reformer favours obedience. It can also explain why another favours dissent. To adjudicate between them, further considerations are needed concerning harms, interests, rights, social cooperation, or other defensible standards.

Skinner's challenge remains valuable: those further considerations must become effective in actual social life. Nevertheless, the need to explain their effectiveness does not make their justification dispensable.

Cultural survival is especially problematic as a final standard. Fernanda Brunkow and Alexandre Dittrich identify ambiguities in Skinner's usage: survival may concern the continued existence of people or the persistence of practices, and it may concern a particular group or humanity as a whole. These meanings cannot be freely substituted for one another. Brunkow and Dittrich, "Cultural Survival in B. F. Skinner" (2021).

An institution can survive while treating its members badly. A practice can disappear while the people who abandoned it flourish. A reform may preserve human life precisely by eliminating an established cultural arrangement.

My assessment is therefore that survival supplies a powerful warning against self-destructive practices, but an incomplete criterion for selecting among sustainable ways of life. Even if two social orders are equally durable, their distribution of opportunities, treatment of minorities, and accommodation of dissent may differ substantially.

There is also a problem of knowledge. Long-term cultural survival cannot ordinarily be observed within the time available for evaluating an intervention. Reformers must use intermediate indicators, and those indicators themselves require justification.

A second major objection concerns the extrapolation from controlled experiments to complex societies. Noam Chomsky's 1971 critique challenged Skinner's movement from experimentally specified reinforcement to broad explanations of human life. He also argued that ostensibly positive rewards can operate within relationships of dependency: access to employment or advancement can make conformity highly consequential without overt physical coercion. Chomsky, "The Case Against B. F. Skinner".

These criticisms should be distinguished from a wholesale dismissal of learning through consequences. The narrower and more persuasive question is whether a particular social explanation identifies the relevant relationships independently and predicts something beyond what was already observed.

Calling every successful influence "reinforcement" after the event would add little. A useful explanation needs to specify what consequence matters, for whom, in which circumstances, and with what observable effects.

Similarly, an increase in measurable compliance does not by itself demonstrate moral improvement. People may conform only under observation, become dependent on a reward, or learn to manipulate the measurement. Conversely, meaningful change may involve capacities—

such as resisting a corrupt demand—that reduce compliance with an institution's immediate expectations.

These problems call for richer evaluation, rather than the assumption that behavioural measurement is either sufficient or worthless.

Skinner is particularly significant for the idea that morality might become consciously revisable in a manner analogous to legislation. The analogy is useful, but it needs qualification.

A deliberately enacted rule can change an official normative position at a specified moment. Its social influence may nevertheless develop gradually. With morality, the distributed practices of families, peers, workplaces, religious communities, and other associations make coordinated replacement still more difficult.

From Skinner's perspective, changing a moral code would leave a substantial task unfinished if the old pattern of approval, shame, advantage, and exclusion continued. A formally repudiated expectation might retain considerable practical power.

The following comparison is my analytical reconstruction:

Dimension	Deliberate rule replacement	A fuller change in moral practice
Formulation	A new standard is stated.	People learn to recognize and apply it.
Withdrawal	The old standard loses official endorsement.	Old habits of approval and condemnation gradually change.
Institutional support	Procedures are amended.	Everyday opportunities and consequences become compatible with the new expectation.
Evaluation	Compliance with the stated standard is examined.	Effects on conduct, relationships, welfare, and dissent are examined.
Further revision	The rule can be amended again.	Supporting practices must also remain open to criticism and adjustment.

This comparison reveals the originality of Skinner's proposal. The object of deliberate intervention extends beyond the text of moral norms to the social processes that give them force.

It also reveals the political danger. Reformers who can influence rewards, education, reputation, and opportunities may exercise substantial moral power without holding anything resembling a publicly recognized legislative office.

The question of authority therefore becomes unavoidable: who may initiate such changes, through which procedures, and with what possibilities of refusal or revision? Behavioural knowledge makes that question more pressing; it does not answer it by itself.

His project did generate practical experiments, although their evidential significance must be stated carefully. Comunidad Los Horcones, founded in Mexico in 1973, described its undertaking as an application of behavioural analysis to cultural design. Its members emphasized experimental methods rather than reproducing every feature of the fictional *Walden*

Two. Their account documents an attempt to organize communal life around these ideas. Comunidad Los Horcones, “Walden Two and Social Change” (1989).

Such experiments establish that the programme could inspire sustained practical activity. They do not, simply by existing or by reporting favourable experiences, prove that a complete society can be successfully redesigned on Skinnerian principles. Voluntary membership, participant selection, local circumstances, and dependence on surrounding institutions complicate the inference.

For evaluating conscious moral change, the enduring value of Skinner’s work is a demanding set of questions. What actually sustains the conduct a society condemns? What makes the conduct it praises difficult? What rewards authorities for continuing ineffective or harmful practices? How can people acquire the capacities needed for a different way of life?

Those questions expose the weakness of reforms confined to exhortation. But they also need to be directed at the reformers themselves. **A defensible programme of deliberate moral change must explain how its practices work, justify the ends they serve, and make its own exercise of power open to effective challenge.** Skinner provides unusually powerful resources for the first task and forces attention onto the institutional conditions of the third; the justification of the ends remains the central unresolved problem in his theory.

****

Howard S. Becker

Howard S. Becker's central contribution is an explanation of how people deliberately change the rules by which conduct is morally judged, and how those changes acquire social force. He also explains how sociological research can prompt people to revise their moral judgments. These are substantial contributions to understanding conscious moral change, although Becker did not develop a comprehensive ethical system specifying what a morally improved society should look like.

His work therefore addresses two connected questions: **How do organized actors establish new moral definitions? And how can investigation expose reasons for reconsidering existing ones?** The distinction matters because successfully establishing a norm does not, by itself, demonstrate that the norm is morally justified.

The account below distinguishes Becker's explicit arguments from the broader implications I draw from them for deliberate moral transformation.

Becker's approach emerged from empirical sociology. His early studies concerned musicians, marijuana users, occupational careers, and professional socialization. Their common concern was how people acquire shared understandings and organize their activities. His principal book for the present question is *Outsiders: Studies in the Sociology of Deviance* (1963), especially its treatment of rule enforcement and moral entrepreneurs. Other relevant writings include "Notes on the Concept of Commitment" (1960), "Whose Side Are We On?" (1967), the essay written with Irving Louis Horowitz on radical politics and research (1972), and "Labelling Theory Reconsidered," included in the expanded 1973 edition of *Outsiders*. Becker's own bibliography establishes this development across the sociology of occupations, deviance, and knowledge. Becker's bibliography

The starting point is that moral classification is itself something to explain. When a society condemns a practice, the sociologist must investigate both the practice and the activities through which people identify, publicize, judge, and sanction it.

Erich Goode's examination of Becker usefully emphasizes this change of perspective: Becker made unconventional conduct intelligible by studying its organization and the processes through which it became socially defined as deviant. The familiar questions about why people engage in an activity remain relevant, but they must be joined by questions about why others condemn it and what their condemnation accomplishes. Goode, "Howard Becker, Drug Use and the Sociology of Deviance"

For the analysis of moral change, this suggests three analytically independent developments:

- People may change what they do.
- Observers may change how they evaluate an unchanged practice.
- Institutions may change how they respond to that practice.

Consider a hypothetical community in which an established recreational activity becomes the target of a reform campaign. Participation might remain constant while public condemnation increases. Alternatively, public attitudes might become more tolerant while an old prohibition remains enforceable. A third possibility is that authorities stop imposing sanctions even though many citizens retain their disapproval.

Calling all three developments "moral change" conceals their different causes. A Becker-inspired investigation would examine their timing and connections separately.

This also clarifies what social construction means here. The existence of organized classification does not make bodily injury, coercion, or material loss imaginary. The relevant question is how societies select conduct for moral attention, describe its significance, assign responsibility, and decide upon responses. Evidence about consequences and analysis of classification belong together.

Becker's most directly relevant concept is the moral entrepreneur. In *Outsiders*, he distinguishes actors who create rules from those who enforce them. The exemplary creator is a committed reformer who regards some condition as intolerable. Such reformers can have sincere humanitarian purposes. Once rules become institutionalized, their administration introduces occupational interests that may differ from the reformers' original intentions. Becker, *Outsiders*, chapter 8

The word *entrepreneur* directs attention to initiative. For analytical purposes, it is useful to distinguish possessing a moral conviction from undertaking the work needed to make it consequential for other people. Many individuals privately disapprove of something. A smaller number create associations, formulate proposals, seek publicity, recruit supporters, or press institutions to act.

This distinction gives Becker's approach particular relevance to conscious moral change. It identifies a missing middle between moral philosophy and established public morality: the practical work through which a judgment becomes a collective demand.

A moral principle does not organize its own campaign. Someone must decide which issue deserves attention, which audience to address, what examples to present, which institution can intervene, and what change to request. These decisions partly determine what the principle will mean in practice.

The concept should therefore be used symmetrically. Identifying someone as a moral entrepreneur does not establish that the person is dishonest, authoritarian, or mistaken. Nor does it establish moral superiority. Opposed campaigns can exhibit comparable forms of organization while advancing very different purposes.

A further implication concerns sincerity and interest. These should not be treated as mutually exclusive explanations. A reformer may care deeply about an issue while also acquiring prestige

through campaigning. An organization may provide a valuable service while seeking a larger jurisdiction. The empirical task is to establish how motives, opportunities, and institutional consequences interact.

Deliberate moral change often involves altering the application of an accepted value rather than inventing an entirely new value.

This is an analytical reconstruction particularly useful for extending Becker's account. Suppose the broadly accepted value is protecting vulnerable people. Agreement on that value does not determine:

- Which people count as vulnerable.
- What constitutes a threat to them.
- Who bears responsibility.
- Which restrictions are justified.
- How competing interests should be accommodated.

A reform movement can transform public morality by changing the answers to these questions while continuing to invoke familiar ideals.

For example, two hypothetical campaigns might both appeal to personal responsibility. One could argue that responsibility requires prohibition of a risky activity. Another could argue that responsibility requires informed individual choice. Their disagreement concerns the interpretation and institutional consequences of an ostensibly shared value.

This helps explain why moral transformation is often presented as fidelity to morality rather than its replacement. Reformers may claim that society has failed to recognize the implications of its own commitments. The sociologist must then investigate how that claim becomes persuasive, who accepts it, and what practical obligations it generates.

The argument's philosophical validity and its social effectiveness remain distinguishable. An excellent argument can remain politically marginal. A weak argument can become influential through favorable organizational circumstances. A complete account should investigate both without reducing one to the other.

The conversion of a moral demand into an operational rule is a substantive transformation.

To make the process concrete, consider a hypothetical professional association that declares respect for clients to be a central obligation. That declaration does not yet determine how members must behave. Operationalization requires decisions about disclosure, confidentiality, conflicts of interest, complaints, evidence, and sanctions.

Those decisions distribute both protection and authority. They determine who can complain, whose testimony matters, which conduct falls within the association's jurisdiction, and who judges disputed cases.

Consequently, the people who translate moral principles into procedures can influence the resulting morality considerably. Lawyers, administrators, professional specialists, and committee members may resolve questions that the initial campaign left open. Their

contribution cannot be assessed merely by asking whether they accurately copied the reformers' language.

Later research illustrates this implication. Mira D. Vale's study of digital psychiatry distinguishes efforts to identify new ethical problems from efforts to establish which professionals are entitled to regulate them. Moral entrepreneurship can help constitute an authorized community of decision-makers as well as identify unacceptable conduct. This is a later development of the approach, rather than a claim that Becker himself analyzed digital psychiatry. Vale, "Moral Entrepreneurship and the Ethics of Artificial Intelligence in Digital Psychiatry"

The broader inference is that **moral reform can redistribute the authority to speak morally**. Its outcome may include a new norm, a new procedure, and a new class of recognized interpreters.

Public communication is another part of the process, because a condition must become an issue for audiences beyond those already concerned.

Becker's discussion of the American Marihuana Tax Act of 1937 became a major example in this research tradition. Stanley Cohen explicitly identifies Becker's account as showing how organizational initiative, publicity, and coordinated action contribute to the establishment of prohibitions. Cohen subsequently developed his own account of moral panics, particularly through the study of reactions to the Mods and Rockers. Moral entrepreneurship and moral panic should therefore remain distinct concepts: the former does not require the latter. Cohen, *Folk Devils and Moral Panics*

The general analytical lesson is that publicity does more than increase the number of people who know about an issue. It can also establish the categories through which they understand it.

A campaign might describe an activity as a private preference, a public danger, an instance of exploitation, or a professional failure. Each description suggests different responsible actors and different remedies. The choice of description can therefore shape the eventual institutional response.

Nevertheless, one should not infer from the existence of campaigning that the danger was invented. Three questions require separate investigation: whether a harmful condition exists, whether its public presentation is accurate, and whether the proposed response is justified.

Nor should the 1937 example be treated as a sufficient historical explanation of every drug prohibition. Its value for the present discussion lies in identifying processes that can be investigated elsewhere, not in supplying a universal historical formula.

Enforcement is a further site at which the practical meaning of morality is formed.

The distinction between rule creation and rule enforcement allows us to identify an implementation problem. A campaign may define success as eliminating a wrong. The institution subsequently responsible for enforcement must manage evidence, personnel, budgets, procedures, and competing demands.

These requirements can produce divergence without anyone deliberately betraying the original principle.

Consider a hypothetical institution created to prevent professional misconduct. Its staff may find some violations easier to document than others. Their reports may consequently emphasize the measurable cases. Decision-makers may then allocate more resources to those cases because they dominate the statistics. Over time, the institution's practical definition of misconduct can narrow around what it is equipped to detect.

The same issue arises with performance measures. A higher number of recorded violations may indicate greater underlying misconduct, improved detection, a broader definition, or more intensive enforcement. It cannot be interpreted adequately without examining the process that generated the records.

An agency also faces a potential tension between demonstrating effectiveness and demonstrating continuing need. This is an organizational problem worth investigating; it is not evidence, by itself, of deliberate fabrication.

For conscious moral reform, the implication is substantial: **a reform's institutional effects cannot be inferred directly from its founders' intentions**. Assessment must extend to the routines that develop after adoption.

Moral campaigns can also change the social identities and opportunities of the people they target.

The wider literature on societal reactions distinguishes the creation of categories from the consequences of applying them. Research on labeling examines how institutional treatment can affect subsequent relationships, opportunities, and behavior. These consequences are empirical questions, with variations across settings and populations. "Societal Reactions to Deviance," *Annual Review of Sociology*

Applied to moral reform, this raises a question that is easily neglected: does a campaign seek to prevent particular acts, or does it establish a durable classification of particular people?

The difference matters. A person can acknowledge a specific wrongdoing while resisting a judgment that the wrongdoing defines the whole person. An institution can prohibit conduct while offering routes to correction, or it can make exclusion effectively permanent.

Suppose a disciplinary sanction restricts employment long after the original offense. Its consequences then extend beyond immediate deterrence or moral expression. It reshapes the person's available future.

A serious evaluation of conscious moral change should therefore examine both its intended protection and its effects on those classified as violators. This does not entail opposition to sanctions. It entails asking what particular sanctions actually accomplish.

The converse deserves equal attention. Removing a prohibition may leave a stigma intact. Formal rehabilitation may fail to restore access to employment, friendship, or professional standing. Institutional reform and social acceptance can proceed at different speeds.

Becker's earlier work adds another dimension: people's evaluations can change through participation and learning.

In "Becoming a Marihuana User" (1953), Becker studied how interpreting an experience involves social learning. His account distinguished learning the practice, recognizing its effects, and interpreting those effects as enjoyable. The significance for the present inquiry is that motives and evaluations can develop within a sequence of participation rather than being fully formed beforehand. Becker, "Becoming a Marihuana User"

"Marihuana Use and Social Control" (1955) examined how access, concealment, and moral reservations affected continuing participation. It consequently connected individual conduct with relationships to conventional audiences and alternative social circles. These were studies of a particular activity, rather than demonstrations of a universal model of moral conversion. Becker, "Marihuana Use and Social Control"

Their broader relevance can be stated as a hypothesis: changing the social setting in which people encounter and interpret a practice may alter moral judgment, even without a comprehensive argument addressed to them in advance.

A person entering a new professional setting, for instance, may encounter cases that complicate earlier convictions. Colleagues supply distinctions and examples. Repeated participation changes what the person notices and regards as significant. Moral learning may then occur through experience interpreted with others.

This possibility makes the relationship between moral education and institutional organization especially important. The explicit lesson taught in a classroom can differ from the practical lesson taught by rewards, responsibilities, and everyday interactions.

However, one must not equate learning to accept something with learning that it is morally right. Social learning can sustain admirable practices, harmful practices, and mixed ones. Its explanatory power does not supply the missing ethical judgment.

Resistance to moral change can also arise from commitments embedded in social life.

In "Notes on the Concept of Commitment," Becker explains continuity through "side bets": interests that become connected to maintaining a course of action. Participation in organizations can accumulate commitments that make departure costly. His argument therefore directs attention beyond a person's stated attachment to an activity toward the other things placed at stake by abandoning it. Becker, "Notes on the Concept of Commitment"

Applied to moral change, this suggests that revising a conviction may have consequences beyond the immediate issue. A person might risk relationships, occupational standing, group membership, or a valued understanding of their own past.

This is my extension of Becker's argument, rather than a moral-reform program formulated in that article. It nevertheless identifies an important obstacle to purely argumentative theories of moral change. Someone can recognize the force of a criticism while finding public revision personally costly.

The practical implication is conditional: where these commitments matter, opportunities for people to revise their positions without unnecessary humiliation may affect whether reform succeeds. Conversely, campaigns that treat every hesitation as proof of bad character may increase the social costs of reconsideration.

A further analytical distinction follows. Persistence can reflect sincere conviction, accumulated dependence, or some combination of both. Observing continued adherence alone does not establish which explanation is correct.

Becker's contribution to moral revision becomes most explicit when he discusses knowledge and ethical judgment.

In the 1973 reconsideration of labeling, he argues that factual inquiry and ethical evaluation interact. Facts do not logically generate obligations by themselves, but judgments depend on beliefs about how people and institutions work. He also insists that evidence must be capable of defeating the researcher's preferred interpretation. Becker, *Outsiders*, expanded edition, pp. 199–204

The significance can be developed through a simple analytical example.

Suppose someone accepts the moral principle that avoidable suffering should be reduced. They also believe that a particular punitive policy reduces suffering. Their support for that policy rests on both a normative commitment and an empirical proposition.

If reliable inquiry undermines the empirical proposition, the person may have reason to revise the policy judgment while retaining the original moral commitment.

This is a form of conscious moral change that requires no conversion to an entirely different ultimate value. What changes is the understanding of what the value requires under actual conditions.

The same structure applies to judgments about responsibility. If an outcome previously attributed to individual irresponsibility is found to depend heavily on institutional arrangements, the appropriate object of reform may change. Conversely, evidence might reveal that an institutional explanation has neglected consequential individual choices.

The process can be formulated as follows:

Element of moral judgment	Question that inquiry can clarify
Identification of conduct	What actually happened?
Attribution of responsibility	Who had knowledge, influence, or alternatives?
Assessment of consequences	What harms and benefits followed?
Evaluation of intervention	What did the proposed remedy accomplish?
Comparison of alternatives	Could the same moral aim be pursued differently?

This is an analytical reconstruction of the relationship between evidence and moral judgment. The normative commitments remain open to philosophical argument; empirical investigation clarifies the situations to which they are applied.

Becker and Horowitz's 1972 discussion is relevant here because it examines the relationship between research, political commitments, and criticism of established institutions. Its importance lies in taking the possible transformative effects of inquiry seriously without making political desirability a substitute for investigation. Becker and Horowitz, "Radical Politics and Sociological Research"

The question of whose testimony receives credibility is central to this process.

In "Whose Side Are We On?" Becker analyzes the tendency to privilege accounts supplied by people higher in an institutional hierarchy. Research that takes subordinate perspectives seriously can consequently appear biased simply because it challenges customary assumptions about who is entitled to describe an institution. Yet Becker also requires researchers to confront unwelcome evidence and delimit what their studies establish. Becker, "Whose Side Are We On?"

For moral change, the implication is that widening the range of credible testimony may alter the perceived character of a practice. An institutional routine described by its administrators as beneficial may appear different when investigated through the experiences of those subjected to it.

But the appropriate inference is to examine both accounts. It is not to assume that subordinate status guarantees accuracy.

Martyn Hammersley's analysis is especially useful on this point. He challenges the familiar interpretation that Becker simply endorsed partisan research or abandoned objectivity. Taking perspectives seriously and recognizing the influence of values do not authorize distortion of findings. Hammersley, "Which Side Was Becker On?"

This supplies a demanding condition for reflexive moral reform: those who expose other people's unexamined assumptions must remain willing to investigate their own.

For the analogy between conscious moral change and legislation, Becker provides a productive approach but leaves important distinctions intact.

The following comparison is my analytical extension of his framework:

Question	Deliberately enacted legal rule	Deliberately promoted moral norm
Who introduces it?	An institution or actor with a recognized procedural role	Potentially any organized group or influential participant
What establishes its standing?	Recognized legal procedures within a particular order	Recognition by relevant audiences, which may remain divided
How is it maintained?	Institutional interpretation and enforcement	Approval, disapproval, education, organizational procedures, and sometimes law
How is it withdrawn?	Repeal, amendment, invalidation, or other recognized mechanisms	Abandonment, reinterpretation, competing judgments, or loss of support

Question	Deliberately enacted legal rule	Deliberately promoted moral norm
What may persist afterward?	Social practices and attitudes surrounding the former rule	Stigma, habits, loyalties, and institutions associated with the former norm

The comparison highlights the possibility of **partial institutionalization of morality**. A professional code, disciplinary procedure, or organizational ethics policy can make a moral demand explicit and assign responsibility for applying it. This resembles legislation in important respects.

Yet no single organization ordinarily possesses a universally accepted competence to determine public morality as a whole. Different communities can continue to recognize different obligations. A norm may be withdrawn in one institutional setting while retaining force elsewhere.

For research, it is therefore useful to distinguish at least four outcomes: adoption of a formulation, establishment of an enforcement procedure, alteration of behavior, and internal acceptance of the judgment. Evidence of one should not automatically be presented as evidence of the others.

This also prevents an overly simple evolutionary account. Moral regulation does not necessarily move in one direction from diffuse custom to centralized codification. Codification can coexist with informal enforcement, provoke resistance, or eventually lose practical force.

Several limitations prevent Becker's approach from becoming a complete theory of morally justified transformation.

First, it cannot settle substantive ethical disputes merely by explaining their social organization. The discovery that a reform required campaigning tells us little, by itself, about whether the reform was right. Institutional success is compatible with moral improvement, moral deterioration, or a mixture of both.

Second, explanations of classification do not exhaust explanations of conduct. Understanding how an activity becomes condemned does not fully explain why people engage in it. Nor does examining a campaign's organization fully explain the origins of its participants' convictions.

Third, the approach requires a more developed account of structural inequality when applied to large transformations. It is necessary to investigate ownership, political authority, organizational resources, and access to communication, rather than merely assert that some actors have greater power.

Alexander Liazos's 1972 critique of the sociology of deviance pressed this issue sharply. He argued that attention to conventionally stigmatized groups could leave harmful activities of powerful actors insufficiently examined, while declarations about the importance of power exceeded its substantive investigation. His criticism concerned the field more broadly, but it identifies a serious test for applications of Becker's perspective. Liazos, "The Poverty of the Sociology of Deviance"

Fourth, conscious initiation does not entail comprehensive control. A campaign encounters other organizations, resistant audiences, practical constraints, and unforeseen consequences. Its final effects cannot be read backward as the execution of a single coherent intention.

Fifth, counter-movements must be included. People initially classified as outsiders can contest the classification and become reformers themselves. Later scholarship has developed this possibility explicitly: Jonas Lindblom and Kerstin Jacobsson describe activists who combine characteristics of stigmatized outsiders with attempts to transform prevailing norms. Lindblom and Jacobsson, "A Deviance Perspective on Social Movements"

These limitations suggest a demanding way to use Becker in evaluating an actual attempt at conscious moral change. One would investigate the proposed norm, its justification, the coalition advancing it, the authority gained by its interpreters, the procedures used to apply it, the experiences of those affected, and the consequences over time.

The final evaluative question would then remain open: **after examining those processes and consequences, what reasons support retaining, revising, or rejecting the norm?** Becker's distinctive contribution is to make that judgment better informed—and to ensure that the makers and administrators of morality are themselves included among the objects of inquiry.

****

John Rawls

John Rawls's contribution to the theory of conscious moral change is a theory of how people can critically revise their moral judgments, reform their institutions, and develop a more durable commitment to justice—while respecting one another's freedom to hold different religious and moral beliefs. His work therefore addresses both the possibility of deliberate moral transformation and the limits within which that transformation can legitimately be pursued.

An initial qualification is essential. Rawls did not develop a general technique for replacing a population's inherited morality with a new moral code. His principal concern was justice: the terms on which citizens live together, exercise political power, and distribute the benefits and burdens of social cooperation. Reading him as a theorist of conscious moral change requires connecting several parts of his philosophy that he did not present as one unified programme of moral transformation.

We should distinguish five questions within that reconstruction:

- How can inherited moral judgments be critically reconsidered?
- What principles should guide changes to social institutions?
- How do people acquire the motivation to act justly?
- How can citizens with different worldviews develop shared political commitments?
- What limits prevent moral reform from becoming coercive indoctrination?

Rawls's answers developed over several decades. His early methodological work began with "Outline of a Decision Procedure for Ethics" in 1951; "The Sense of Justice" appeared in 1963; *A Theory of Justice* in 1971 integrated justification, institutions, and moral development. His later writings, culminating in *Political Liberalism* in 1993 and "The Idea of Public Reason Revisited" in 1997, reconsidered how shared justice could survive enduring moral disagreement. *Justice as Fairness: A Restatement* in 2001 clarified the institutional implications. This development matters because the later Rawls places stronger limits on the ambition to unify society around one comprehensive moral philosophy. Rawls bibliography.

Rawls's starting point is that **moral judgment can be subjected to disciplined reflection without being reduced to either inherited custom or the commands of an authority**. His 1951 essay asks whether a reasonable procedure can adjudicate competing interests and whether that procedure can itself be rationally defended. Moral philosophy consequently has a constructive task: it should help us determine what we have reason to accept, rather than merely describe prevailing convictions. Rawls, "Outline of a Decision Procedure for Ethics".

Applied to conscious moral change, this establishes an important distinction between the *social existence* of a norm and its *justification*. A community may have observed a rule for centuries. That fact explains its authority within the community, but does not settle whether it deserves continued allegiance. Conversely, the novelty of a proposed norm gives it no automatic superiority. Deliberate reform must provide reasons for replacing the old judgment.

The central method for such reconsideration is **reflective equilibrium**. We begin with considered judgments: convictions formed under conditions in which ignorance, fear, self-interest, and other distortions have been reduced. We then examine how these judgments fit general principles and relevant background beliefs. Where they conflict, revision can occur at any level. A particular judgment may be abandoned, a principle reformulated, or an assumption about society corrected. In its wider form, the process also considers competing moral theories and arguments against one's initial position. Carl Knight, "Reflective Equilibrium".

Consider an illustrative case. Someone believes that public offices should be open to equally qualified applicants, but also believes that members of a historically dominant religious community should receive preference. Reflection exposes a tension. The person must explain why religious membership is relevant to public qualification, reconsider the preference, or modify the general commitment to equal opportunity.

The example shows why reflective equilibrium can generate moral change without beginning from a demand to discard an entire tradition. A commitment already accepted in one area can become a reason to revise an inherited exception elsewhere. The change may be substantial even when its initial justification comes from familiar moral language.

But reflective equilibrium also permits revision of apparently attractive general principles. It does not require concrete experience to surrender whenever it conflicts with an elegant theory. John Mikhail's analysis emphasizes that Rawls's method includes an attempt to understand our moral capacities and considered judgments; treating it simply as the imposition of abstract principles misses part of its original function. Mikhail, "Rawls' Concept of Reflective Equilibrium and Its Original Function in *A Theory of Justice*".

There is nevertheless a difficulty. A coherent set of beliefs may still reflect a restricted social experience. If everyone consulted shares the same prejudices, improving consistency can make those prejudices more systematic. My assessment is that reflective equilibrium becomes a stronger instrument of reform when the range of experiences and objections considered is broad. Logical consistency alone cannot substitute for learning how institutions affect people whose lives differ from one's own.

Rawls's **original position** provides a further means of gaining critical distance. It asks which principles representatives would choose under fair conditions, behind a veil of ignorance that conceals the particular social positions, talents, and comprehensive commitments of those they represent. General knowledge about social life remains available. The situation is hypothetical: it represents constraints on justification rather than describing an actual constitutional meeting. Its purpose is to prevent bargaining advantages rooted in existing inequalities from determining the principles of justice. Leif Wenar, "John Rawls," discussion of the original position.

Its significance for moral reform is easiest to see through an example. A person who supports unequal educational opportunities because their own children benefit is asked to reconsider the arrangement without relying on knowledge that they belong to the favoured group. The

intellectual operation is deliberate: suspend the advantage that makes the existing order attractive, and ask whether the rule remains defensible.

This does not mean that ignorance itself produces morality. The situation has already been designed to embody moral ideas about equality and fairness. Its critical force depends on the defensibility of those ideas. Rawls therefore cannot avoid substantive moral argument by appealing to an apparently neutral procedure.

That point also clarifies his use of **constructivism**, explicitly developed in his 1980 Dewey Lectures. Constructing a conception of justice means articulating principles through an appropriately justified practical standpoint. It does not mean that a government or intellectual group can make any chosen norm morally valid simply by announcing it. The conditions of construction themselves require justification. Rawls, "Kantian Constructivism in Moral Theory".

The direction of reform is specified by Rawls's principles of justice. In paraphrase, they require:

1. **Equal basic liberties:** everyone must enjoy an adequate, compatible scheme of fundamental freedoms.
2. **Fair equality of opportunity:** access to offices and positions must involve substantive opportunities, rather than merely the formal absence of prohibitions.
3. **The difference principle:** permissible social and economic inequalities must work to the greatest advantage of the least advantaged.

The first principle has priority over the second, and fair equality of opportunity has priority over the difference principle. Consequently, greater prosperity cannot by itself justify sacrificing basic liberties. In his later interpretation, these principles belong to a political conception addressed to the basic institutions of constitutional democracy. Rawls, "Justice as Fairness: Political not Metaphysical".

These principles make Rawls's contribution more specific than a general recommendation to become fairer. They supply standards by which a community can examine its inherited arrangements. An educational system may satisfy formal equality while allowing family wealth to determine effective access. A prosperous economy may improve average living standards while retaining avoidable disadvantages for those at the bottom. A popular reform may remain objectionable because it violates fundamental liberties.

The resulting moral transformation concerns the language in which advantages are justified. Institutional benefits must be defended to those who occupy less favourable positions. The question becomes whether the rules establish fair cooperation, including for people who cannot rely on wealth, inherited status, or political influence.

Rawls also identifies two moral powers: the capacity to understand and act from principles of justice, and the capacity to form, revise, and pursue a conception of the good. Basic liberties protect the development and exercise of these powers. Their importance helps explain why a programme that secures conformity by disabling independent judgment would conflict with his account of citizenship. Rawls, "The Basic Liberties and Their Priority".

For conscious moral change, this introduces a demanding condition: **the people whose convictions are being changed must remain participants in evaluating the change**.

Reformers cannot regard their superior confidence as sufficient authority to determine everyone else's moral life. A defensible reform should strengthen people's ability to understand the terms under which they live.

The institutional dimension is equally central. Rawls's "basic structure" comprises the major institutions whose combined operation assigns rights, duties, opportunities, and social advantages. His reason for concentrating on it includes the cumulative effects of institutional arrangements: individually permissible transactions can gradually produce unjust background conditions. Justice therefore requires attention to the system within which choices occur. Rawls, "The Basic Structure as Subject".

This gives us an important inference about moral change. Appeals to personal virtue may have little effect when everyday institutions reward contrary conduct. If appointments depend on patronage, citizens repeatedly learn that loyalty to powerful people matters more than equal qualification. Introducing credible, transparent procedures could alter both behaviour and expectations. Over time, conduct once regarded as an unavoidable necessity might become recognizable as an abuse.

That is an illustrative application of Rawls's framework, rather than a historical case he investigated. It shows how deliberate reform can operate through the social conditions in which moral judgments are formed. People learn from what institutions actually do, including whether declared principles govern the behaviour of those with authority.

Rawls's economic proposals make the scale of this institutional ambition clear. In *Justice as Fairness: A Restatement*, he distinguishes property-owning democracy and liberal socialism from welfare-state capitalism. He argues that generous assistance can coexist with excessive concentrations of productive wealth and political power. Property-owning democracy seeks a wide distribution of productive assets and educational capacities; liberal socialism is another possible institutional realization of his principles. He does not settle the choice independently of historical circumstances. Rawls, *Justice as Fairness: A Restatement*, §§41–42.

The moral implication is that changing the relationship between citizens may require changing their material independence. A society can proclaim equal respect while organizing daily life around dependency and domination. Conversely, institutional arrangements that make cooperation less humiliating may give citizens stronger reasons to regard one another as partners.

Rawls's account of **moral development** explains how such reasons might acquire motivational force. "The Sense of Justice" presents a hypothetical development through three relationships:

- Attachment to caring authorities makes their expectations morally significant.
- Participation in associations develops trust, role understanding, and willingness to reciprocate.
- Recognition of just principles extends concern beyond particular people and groups.

The movement joins affection, social experience, and understanding. Rawls also recognizes that emotions can outlast the judgments that originally supported them: someone may retain uneasiness about an activity after rejecting the religious prohibition under which they were raised. Rawls, "The Sense of Justice," §§III–VI.

This last observation is especially relevant to conscious moral change. Revising a judgment and altering the emotional habits attached to it are different processes. A person may accept equality intellectually while continuing to experience inherited discomfort in particular relationships. Conversely, emotional identification with others may precede a fully articulated principle of equality.

A plausible programme of moral education must therefore address understanding and experience together. Merely supplying a new vocabulary does not establish that the underlying dispositions have changed. Nor does producing an emotional response establish that the judgment associated with it is justified.

In *A Theory of Justice*, Rawls explicitly rejects an account of moral education as the installation of compulsive mechanisms by authorities. Instruction should provide reasons appropriate to developing understanding. Yet the developmental discussion has a crucial limitation: it principally investigates how a sense of justice arises when just institutions are already established. Rawls also denies that developmental psychology can by itself establish the philosophical superiority of his particular conception of justice. Rawls, *A Theory of Justice*, §§69–76 and §78.

Thus, there are two separate tests. A moral conception must be defensible, and people must be capable of acquiring a stable commitment to it. Psychological effectiveness cannot substitute for justification; philosophical attractiveness cannot substitute for a plausible account of motivation.

The gap between moral reproduction and moral transformation remains significant. Hannes Kuch's study identifies the difficulty sharply: much of Rawls's account explains how one newcomer develops a sense of justice within a surrounding society already organized around it. That leaves collective transformation less fully explained. Kuch reconstructs resources for a more reciprocal account in which changing institutions and changing motivations develop together. Kuch, "Difficulties in Nurturing a Sense of Justice".

This is a limitation of scope rather than a simple logical contradiction. A theory may legitimately explain how a just order reproduces itself without explaining every route by which it could emerge. Nevertheless, for the particular question of consciously creating moral change, the route from unjust conditions to more just ones is indispensable.

Rawls's later political philosophy addresses part of that problem through its treatment of pluralism. He came to regard the stability argument in *A Theory of Justice* as requiring revision: a free society could not reasonably be expected to converge on the same comprehensive philosophical account. Different religious, moral, and philosophical doctrines would persist. His later project asks how citizens can share political principles despite these differences. David Reidy, "John Rawls," discussion of *Political Liberalism*.

The important transformation here concerns what people demand of one another. They may continue to believe that their own doctrine is true while recognizing limits on the use of collective coercion to enforce it. A shared political morality can emerge without agreement about salvation, the ultimate meaning of life, or the best form of personal fulfilment. This establishes a limit on moral reform itself. The fact that a group regards its worldview as more enlightened does not, within Rawls's framework, authorize it to make that worldview

compulsory. The relevant political question is whether coercive arrangements can be justified to citizens who remain free and equal despite their doctrinal disagreements.

The idea of an **overlapping consensus** explains how common political commitments might nevertheless become morally substantial. Different doctrines can support the same political conception for different reasons. Such support is stronger than a temporary accommodation to an unfavourable balance of power: people continue to endorse fair terms even when they could gain from abandoning them. Rawls also insists that a political conception should not simply be fashioned to legitimate the outcome preferred by a dominant group. Rawls, "The Domain of the Political and Overlapping Consensus".

For illustration, one citizen might defend religious freedom because sincere faith requires freedom from coercion; another because individual conscience deserves protection; a third because experience demonstrates the dangers of religious persecution. They need not resolve their disagreements about religion to defend a shared freedom.

The potential moral change is considerable. Someone may move from tolerating others only while politically weak to recognizing that others possess a claim to liberty even when their own group becomes dominant. The external conduct might initially look similar, but its moral basis has changed.

Rawls's 1987 account offers an explicit mechanism through which this could happen. Successful cooperation can generate trust and independent attachment to political principles. Because comprehensive doctrines are often incomplete and loosely organized, citizens may subsequently revise parts of them to accommodate those political commitments. Experience can also demonstrate that stable cooperation does not require religious unanimity. Rawls presents this as a plausible possibility, not an inevitable historical law. Rawls, "The Idea of an Overlapping Consensus," §§VI–VII.

In *Political Liberalism*, the account becomes more differentiated. A *modus vivendi* rests on circumstances and interests. A *constitutional consensus* establishes allegiance to certain liberties and democratic procedures without agreement on their deeper foundations. An *overlapping consensus* provides more extensive and deeper moral support. Political actors seeking support beyond their own communities have incentives to develop arguments others can accept. Rawls characterizes the supporting psychology as philosophical rather than as a deduction from experimental findings. Rawls, *Political Liberalism*, Lecture II §8 and Lecture IV §§6–7.

This is perhaps his most direct contribution to a theory of consciously facilitated moral change. Reform need not wait until everyone already possesses the appropriate convictions. Fair arrangements may initially receive imperfect support; their successful operation may then help deepen that support.

However, the mechanism can also fail. If institutions proclaim reciprocity while distributing burdens selectively, participation may deepen distrust. The practical implication is that reformers must evaluate the credibility of institutional experience. Public instruction cannot reliably compensate for repeated encounters with injustice.

Education has a definite but limited role. Rawls supports teaching constitutional rights, the legal freedom to leave a religious community, capacities for social cooperation, and the political

virtues needed for citizenship. He distinguishes these purposes from educating everyone into a comprehensive liberal ideal of personal autonomy. At the same time, he acknowledges that civic education can have wider effects on beliefs and that liberal institutions cannot be neutral in their effects on every way of life. Rawls, "The Priority of Right and Ideas of the Good," §VI.

This produces a genuine tension. Enabling children to understand their rights may weaken inherited forms of authority even when weakening those traditions is not the educational aim. A defensible account must acknowledge that consequence. Rawls's distinction concerns the justification and purpose of public education; it does not promise that civic formation will leave every surrounding belief untouched.

Education also occurs outside schools. Steinar Bøyum's interpretation highlights Rawls's idea that a publicly available political conception can educate through political culture and institutions. Citizens encounter a representation of themselves as people entitled to fair terms, and this representation can influence how they interpret their social position. Bøyum, "Rawls's Notion of the Political Conception as Educator".

Public reason adds a moral discipline to the exercise of political power. When fundamental political questions are at stake, officials and citizens should offer reasons consistent with reciprocity among free and equal people. Rawls's later formulation permits religious and other comprehensive arguments, provided sufficient public reasons are supplied in due course. It accommodates a family of reasonable political conceptions, rather than requiring adherence to his own theory. The associated duty of civility is moral, not legally enforceable, and does not govern all discussion in civil society or the media. Rawls, "The Idea of Public Reason Revisited".

The intended change is partly a change in political character. Citizens learn to distinguish the reasons that move them personally from the reasons sufficient to justify coercive rules to others. They must also accept that losing a reasonable political disagreement does not necessarily mean being subjected to illegitimate power. Applied to moral reform, this discourages treating disagreement as evidence of moral inferiority. It requires an effort to explain why a proposed institutional change respects those who oppose it. That requirement can constrain the language, ambition, and means of reform even when reformers remain convinced of their goal.

Rawls also provides a place for deliberate moral criticism originating among citizens through **civil disobedience**. His 1969 discussion concerns reasonably just constitutional democracies and interprets justified disobedience as a political appeal to the majority's sense of justice. Its relevance is that existing authorities need not be the sole initiators of moral learning. Those subject to injustice can expose a conflict between publicly affirmed principles and established practices. Rawls, "The Justification of Civil Disobedience," in *Collected Papers*.

The reforming argument can therefore be internal to a society's commitments: citizens demand that principles already professed be applied more consistently. Such an appeal can change how the public understands those principles, not merely how it evaluates one disputed law. Rawls's restricted setting should nevertheless be remembered; this discussion does not supply a complete theory of resistance under every kind of regime. The relationship between this project and the conscious creation of positive law can now be made more precise. The following comparison is an analytical reconstruction:

Question	Deliberate legal change	Moral change within a Rawlsian framework
How is a proposal formulated?	Through legal drafting and institutional deliberation	Through reflection, argument, and public justification
What gives it authority?	Recognized procedures confer legal validity	Reasons must support its moral acceptability
How is it implemented?	Enactment, administration, and adjudication	Institutional reform, education, participation, and persuasion
Can acceptance be required immediately?	Conduct may become legally obligatory on a specified date	Moral conviction cannot be created by declaring an effective date
How is an old norm displaced?	Repeal or replacement changes its legal status	Criticism and changing experience may weaken its authority gradually
What remains open to criticism?	Both the law and its justification	Both the proposed principle and the means used to spread it

Rawls thus provides resources for the explicit formulation and institutional embodiment of morality, but no equivalent of a sovereign moral legislature. A valid statute may remain unjust. A majority may remain mistaken. A philosopher's argument may be influential without conferring authority to govern others. This also means that changing public morality and changing the law cannot be assumed to move together. Legal reform may precede conviction; moral criticism may precede legal change; either may proceed without securing the other. A serious theory must examine these relationships rather than treating enactment as evidence that moral transformation has occurred.

Several criticisms identify where Rawls's contribution needs supplementation.

Habermas raises the issue of democratic authorship. His 1995 critique questions whether Rawls assigns too much work to the philosophical construction of principles and insufficient work to citizens' continuing exercise of political autonomy. Applied to moral change, the concern is that citizens should participate in determining their shared norms, rather than merely receive conclusions already worked out by philosophy. Rawls disputed this interpretation of his position. Habermas, "Reconciliation Through the Public Use of Reason".

The practical issue extends beyond their disagreement. A reform can possess an elaborate public justification while still being socially experienced as something imposed by a narrow professional group. The adequacy of the argument and the inclusiveness of the process are related but distinguishable questions.

Susan Moller Okin exposes the importance of justice within families. Rawls relies on family relationships in explaining moral development, yet his early account insufficiently examines their gender hierarchy. Children who regularly observe unequal authority and unequal domestic burdens may learn lessons that conflict with public equality. Okin develops this challenge partly through Rawls's own principles. Okin, *Justice, Gender, and the Family*, chapter 5.

Her objection shows why civic education cannot be assessed only through its stated curriculum. Everyday relationships can confirm, qualify, or undermine the principles taught explicitly. Moral reform therefore requires attention to the places where expectations are learned before people encounter formal political argument.

G. A. Cohen challenges the division between just institutions and personal economic choices. If productive citizens demand substantial rewards before contributing, the supposed necessity of incentives partly reflects their motivations. Cohen argues that distributive justice also requires an egalitarian ethos in everyday conduct. This does not establish that Rawls ignored personal morality; it challenges the restricted application of his distributive principles. Cohen, "Where the Action Is: On the Site of Distributive Justice".

For conscious moral change, the disagreement concerns how deeply reform must penetrate personal motivation. Institutional rules may constrain conduct, but some forms of cooperation depend on what people voluntarily demand, refuse, or contribute within those rules.

Charles W. Mills challenges idealization that obscures actual domination. Principles for a just society do not automatically identify appropriate remedies for societies shaped by slavery, racial hierarchy, and exclusion. Mills argues that historical structures and the distortions they create must enter the analysis. His criticism does not reject moral ideals or abstraction as such. Mills, "'Ideal Theory' as Ideology".

This objection bears directly on the starting point of reform. Distrust may be a rational response to past and continuing treatment. Asking people to adopt the attitudes appropriate to fair cooperation, before establishing credible changes in their circumstances, can mistake the consequences of injustice for deficiencies in the people experiencing it.

Taken as an object of analysis, a proposed Rawlsian reform would therefore need to pass several distinct tests. Its principles would require justification; its institutional design would need to express those principles; its educational practices would need to respect independent judgment; and its operation would need to provide credible grounds for trust. These are my reconstruction of the demands arising from the preceding arguments, rather than a programme Rawls explicitly enumerated.

An example makes their independence clear. Suppose a government introduces equal-access rules for public employment. The legal rule may be defensible, but its moral effects depend on whether selection procedures actually follow it, whether excluded groups can obtain the necessary preparation, whether decisions can be challenged, and whether officials regard equal access as a public obligation. Declaring the principle addresses only part of the task.

Rawls's framework also changes how success should be evaluated. Immediate obedience, favourable opinion surveys, or repetition of approved language would not by themselves demonstrate moral improvement. The stronger evidence would concern whether people understand the principle, can question its application, recognize its claims when personally disadvantaged, and have adequate grounds for believing that others will reciprocate.

The deepest issue is whether conscious moral reform can create the conditions for its own free endorsement. Rawls supplies a demanding account of what such endorsement would involve and why institutions matter to its development. He offers a less complete explanation of how reformers can begin that process amid entrenched domination, unequal influence, and justified distrust. That remaining problem requires historical and sociological inquiry alongside the philosophical work of justification.

****

Jürgen Habermas

Jürgen Habermas's central contribution to the question of conscious moral change is a theory of how people can collectively reconsider the norms governing their lives while remaining the authors of those norms. Moral reform becomes legitimate when those affected can challenge existing rules, articulate their experiences, examine proposed alternatives, and participate as equals in their justification.

This is a reconstruction of the implications of his work for your question. Habermas did not present a single, self-contained programme called "the conscious transformation of morality." His account emerges from the connection between his theories of communicative action, moral development, discourse ethics, the public sphere, and democratic law.

The crucial distinction is between **successfully changing people's moral beliefs** and **giving them good reasons to change those beliefs**. A campaign can achieve the first through authority, social pressure, selective information, or manipulation. Habermas's distinctive concern is the second—and the social conditions under which reasoned reconsideration can become effective.

His theory therefore addresses several different questions that should remain separate:

Dimension	Central question
Moral justification	Is the proposed norm acceptable from the standpoint of everyone affected?
Moral learning	How do people acquire the capacity to reconsider inherited obligations?
Social acceptance	How does a revised norm become established in everyday expectations?
Democratic legitimacy	When may a political community make a rule legally binding?
Moral motivation	Why would people act on an accepted norm when doing so is inconvenient?

These distinctions explain both the ambition and the limitations of his approach.

The theory developed through a succession of problems rather than through the elaboration of one original blueprint. In *The Structural Transformation of the Public Sphere*—published in German in 1962—Habermas investigated the emergence of public criticism outside established political authority. His writings of the 1970s connected social development with changes in normative consciousness. *The Theory of Communicative Action* (1981) supplied the broad account of communication and social integration. *Moral Consciousness and Communicative Action* (1983; English translation 1990) developed discourse ethics; *Justification and Application* refined it; and *Between Facts and Norms* (1992; English translation 1996) reconstructed its relationship with democratic institutions.

A significant development within this sequence was his increasingly explicit distinction between universally valid morality, particular ethical commitments, and legitimate law. Consequently, formulations from different periods cannot always be treated as interchangeable. Finlayson and Rees, "Jürgen Habermas."

The intellectual background combines several traditions. Kant supplies the demand for impartial justification; Hegel highlights the social conditions of moral life; Marx raises questions about domination and institutional transformation; Mead explains the formation of the self through interaction; and Piaget and Kohlberg provide developmental accounts of increasingly differentiated judgment. Karl-Otto Apel was especially important in developing discourse ethics. Habermas, however, adopted a more fallible, reconstructive approach than Apel's stronger project of ultimate justification. Pieter Duvenhage, "Reflections on Habermas's Discourse Ethics."

For the problem of deliberate moral change, this combination matters because it joins two questions often studied separately: how norms can be justified and how people become capable of questioning them.

The foundation is Habermas's distinction between communicative and strategic action. In strategic interaction, participants seek to influence one another in pursuit of their objectives. In communicative interaction, they coordinate their actions through an understanding they can mutually accept.

Communication involves claims that others may question: whether a statement is true, whether an action or expectation is normatively right, and whether a speaker is sincere. When an accepted norm becomes controversial, participants can suspend their ordinary reliance on it and examine its justification through argument. Habermas calls this reflective examination *discourse*. It is a particular form of communication, rather than a synonym for any conversation. Habermas, *The Theory of Communicative Action*, volume 1.

Consider a workplace in which employees have traditionally been expected to remain available after working hours. The expectation may function without explicit justification until someone asks whether it unfairly disadvantages employees with caring responsibilities.

At that point, several distinct responses become possible. Management could threaten dissenters, offer incentives for compliance, invoke tradition, or explain why the expectation is necessary and submit that explanation to criticism.

The Habermasian significance of the last response is that employees are recognized as people entitled to assess the rule. The transition to conscious moral reconsideration occurs when a previously unquestioned expectation becomes a claim requiring justification.

This also shows why communication alone is an inadequate indicator of moral progress. A highly interactive campaign can remain strategic if its organizers have already fixed the desired conclusion and regard participants' objections merely as obstacles to persuasion.

Habermas's developmental theory explains how people become capable of this reflective stance. Drawing on Kohlberg, he distinguishes broadly between preconventional, conventional, and postconventional forms of moral judgment.

At a preconventional level, judgment is organized largely around consequences for oneself, sanctions, and exchanges. Conventional judgment takes established roles and the maintenance of a recognized social order as authoritative. Postconventional judgment makes those arrangements themselves open to assessment through more general principles. His interest is in changes in the structure of justification, rather than simply the replacement of one approved opinion by another. "Jürgen Habermas," *Internet Encyclopedia of Philosophy*.

For example, the same person might defend equal treatment for three different reasons: because discrimination attracts punishment; because equality is an established constitutional commitment; or because unequal treatment cannot be justified to the people disadvantaged by it. Identical outward conduct can therefore express different forms of moral understanding.

This distinction places an important qualification on conscious moral reform. Teaching people a new list of approved answers does not necessarily develop their moral autonomy. A population may abandon an inherited convention only to adopt another convention with the same unquestioning dependence on authority.

Habermas's writings on social evolution extend the learning idea beyond individuals. Societies can institutionalize more demanding forms of justification, although individual development and historical change are not identical processes. His reconstruction of historical materialism gives moral-practical learning a significance that cannot be reduced to advances in productive technology. Habermas, *Communication and the Evolution of Society*.

Analytically, this suggests that a society's capacity for moral change depends partly on whether its institutions preserve opportunities to question inherited expectations. It does not establish that every later historical arrangement is morally superior. Nor does a person's proficiency in discussing universal principles establish superior character: sophisticated moral reasoning can coexist with selfish conduct.

Discourse ethics supplies the central test for proposed moral changes. Two principles are particularly important, conventionally designated **D** and **U**. The following are explanatory paraphrases.

The discourse principle, in its familiar moral formulation, holds that a norm can claim validity only if everyone affected could accept it as a participant in practical discourse.

The universalization principle requires consideration of the foreseeable consequences and side effects of everyone's following the norm. Those consequences must be acceptable to each affected person, with alternatives also taken into account.

D identifies the connection between validity and discursive acceptability; U specifies the impartial perspective needed in moral argument. Habermas, *Moral Consciousness and Communicative Action*.

These principles have several implications for deliberate reform.

First, the reformer must consider the burdens imposed on particular people. An aggregate improvement cannot by itself settle whether a norm is justified. A rule might benefit most participants while exposing a minority to a burden they have strong reasons to reject.

Second, affected people must have an opportunity to correct descriptions of their own interests. An observer's sympathetic imagination is valuable, but it cannot reliably substitute for hearing experiences that the observer does not share.

Third, the reformer's position must remain open to revision. A consultation whose outcome cannot be changed by participants' reasons would fail the very test being invoked.

Fourth, a norm's popularity and its moral validity remain different questions. Agreement reached under intimidation, misinformation, or systematic exclusion is not sufficient evidence of justification.

The relevant procedural expectations include opportunities to question claims, introduce reasons, express needs, and challenge exclusions without coercion. Discourse ethics consequently evaluates the conditions under which agreement is reached, as well as the agreement itself. It does not provide a formula from which philosophers can mechanically deduce a complete substantive moral code. Michael Hemmingsen, "Jürgen Habermas and Discourse Ethics."

There is nevertheless a difficulty in the expression "could accept." Actual assent by everyone is rarely available. Future people cannot participate; some affected individuals cannot articulate their interests; and even extensive discussion may remain inconclusive.

The distinction between actual assent and justified acceptability is therefore indispensable, but it creates a danger. An advocate might claim that all properly rational participants would agree, then dismiss actual dissent as evidence of irrationality.

A defensible application of Habermas's approach must keep such claims contestable. The thought that others *would* agree cannot give an advocate immunity from answering the objections they actually raise.

Habermas seeks a foundation for these requirements in the presuppositions of argumentation. Someone who offers an argument presents it as deserving assessment through reasons. Practices of justification therefore carry expectations that cannot simply be replaced by force without changing the character of the activity.

His later account also acknowledges that the move from these expectations to universal morality requires more than identifying conversational rules. It involves a reconstruction of moral understanding and historical learning. In his discussion in *The Inclusion of the Other*, he expressly distinguishes argumentative presuppositions from obligations governing conduct outside discourse. Habermas, "Communicative Ethics," excerpt from *The Inclusion of the Other*.

The philosophical issue is substantial. Why should someone's participation in reason-giving establish obligations toward everyone affected, including people outside the discussion? Habermas offers an account of the normative commitments involved in impartial justification. Whether that account establishes the full authority of universal morality remains disputed.

For conscious moral change, its practical value is clearest as a discipline of justification: whoever seeks to alter others' obligations must provide reasons that take their standpoint seriously.

Habermas also limits moral reform by distinguishing morality from other kinds of practical judgment. In his mature terminology, "ethical" questions concern the good life and the identity of a person or community; "moral" questions concern what everyone owes to everyone else.

Kind of question	Example	Relevant consideration
Pragmatic	Which arrangement will achieve an agreed objective?	Effectiveness and consequences
Ethical	What kind of community do we want to become?	Identity, commitments, and a shared interpretation of the good
Moral	What can we justifiably require of everyone affected?	Impartial consideration and reciprocal acceptability

A community can rationally discuss its traditions without every conclusion becoming universally obligatory. Habermas's distinction allows ethical diversity while preserving criticism of injustice. Habermas, *Justification and Application*.

This distinction is especially useful when evaluating campaigns that present a preferred way of life as a moral necessity.

For example, a community might value intense professional achievement, religious devotion, artistic cultivation, or continuity between generations. Members may deliberate intelligently about these commitments. Yet the argument that a practice expresses their identity differs from the argument that everyone must adopt it.

Conversely, calling a practice "our tradition" does not answer an objection that it unjustifiably subordinates some members. Ethical identity is relevant, but it cannot automatically terminate moral inquiry.

Another distinction concerns justification and application. Establishing a norm's general validity does not settle every case in which it might apply. Drawing on Klaus Günther, Habermas recognizes a further task: determining which valid norm is appropriate in a situation

whose relevant features must be adequately understood. Klaus Günther, *The Sense of Appropriateness*.

Suppose a school recognizes both equal treatment and the need to accommodate disability. Discovering that an identical examination format unfairly disadvantages some pupils need not mean abandoning equality. It may require a more adequate interpretation of equality in that situation.

Consequently, what appears publicly as "a new morality" may involve different intellectual operations: rejecting a norm, widening its scope, correcting its application, or recognizing a previously neglected conflict between obligations. These should not be collapsed into one undifferentiated process of replacement.

The relationship between these ideas can be reconstructed as a model of deliberate moral change. The following sequence is an analytical synthesis, rather than a numbered programme written by Habermas.

1. **An experience becomes a grievance.** People identify something as unfair, humiliating, exclusionary, or inconsistent with an acknowledged commitment.
2. **The grievance receives public expression.** An experience previously treated as personal misfortune becomes a claim that others can assess.
3. **An inherited norm becomes questionable.** Its authority no longer suffices to settle the dispute.
4. **Participants examine alternative arrangements.** They assess reasons, consequences, interests, and the distribution of burdens.
5. **A revised understanding gains support.** Some participants change their judgments; others may continue to object.
6. **The result enters practices or institutions.** Organizational rules, educational practices, public expectations, or law may change.
7. **Experience exposes further problems.** New objections or unintended effects reopen the discussion.

This model identifies several points at which an apparent success can be misleading. Public attention may increase without improving understanding. A rule may be amended while everyday conduct persists. Compliance may increase while resentment deepens. A reform may also solve one injustice while producing another.

The contribution of the model is therefore partly diagnostic: it helps identify *what kind* of change has occurred and which further claims remain unproven.

Habermas's public-sphere theory addresses the movement from private experience to public criticism. The public sphere consists of communication through which citizens discuss matters of common concern and scrutinize political authority. It depends on freedoms of expression, association, and access to information; it cannot be identified simply with state institutions or with a crowd of individuals. Habermas, "The Public Sphere: An Encyclopedia Article."

For moral transformation, the significance is that those outside official decision-making can bring problems into view. A parliamentary agenda may omit an issue because the prevailing vocabulary does not yet recognize it as a public injustice.

Associations, social movements, investigative reporting, and public testimony can help change that situation. Their contribution includes discovering questions that existing institutions have failed to ask. Their claims still require examination: distance from government does not confer moral infallibility.

Habermas's treatment of civil disobedience gives this critical activity a sharper form. He regards publicly accountable challenges to law as potentially important to the self-correction of a constitutional democracy. The state's existing legal order does not exhaust the standards by which its legitimacy can be judged. Habermas, "Civil Disobedience: Litmus Test for the Democratic Constitutional State."

This makes moral change possible within an established political order without assuming that every legitimate challenge must already have official permission.

The relationship between moral criticism and law is where Habermas most directly addresses the possibility of consciously reorganizing normative life. His Tanner Lectures, *Law and Morality*, examine the transition from sacred and customary foundations toward deliberately enacted, alterable law.

That transition produces a problem: once law can be changed by decision, its authority requires justification beyond the fact that someone had the power to enact it. Habermas connects legitimacy with procedures that institutionalize practical argument while preserving its critical force.

He also recognizes why moral reasoning cannot perform all the work of law. Societies require decisions within limited periods, coordinated implementation, and reliable compliance. Law supplies institutional capacities and sanctions that moral conviction alone cannot guarantee. Habermas, *Law and Morality*.

In *Between Facts and Norms*, law and morality become more clearly differentiated. The general discourse principle is specified differently for moral justification and democratic legislation. Political decisions involve pragmatic judgments, ethical interpretations, moral considerations, and compromises.

The associated account of rights and democracy is often expressed through the co-originality of private and public autonomy: citizens require protected freedoms to participate independently, and democratic participation to determine the concrete form of their common legal order. Habermas, *Between Facts and Norms*.

This prevents an uncomplicated identification of democratic lawmaking with moral legislation.

A legislature may legitimately choose among several permissible arrangements. It may settle a budget, adopt an administrative scheme, or broker a fair compromise without demonstrating that its decision is the uniquely correct moral answer.

Habermas's "Three Normative Models of Democracy" connects informal public discussion with institutional decision-making. Public communication can generate influence, which democratic procedures translate into legitimate political power. Yet public opinion does not itself administer society. The process includes different forms of reasoning and bargaining,

rather than requiring one encompassing ethical consensus. Habermas, "Three Normative Models of Democracy."

The implications for the idea of a "positivization of morality" are particularly revealing. That expression can mean at least three things:

Possible meaning	Assessment from a Habermasian perspective
Making inherited norms explicit and open to revision	A central feature of reflective moral life
Creating institutions in which norms can be examined and reformulated	An important condition of legitimate collective action
Giving an authority power to determine moral validity by enactment	Incompatible with keeping validity open to reasoned criticism

A parliament can establish when a statute enters into force. A professional association can amend its code. A university can repeal a disciplinary regulation. Those acts create or remove institutional obligations according to their respective procedures.

Their moral justification remains a separate question.

An association could unanimously adopt a morally objectionable rule; a dissenter might have the strongest justification. Conversely, a morally compelling principle might lack recognition in every existing institutional code.

Thus, Habermas helps explain how morality can become more explicit, organized, and reflexive. His approach does not yield an institution with authority to settle moral rightness simply by issuing a decision.

This has consequences for philosophers, constitutional courts, expert committees, and reform movements. Their expertise or institutional position may give them important responsibilities, but it cannot make their own interpretation equivalent to the assent of everyone affected.

An institution invoking universal moral principles must still explain its mandate, its interpretation, the interests it has considered, and the objections it has answered. This is my analytical implication of the theory: **a claim to speak for universal reason increases the burden of justification.**

The remaining difficulty is motivation. People may understand and even accept an argument while failing to act on it.

Habermas acknowledges this gap. His account of communicative morality does not establish an automatic passage from justified judgment to conduct. Other motives can prevail; solidarity and social relationships matter to the practical realization of moral insight. Habermas, "Communicative Ethics."

His broader social theory also distinguishes the *lifeworld*—the background of shared meanings, relationships, and socialization—from systems coordinated through such mechanisms as money and administrative power. The concern about the colonization of the lifeworld arises

when those mechanisms displace forms of communication necessary for cultural reproduction, solidarity, and personal development. Habermas, *The Theory of Communicative Action*, volume 2.

An implication for education is that learning to justify norms requires experience of being heard, answering others, and accepting correction. A lesson on respectful discussion will have limited credibility in an institution where questioning authority predictably attracts humiliation.

An implication for organizations is similar. Asking employees to internalize a public-interest ethic while rewarding only competitive advantage creates conflicting practical expectations. Moral exhortation cannot by itself resolve that conflict.

These are applications of Habermasian reasoning, not evidence that he supplied a comprehensive educational curriculum or a tested programme for organizational reform.

A concrete hypothetical example brings the mechanisms together. Imagine a profession whose established culture treats reporting a colleague's misconduct as disloyal.

Initially, the accepted expectation may appear to protect mutual trust. A scandal then reveals that silence has exposed clients to serious harm.

A reflective reform would need to distinguish several issues:

- Whether loyalty to colleagues can justify concealing harm to outsiders.
- Whether allegations can be investigated fairly.
- Whether people reporting concerns face retaliation.
- Whether confidential advice and public accusation require different rules.
- Whether new procedures produce compliance through fear or become recognized as justified.

An expert committee could draft a revised code, but the code's existence would answer only part of the problem. Its justification would require hearing those exposed to harm, those accused, and those expected to report. Its effectiveness would depend on credible procedures. Its internalization would depend partly on whether experience confirms that the system treats people fairly.

The reform might ultimately redefine professional loyalty as responsibility for maintaining trustworthy practice. Here moral change involves both criticism and reinterpretation: an established value receives a different meaning through examination of its consequences and relationships.

Habermas's later writings add further qualifications. In "Religion in the Public Sphere," he treats religious and secular citizens as capable of mutual learning. Religious citizens must accommodate pluralism and the requirements of constitutional government, while secular citizens should remain open to possible insights expressed in religious traditions. Religious contributions can appear in informal public debate; generally accessible reasons are required for their authoritative institutional use. Habermas, "Religion in the Public Sphere."

For moral reform, this means that inherited religious language cannot simply be classified as exhausted or irrational before its claims are considered. Translation and reinterpretation can be forms of learning. Tradition supplies material for criticism as well as objects of criticism.

His essay on human dignity similarly connects normative development with experiences of humiliation and exclusion. Human rights acquire more determinate content through responses to injuries that reveal inadequacies in existing protections. Dignity links moral claims about equal standing with their legal articulation. Habermas, "The Concept of Human Dignity and the Realistic Utopia of Human Rights."

This helps explain how moral innovation can be both creative and continuous with existing commitments. People may discover that a principle they already acknowledge has implications they previously ignored.

His later work on digital communication introduces another qualification. Expanded opportunities to publish do not necessarily produce a shared environment for critical assessment. Habermas examines both the enlargement and fragmentation of the public sphere as participants become potential authors and established media lose their former position. Habermas, "Reflections and Hypotheses on a Further Structural Transformation of the Political Public Sphere," abstract reproduced by *Theory, Culture & Society*.

Applied to your question, this creates a distinction between the speed of normative diffusion and the quality of moral learning. A new judgment may spread rapidly through repetition, emotional identification, or reputational pressure while receiving little examination. Equally, digital communication may expose previously ignored experiences and make criticism possible. The medium does not settle which process predominates.

The strongest criticisms concern whether Habermas's demanding standards can guide actual moral transformation without concealing the inequalities within it.

Nancy Fraser argues that formally open participation does not neutralize social differences in authority, credibility, vocabulary, or access. Subordinated groups may need their own counterpublics to articulate experiences that dominant discussion excludes. The boundary between private matters and public concerns is itself politically contested. Fraser, "Rethinking the Public Sphere."

This criticism deepens the account of moral change. Before an argument can receive fair consideration, people may need to develop the language in which to express it. Public reasoning depends on earlier work of interpretation, association, and collective self-understanding.

Joel Anderson identifies a related issue: discourse presupposes capacities to interpret and articulate needs, yet those capacities are unequally developed and socially evaluated. What counts as competent self-expression can itself become a matter of dispute. Anderson, "Competent Need-Interpretation and Discourse Ethics."

The resulting problem is acute. Accepting every expressed preference at face value may overlook dependency or adaptation to deprivation. Allowing experts to determine people's "real" interests risks paternalism. A procedure centred on participation needs an account of how participation becomes substantively possible.

Amy Allen challenges the historical narratives of progress underlying Habermas and related critical theorists. Her concern is that Western modernity can become the assumed developmental standard against which other societies are judged. She seeks to retain emancipatory criticism while questioning that historical hierarchy. Allen, *The End of Progress*.

For deliberate moral change, the corresponding question is whether reformers have demonstrated the superiority of a proposal or merely identified it with the presumed direction of history. Describing a norm as "modern" cannot substitute for justification.

There are also difficulties internal to the theory. How should participants identify everyone affected? Which consequences are sufficiently foreseeable? What happens when informed, sincere participants continue to disagree? How can unequal argumentative abilities be accommodated without handing interpretive authority to specialists?

These questions do not disappear by invoking an ideal discourse. They become practical tasks for any institution claiming to approximate one.

Finally, the theory's justificatory strength should not be confused with a complete causal explanation. It offers criteria for assessing whether moral change is defensible and concepts for reconstructing learning. It does not by itself predict which innovations will prevail, how quickly they will spread, or how deeply people will internalize them. Those questions require empirical investigation of organizations, interests, socialization, media, and power.

Habermas's most distinctive achievement is therefore to place the legitimacy of deliberate moral change inside the relationship between reformers and those whose lives they seek to change. A reform's content matters, but so does whether people can scrutinize it, dispute its description of their interests, and influence its revision.

The demanding test is whether an initiative develops that capacity for judgment in the people it addresses. When it does, conscious moral change can become an exercise of collective autonomy. When it secures assent while insulating its premises and authorities from criticism, its success provides little evidence of the moral learning Habermas sought to explain.

****

Philip Kitcher

Philip Kitcher offers an explicit philosophical programme for making moral change more deliberate, better informed, and less dependent on suffering, political accident, and violent conflict. His central proposal is that societies should develop practices of collective moral inquiry: procedures through which people identify failures in their inherited norms, understand the experiences of those affected, consider alternatives, and revise their ways of living together.

He approaches morality as a human achievement that remains open to improvement. Consequently, the task of moral philosophy includes investigating **how people can improve morality itself**. In his own formulation of the project, understanding earlier advances should help make future progress more systematic and less costly. Kitcher, "A Method for Moral Progress."

This makes him particularly relevant to the question of whether moral norms can undergo something resembling the transition from inherited customary law to consciously revised law. However, his proposal centres on collective inquiry and institutional learning; it does not establish a sovereign moral legislator whose enactments automatically determine what is right.

The development of his position can be followed through these principal works:

Work	Contribution to the problem of conscious moral change
"Psychological Altruism, Evolutionary Origins, and Moral Rules" (1998)	Explains the psychological capacities and limitations from which moral practices could emerge.
The Ethical Project (2011)	Develops his historical, naturalistic account of ethics as an evolving human practice.
Moral Progress (2021), based on his 2019 Munich lectures	Presents his most concentrated methodology for identifying moral problems and improving collective responses.
The Main Enterprise of the World: Rethinking Education (2022; published online in 2021)	Connects moral development with education, citizenship, and the institutions supporting worthwhile lives.
The Rich and the Poor (2025)	Extends the programme into economic policy, poverty, inequality, and international cooperation.

The following analysis distinguishes his published proposals from the further institutional implications that can be drawn from them.

The starting point is the limited character of human responsiveness to other people. Kitcher's early work distinguishes biological altruism from psychological altruism. Biological altruism concerns effects on reproductive success; psychological altruism concerns an agent's motives. A person is psychologically altruistic when recognizing another person's needs changes what that person wants to do, with the other's benefit mattering in its own right.

Such responsiveness varies. Someone may care strongly about a few people, respond generously only in certain situations, or misunderstand what another person actually needs. These differences matter because moral development involves more than increasing the intensity of benevolent feelings: it can also involve widening their range and improving the understanding on which they depend. Kitcher explicitly rejects the straightforward derivation of moral prescriptions from evolutionary facts. Kitcher, "Psychological Altruism, Evolutionary Origins, and Moral Rules."

This provides a useful analytical distinction for deliberate reform. Consider an employer who treats familiar employees considerately but ignores workers employed by a distant subcontractor. The difficulty may concern the *range* of concern rather than an absence of benevolence. Alternatively, the employer might care about those workers but possess a mistaken picture of their conditions. That would be partly a failure of knowledge. A third possibility is that the employer recognizes their circumstances but consistently gives competitive advantage overriding importance.

These are different obstacles. Information alone might help with the second; the first requires widening attention; the third raises questions about motives, incentives, and institutional constraints. Kitcher's approach encourages this differentiated diagnosis.

In *The Ethical Project*, the central historical hypothesis is that early moral practices helped people manage recurring failures of altruistic responsiveness. Rules could stabilize cooperation when spontaneous sympathy proved unreliable. A capacity for **normative guidance** allowed individuals to check an impulse against a socially acquired directive before acting. Over time, injunctions, emotional dispositions, expectations, and practices of accountability could reinforce one another.

Kitcher calls this an analytical history. Parts of it are explicitly reconstructions of how morality *could* have developed, given the limited evidence concerning prehistoric life. It should therefore be read as a philosophically informed explanatory hypothesis, rather than an established account of one identifiable founding event. Kitcher, "Précis of The Ethical Project."

The significance for conscious change is substantial. If moral arrangements were historically developed responses to difficulties in social life, then their revision need not be understood as a departure from morality. Revision can be a continuation of the activity through which morality arose.

Nevertheless, historical origin cannot by itself justify present authority. Showing that a rule once promoted cooperation does not establish that it should govern contemporary society. Nor does showing that a practice survived establish that it was beneficial to everyone. Survival

might reflect coercion, exclusion, or advantages enjoyed by powerful groups. An adequate account therefore needs a criterion for distinguishing improvement from successful domination.

Kitcher's description of ethics as a "social technology" makes this problem especially clear. Moral practices help organize human interaction, and their historical development generates new possibilities alongside new difficulties. In his own account, religious interpretations could reinforce rules whose practical beginnings lay in social problem-solving; later moral reform could involve reconsidering those rules and their inherited justifications. He presents informed attention to consequences and an expansion of sympathy as central resources for continuing this process. Kitcher, "The Evolution of Ethics."

The technological analogy should be interpreted carefully. It does not imply that a society can be redesigned with the precision of a machine. Human beings interpret reforms, contest their purposes, alter their behaviour, and develop aspirations that reformers did not anticipate.

Its strongest implication is more modest: institutions and norms should be assessable in relation to the human problems they address. Their antiquity, sacred status, or familiarity cannot alone place them beyond examination.

For example, an inherited professional hierarchy might once have supported apprenticeship and the transmission of knowledge. Later, the same hierarchy might unnecessarily obstruct competent newcomers. Assessing it requires distinguishing the valuable function from a historically contingent arrangement that performs that function imperfectly. Reform might preserve mentoring while changing access, authority, or accountability.

Kitcher's normative position also develops beyond the prevention of conflict. His account includes a concern with the opportunities people have to lead worthwhile lives. In the 2011 book, this is associated with **dynamic consequentialism**: both the arrangements being evaluated and the understanding of what makes life better can develop through inquiry. His conception of improvement is consequently broader than maximizing a fixed quantity of pleasure or satisfying existing preferences without examining how those preferences arose. Ron Mallon's analysis of *The Ethical Project*.

This is an important feature of deliberate moral change. A society might initially pursue access to education because education increases productivity, then discover that intellectual participation, independence, and cultural enjoyment have value beyond their economic contribution. The reform has changed both opportunities and the public understanding of what opportunities matter.

That flexibility is also a source of difficulty. If the ends themselves are revisable, reformers cannot settle every dispute by applying a permanently fixed measure. They must explain why their revised conception deserves acceptance.

His most distinctive answer is to give explanatory priority to moral progress rather than to independently established moral truth. Kitcher proposes understanding progress as movement away from identifiable difficulties, without requiring a complete account of the final moral destination. His preferred analogy is problem-solving: improvement can be recognized even when no finished, ideal form of human life has been specified. Kitcher's account of his research programme at the Wissenschaftskolleg.

This distinction is sometimes misunderstood. A reform can have a definite objective—ending a particular exclusion, for example—without being one step toward a fully specified ultimate society. Rejecting a final destination does not mean abandoning purposes, standards, or judgment.

The attraction is that moral action need not wait for agreement on a complete ethical system. People who disagree about religion, happiness, or the best life may nevertheless recognize a particular institutional cruelty and cooperate in removing it.

The limitation is that identifying a problem already involves evaluation. A loss of status can be experienced as a serious grievance by someone whose status depended on others' subordination. The fact that someone dislikes a change cannot establish that the change is regressive. Kitcher therefore needs a disciplined way of examining complaints, rather than treating dissatisfaction itself as decisive.

In *Moral Progress*, he also considers a pragmatist account of moral truth: some judgments may endure through continued improvement in moral practice. Crucially, the methods used to assess those judgments can themselves be revised. His picture therefore includes development in conduct, habits, judgment, and collective practices, rather than limiting progress to the acquisition of correct propositions. Kitcher, "The Many Modes of Moral Progress."

An analytical consequence is that three questions must remain distinguishable:

- Was the proposed change reasonably justified with the information available?
- Did its implementation actually improve the situation?
- Would the judgment survive further experience and better inquiry?

A conscientious decision can produce harmful unforeseen effects. Conversely, an improvement can occur through fortunate circumstances without having been intelligently planned. Kitcher's programme seeks to improve the relationship between justification and outcome, but cannot eliminate the difference.

The 2021 book turns this philosophical position into a methodology of moral inquiry. Its central historical examples are the abolition of chattel slavery, the expansion of opportunities for women, and increasing acceptance of same-sex love. Kitcher uses these cases to ask how societies might identify and address moral failures earlier, more thoroughly, and with less dependence on exceptional courage and favourable circumstances. *Moral Progress*, publisher's description.

These examples matter because they show that public acceptance and moral adequacy can diverge. A practice can be customary, legally recognized, religiously defended, and personally internalized while remaining open to serious criticism.

They also suggest that conscious reform has at least two targets. It can change the answer society gives to a recognized question. More fundamentally, it can change which questions society is prepared to recognize at all.

A demand for access to an institution, for instance, may initially be dismissed as absurd because the excluded group is considered naturally unsuited to participation. Before arguments about fair access can proceed, that supposed unsuitability must itself become contestable.

Kitcher calls his approach **democratic contractualism**. Its organizing idea is a discussion in which affected perspectives are represented, factual claims are responsibly assessed, and participants try to arrive at arrangements others can live with. His methodological maxims are guides for inquiry rather than a mechanical algorithm. The following table compresses their orientation; it is not a reproduction of his numbered sequence. Kitcher, "Method in Moral Inquiry."

Requirement	Its role in deliberate moral change
Take grievances seriously	Bring neglected experiences into public consideration.
Establish who is affected	Include those affected by existing arrangements and proposed replacements.
Improve the factual basis	Examine beliefs about capacities, circumstances, causes, and consequences.
Develop mutual engagement	Encourage participants to understand the significance of alternatives for other lives.
Consider workable revisions	Search for responses that address the difficulty without merely relocating its burdens.
Continue scrutiny	Treat apparent solutions as open to correction through subsequent experience.

The demanding element is mutual engagement. It requires more than letting opponents speak before restating one's position. Participants must attempt to understand why an arrangement matters to others.

This changes the function of discussion. In ordinary bargaining, people may keep their preferences fixed and negotiate according to relative power. In moral inquiry, learning about other people can alter what participants consider desirable, acceptable, or worth defending.

That difference can be illustrated through a hypothetical workplace rule requiring unrestricted evening availability. Management initially treats availability as evidence of commitment. Employees with care responsibilities experience it as exclusion. A useful inquiry would examine whether evening work is genuinely necessary, how burdens are distributed, and whether other arrangements would meet the organization's needs.

A successful discussion might change the meaning of "commitment" itself. Reliability, quality of work, and cooperation could become more central than constant personal availability. The resulting moral change would concern both an explicit rule and the evaluative culture supporting it.

But the discussion would be inadequate if employees could describe their difficulties only at the risk of losing promotion. Formal inclusion is insufficient when the circumstances of participation make candour costly. This is a practical implication of Kitcher's requirements, and one reason institutional design matters as much as the willingness to hold a meeting.

Kitcher also treats the ideal conversation as a standard for improving actual deliberation. It is not necessary to assemble every affected person in one place. His discussion of climate deliberation includes representation of future generations and the development of institutions through which deliberative groups can communicate. In an interview, he regarded France's

Citizens' Convention on Climate as a promising beginning. Julia Hermann's account of her conversation with Kitcher.

Representation introduces further questions, however. Who can speak adequately for those absent? How can participants discover that an important perspective has been omitted? How should competing accounts of another group's interests be evaluated? These questions become especially difficult for future people, children, and nonhuman animals.

The problem of false consciousness makes Kitcher's programme more ambitious than a system for hearing complaints. Some people protest against restrictions but are ignored. Others may accept restrictive expectations because they have learned that certain lives are inappropriate for people like themselves.

Kitcher's discussion focuses particularly on socially distributed ideals of the self: which aspirations and forms of life are considered available to which groups. Deliberate moral inquiry must therefore examine inherited restrictions periodically, even when no organized complaint has emerged. When exclusion is defended through claims about a group's incapacity or unsuitability, those claims require their own investigation. Kitcher, "Problems of False Consciousness."

This is a powerful diagnosis, but also a dangerous point for any programme of moral reform. A reformer may too easily interpret agreement as evidence of enlightenment and disagreement as evidence of indoctrination.

To assess the difficulty, consider two people who choose the same traditional occupation. One has considered alternatives and prefers it. The other has never encountered a credible alternative and believes departure would be shameful. Their outwardly identical choices do not necessarily have the same significance.

Yet an observer cannot infer the second situation merely from the traditional character of the choice. The relevant investigation concerns access to information, practical opportunities, freedom from coercion, and the person's ability to reconsider.

A defensible institutional application would therefore expand opportunities for exploration while preserving the possibility that people will continue to prefer familiar lives. Otherwise, "liberation" becomes a requirement to adopt the reformer's preferred identity.

Kitcher's discussion of experiments in living addresses part of this problem. Where abstract arguments cannot settle whether alternative lives are possible or worthwhile, voluntary exploration can provide experience. He recognizes the ethical difficulties surrounding risk, effects on others, and disagreement over when an experiment has failed. These are especially important qualifications: his approach does not establish that every challenge to an inherited norm will succeed. *Moral Progress*, discussion of false consciousness, pp. 66–72.

My interpretation is that the strongest version of this proposal treats reform as creating conditions for informed agency. It gives people greater ability to test inherited assumptions without prejudging every choice they should subsequently make.

A further development concerns the relationship between individual beliefs and interlocking institutions. Rahel Jaeggi argues that moral problems arise within larger social

arrangements: work, family, economic organization, law, and cultural expectations can reinforce one another. The resulting restrictions may remain difficult to identify until conflicts or changes within that arrangement expose them. Her criticism is that an emphasis on understanding individual perspectives can understate these structural conditions. Jaeggi, "Progress as the Dynamics of Crisis."

The importance of this objection can be seen without assuming that people's experiences are irrelevant. Suppose a society formally opens a profession to everyone but retains training schedules, household expectations, and financial requirements that systematically make participation unavailable to some. Changing the admission rule alone leaves much of the obstacle intact.

Conversely, changing one institution may generate pressures elsewhere. An educational reform can alter employment aspirations; employment changes can alter domestic arrangements; domestic changes can expose previously unexamined expectations about responsibility.

Conscious moral change therefore requires attention to interactions among institutions. Otherwise, reformers risk interpreting predictable structural difficulties as personal failures to adopt the new morality.

Kitcher accepts that Jaeggi identifies a limitation in his initial treatment and proposes extending inquiry to examine the assumptions embedded in connected institutions. This is a significant concession: his methodology is presented as capable of development through criticism, not as an already complete account of social transformation.

Education supplies an important connection between occasional moral inquiry and a society's enduring capacity for reform. In *The Main Enterprise of the World*, Kitcher distinguishes preparation for earning a living, democratic citizenship, and personal fulfilment. He examines education together with the social and economic arrangements needed to make its benefits widely available. Kitcher, *The Main Enterprise of the World*.

An earlier essay makes the underlying distinction particularly clear: education involves acquiring information, developing habits and dispositions, and becoming capable of further learning. Reflective ethical judgment and reasoned exchange with fellow citizens belong within its purposes. Kitcher, "Education, Democracy, and Capitalism."

The implication for moral change is that a society needs more than citizens who know its current rules. It needs people capable of recognizing when those rules require reconsideration.

This creates an educational tension. Children cannot begin social life by independently reconstructing every moral expectation. They need trust, habits, examples, and initially accepted guidance. Yet education that makes inherited expectations immune to questioning undermines the capacity for later correction.

A plausible application of Kitcher's approach would combine the formation of dependable dispositions—honesty, attentiveness, cooperation—with increasing opportunities to examine their meaning and resolve conflicts among them. Respect for authority, for example, would need to coexist with the capacity to recognize abusive authority.

His educational outlook also extends beyond childhood. In a 2025 interview, he advocates opportunities for people to return to education throughout life and criticizes the reduction of education to economic productivity. He connects worthwhile lives with relationships, cooperation, and opportunities for development. Kitcher's interview with *El País*.

This matters because moral change affects people whose formative education occurred under earlier expectations. A society cannot rely exclusively on generational replacement if it hopes to reconsider urgent problems intelligently.

In *The Rich and the Poor*, the programme extends into public policy. Kitcher criticizes the displacement of ethical considerations from economic and political judgment and calls for renewed inquiry into poverty, aid, and collective responsibilities. Kitcher, *The Rich and the Poor*.

The conceptual point is that economic organization also shapes moral dispositions. Institutions may reward competition, normalize indifference, or make cooperation difficult. A programme that asks individuals to become more considerate while leaving those pressures unexamined is incomplete.

This does not mean that a particular economic policy follows automatically from democratic contractualism. It means that the purposes and human consequences of policy must themselves enter collective assessment. Economic expertise can inform that assessment, but selecting the ends cannot be reduced to a technical calculation.

The analogy with the conscious development of law has direct textual support in Kitcher. At the end of *Moral Progress*, he imagines a future observer comparing institutions of moral inquiry with the emergence of courts in Locke's account. The comparison concerns the possibility that organized consideration of competing claims could reduce reliance on superior force. He explicitly presents this as a hopeful imaginative exercise. Kitcher, *Moral Progress*, p. 168.

For the wider question of the deliberate revision of morality, I would distinguish three different kinds of institutionalization:

Form	What becomes organized	Relationship to Kitcher
Formulation of norms	Expectations become explicit and publicly discussable.	Compatible with his approach, but insufficient for assessing their merits.
Institutionalization of revision	Procedures identify problems, hear affected perspectives, consider alternatives, and review outcomes.	The strongest connection with his programme.
Authoritative moral enactment	A designated body's decision determines moral validity by virtue of its authority.	Goes beyond what his methodology establishes.

The distinction matters because legal and moral change have different relations to formal decision. A competent legislature can repeal a legal provision through a recognized procedure. A moral commission's declaration does not, by itself, remove a belief from people's consciences or make its replacement morally adequate.

This leaves at least four tasks for an institutional version of Kitcher's programme: justifying a proposed change, making it practically effective, cultivating the dispositions needed to sustain it, and preserving opportunities to challenge it.

A reform might succeed at one and fail at another. A well-argued recommendation may remain socially ineffective. An effectively enforced rule may remain poorly justified. A widely internalized expectation may later prove harmful.

For that reason, I would describe Kitcher's contribution as a theory of **the organized revisability of morality**. This is my characterization, rather than his technical terminology. Its distinctive aspiration is to make criticism, experimentation, and correction normal features of ethical life.

That aspiration also clarifies the role of philosophers. Their contribution can include conceptual clarification, identifying inconsistent assumptions, examining historical cases, and proposing procedures. None of these activities grants an automatic entitlement to determine everyone else's good.

The relationship to Dewey is especially close: both approaches connect ethics with inquiry into lived difficulties and the reconstruction of habits and institutions. Kitcher develops this inheritance through an evolutionary account of morality and a more explicit methodology for moral progress. His contractualism also has affinities with Rawls and Habermas, but his particular emphasis falls on improving historically situated processes of inquiry. These are differences of emphasis; Rawls and Habermas also address institutions and practical political questions.

The strongest criticism concerns power: people may understand a moral objection and still have compelling interests in resisting it. Amia Srinivasan argues that histories of emancipation cannot be adequately understood through conversation alone. They involve domination, organization, resistance, and changes in the distribution of power. She also questions whether Kitcher's historical argument establishes the inadequacy of moral realism. Srinivasan, "The Limits of Conversation."

The causal issue is fundamental. A dominant group may refuse discussion because it already understands what serious discussion would threaten. A reformer then needs an account of how the conditions for inquiry can be established.

Kitcher's response is **weak optimism**: improved institutions might reduce conflict and contingency in some cases. He expressly accepts the historical importance of violent struggles and does not claim that deliberation can eliminate every struggle over power. His proposal is to investigate how much improvement is possible, rather than assume either universal success or inevitable failure. *Moral Progress*, response to the commentaries, pp. 164–168.

Susan Neiman presses a related motivational objection. A method designed for honest inquiry works most readily among people already willing to inquire honestly. It also needs to account for the deliberate or unconscious suppression of knowledge that would threaten oppressive arrangements. Neiman, "Progress, Regress, and Power."

This exposes a possible circularity in implementation: successful moral inquiry requires dispositions that moral reform is partly intended to produce. Education can help, but it cannot

instantly supply trust, willingness, and mutual concern where political experience has destroyed them.

Frederic Kellogg adds that preferences and conduct are embedded in habits. Conversation may accompany transformation without being sufficient to cause it; conflict and practical experience can themselves generate new possibilities for learning. Kellogg, "Moral Progress: Conversation or Conflict?"

A second major criticism concerns the normative content of the procedure. Inclusion, mutual concern, and responsiveness to suffering are already moral commitments. They cannot all be presented as conclusions produced by a morally neutral discussion.

This does not necessarily invalidate Kitcher's project. A procedure can be valuable without deriving itself from an entirely neutral starting point. But it limits what the procedure proves. Its advocates must defend its commitments openly and explain why people should accept them.

A third problem is control over participation. Who determines whether someone is sufficiently informed or willing to engage? Restrictions intended to prevent domination may protect inquiry, but they may also exclude inconvenient disagreement. Stephen Macedo's review of *The Rich and the Poor* draws attention to this tension in Kitcher's discussion of selecting participants. Macedo's review.

Institutional safeguards would therefore have to address the selection of participants, the framing of questions, the choice of experts, and the possibility of appeal. These are my proposed requirements for developing the programme further; they should not be mistaken for a detailed constitutional design already supplied by Kitcher.

A fourth difficulty concerns moral realism. The absence of an identifiable moment of moral discovery does not demonstrate that there are no moral truths independent of the inquiry. A realist can agree that experience, testimony, protest, and institutional change improve moral understanding. Kitcher's practical methodology may consequently be more persuasive than the strongest version of his argument against rival metaethical positions.

Finally, the programme must confront **regression and mixed outcomes**. A reform can improve one dimension of life while worsening another. It can expand formal opportunities while creating new dependencies. It can reduce one exclusion while generating a different one. Recognizing these possibilities requires evaluating outcomes separately from the moral intentions of the reformers.

Kitcher's enduring contribution is to make the methods of moral revision a central philosophical subject. The practical value of that contribution will depend on whether institutions informed by it can hear previously neglected experiences, correct their own assumptions, and demonstrate improvements that remain open to scrutiny by those who must live with them.

****

Allen Buchanan and Russell Powell

Allen Buchanan and Russell Powell's central contribution is a **biocultural theory of moral progress**: an account of how evolved human psychology, social institutions, material conditions, and cultural innovation interact to make moral improvement possible. Its significance for conscious moral change lies in a practical proposition: **human beings can improve the conditions under which moral convictions develop, are critically examined, and become effective in social life.** Their main joint statement is *The Evolution of Moral Progress: A Biocultural Theory* (2018). Oxford University Press

Their project needs to be understood at two levels. They investigate how changes commonly regarded as moral advances became possible, despite apparently restrictive evolutionary inheritances. They also ask how knowledge of those processes might help sustain existing achievements and facilitate further improvements. They offer an explanatory framework and directions for reform; they do not provide a demonstrated technique for reliably producing whatever moral transformation a reformer desires.

For the question of consciously changing morality, the most important distinction is between **understanding how moral convictions can be changed, establishing that a proposed change is morally justified, and determining who has legitimate authority to promote it**. These are separate problems. Buchanan and Powell make their strongest contribution to the first, adopt substantive commitments relevant to the second, and leave significant questions concerning the third open.

Their joint research developed through several closely connected publications:

Work	Contribution to the project
"The Limits of Evolutionary Explanations of Morality and Their Implications for Moral Progress" (2015)	Challenges claims that evolved human nature places rigid limits on moral inclusion.
"Toward a Naturalistic Theory of Moral Progress" (2016)	Develops an explanation of the circumstances enabling inclusive morality.
"De-moralization as Emancipation" (2017)	Examines progress through abandoning unjustified moral restrictions.
The Evolution of Moral Progress (2018)	Integrates the conceptual, evolutionary, institutional, and historical arguments.

Buchanan's subsequent *Our Moral Fate: Evolution and the Escape from Tribalism* (2020) develops deliberate moral institutional design more explicitly. That later development should be distinguished from the jointly authored theory.

Their conception of progress begins with a plurality of possible improvements. They resist reducing morality to a single function, such as maintaining cooperation, or identifying all improvement with expanding sympathy. Their provisional typology includes ten dimensions:

Dimension	**What may improve**
Observance	People follow justified norms more reliably.
Concepts	Moral vocabulary becomes more adequate to experience.
Virtues	Interpretations of admirable character become more defensible.
Motivation	Moral considerations influence action more appropriately.
Reasoning	Judgments become more consistent and discriminating.
De-moralization	Unjustified prohibitions lose their authority.
Moralization	Previously tolerated wrongs become objects of criticism.
Standing and status	Beings receive appropriate moral consideration.
Conceptions of morality	Understanding of morality's nature develops.
Justice	Its subjects and objects are understood more adequately.

They distinguish improvement involving the exercise or development of moral capacities from a merely fortunate change in circumstances. A beneficial outcome need not demonstrate that anyone has become morally better. Their discussion of the typology

This distinction matters for evaluating deliberate reform. Imagine that an organization reduces discriminatory decisions by introducing anonymous initial assessments. The change might improve outcomes even if evaluators' prejudices persist. Alternatively, evaluators might come to understand why their previous judgments were unjust. Both developments matter, but they represent different achievements. A serious assessment should examine conduct, convictions, institutional procedures, and the effects on those subject to them separately.

Their conception is also **dynamic**. Our present understanding of improvement remains open to correction. Moral progress need not consist in approaching a completely specified final moral order. The authors additionally identify *meta-moral progress*: improvement in the means through which moral advances are pursued, including reductions in violence and other moral costs. This places the conduct of reformers themselves within the scope of moral evaluation. "A Pluralistic, Dynamic Conception of Moral Progress"

The analytical implication is substantial. A movement could promote a defensible substantive objective while using methods that damage other moral goods. Conversely, a society could improve its capacity to debate and revise norms before resolving its principal disagreements. The morality of change therefore includes both the proposed destination and the processes through which people attempt to reach it.

The evolutionary problem that motivates much of their work concerns the relationship between cooperation and exclusion. On an influential evolutionary account, human moral dispositions developed partly because they supported cooperation within groups competing with other

groups. Such an account can explain loyalty, mutual assistance, and punishment of free riders. It also appears to explain why concern is often selective.

Buchanan and Powell challenge the inference that this history fixes the permissible or feasible boundaries of future morality. Their 2015 argument emphasizes developments that an excessively restrictive account struggles to explain: commitments extending beyond kinship, group membership, reciprocal advantage, and the capacity to threaten others. An explanation of morality's origins must accommodate its subsequent transformations. Their 2015 article

The underlying methodological point can be stated independently: an account of why a capacity originally evolved does not, by itself, establish everything that the capacity can subsequently accomplish. Language, inference, imagination, and social learning can operate in circumstances very different from those in which their precursors developed. Whether a particular moral achievement is possible requires an account of those capacities and their present environment.

Their discussion targets two contrasting responses to evolutionary constraints. They call one *evoconservative*: the conclusion that ambitious inclusive moralities are unrealistic because human concern is inherently restricted. They call another *evoliberal*: acceptance of much the same restrictive psychology, combined with the proposal that biomedical intervention may be needed to overcome it. Their own account emphasizes the interaction between psychological plasticity and cultural change. These labels designate positions within their argument, rather than all conservatives or all liberals. The authors' 2019 précis

A particularly important distinction is between **strategic morality and subject-centered morality**. In the former, another person's claim on us depends fundamentally on participation in mutually advantageous relations. In the latter, an individual can possess moral significance independently of an ability to benefit or harm us. This distinction has an earlier foundation in Buchanan's individual work on justice. "Justice as Reciprocity versus Subject-Centered Justice"

Consider a person with profound disabilities who cannot reciprocate assistance. A society may provide support because doing so reassures other citizens, improves its reputation, or satisfies relatives. Those are possible strategic explanations. A different justification holds that the person's own interests supply reasons for protection. The observable conduct might initially be similar, but its normative basis differs.

That difference becomes consequential when strategic benefits disappear. If the person matters only through advantages to others, protection becomes vulnerable when those advantages diminish. If the person has standing independently, the disappearance of reciprocal benefits does not remove the underlying obligation. This illustrates why expanding cooperation and expanding moral recognition cannot simply be treated as the same development.

Buchanan and Powell's positive explanation centers on **adaptive plasticity**. Their hypothesis is that some evolved dispositions toward exclusion are conditional: their development and expression depend on environmental signals, especially signals of danger from outsiders. Under other circumstances, those mechanisms need not obstruct inclusive norms in the same way. Because human activities can alter relevant environments, evolutionary influences can become objects of indirect intervention. Their chapter on the naturalistic explanation of inclusion

This is a probabilistic developmental account, not a claim that one environmental change immediately switches everyone's morality. Habit, previous experience, institutions, and interpretation matter. Nor does it require supposing that natural selection directly produced a commitment to universal human rights. More general capacities can contribute to cultural achievements that were not themselves the original objects of selection.

To illustrate the practical logic, imagine two communities with similar initial distrust. In one, access to basic resources is insecure, public protection is unreliable, and each encounter with outsiders is interpreted as competition for survival. In the other, disputes are handled through credible procedures and cooperation does not routinely expose participants to exploitation.

An identical appeal to equal human worth may encounter very different obstacles in these settings. The thought experiment does not prove the authors' hypothesis. It shows what their framework asks reformers to investigate: the circumstances making a proposed norm credible, practicable, and stable.

Their account is consequently also a theory of **moral regression**. Actual deterioration is only one possible cause. People can come to believe that they face severe group threats even when their perceptions are distorted. Cognitive biases and defective practices of information transmission can help produce that outcome. Political manipulation can therefore undermine inclusive commitments without first producing an equivalent objective change in circumstances. "Naturalizing Moral Regression"

An implication for conscious reform is that moral education cannot be assessed only by the principles it teaches. One must also examine its descriptions of the world. People might sincerely accept a general prohibition on harming innocents while being systematically misinformed about who is innocent, who is dangerous, or who qualifies for protection. In that situation, adding another statement of the principle leaves the decisive error untouched.

This connects the joint theory with Buchanan's earlier **social moral epistemology**. That inquiry examines how social practices and institutions affect the formation, preservation, and transmission of beliefs relevant to moral judgment. Applied ethics, on his account, needs to investigate those arrangements if it is to guide action effectively. "Social Moral Epistemology"

The importance of this approach can be seen in a simple case. Suppose a professional organization formally endorses respect for those affected by its decisions, while its internal procedures systematically exclude their testimony. Members may deliberate conscientiously within an informational structure that prevents relevant experiences from being heard.

Reform would then require more than exhorting members to become respectful. It might require changing complaint procedures, the composition of review bodies, the circulation of information, and opportunities to challenge established interpretations. These are illustrative applications of the approach, rather than a detailed program prescribed by the authors.

The same reasoning applies reflexively to moral reformers. If a group claims privileged knowledge of what morality requires, its own patterns of authority, selection of evidence, and treatment of dissent become proper objects of investigation. Expertise about the causes of behavior cannot itself settle the moral validity of the behavior being promoted.

Their treatment of **de-moralization** is especially relevant to consciously revising inherited morality. The term means withdrawing a mistaken moral condemnation. It does not mean the general collapse of moral commitment. They describe unjustified restrictions as *surplus moral constraints*, which can operate through public sanctions and through internalized conscience. Removing external penalties may therefore leave significant constraints intact. A person may continue to experience guilt or shame after the surrounding institution has changed its rules. Their 2017 article

Their discussion of invalid norms also challenges the inference that a practice's survival through cultural evolution establishes its moral authority. They investigate how harmful restrictions can arise and persist, and how their hold can be weakened. The relevant category is *proper* de-moralization: abandoning a valid constraint would not become progress simply because it enlarged someone's freedom. Their chapter on invalid moral norms

This produces a more discriminating way of understanding moral change. There are at least three possible operations:

- Retain a justified norm and improve compliance.
- Remove an unjustified restriction.
- Establish a justified restriction where conduct was previously tolerated.

For example, rejecting condemnation of consensual relationships and strengthening condemnation of coercion can occur together. The result is a reorganization of moral judgment, rather than a uniform increase or decrease in moral regulation.

The difficult question concerns how to identify invalidity. A reformer cannot infer it from a norm's age, its religious or secular origin, its unpopularity among intellectuals, or the inconvenience it causes. Conversely, longevity cannot establish validity. The relevant inquiry concerns reasons, consequences, competing claims, and the relation between the norm and other practices. This is where a causal theory of change must be supplemented by normative argument.

Buchanan's earlier criticism of Philip Kitcher helps explain the role of **open-ended normativity** in this project. In "The Open-ended Normativity of the Ethical" (2012), Buchanan argues that explaining morality through its original role in reducing the social costs of failures of altruism does not adequately explain its subsequent commitments. Suppressing conflict and taking everyone's interests seriously are different objectives. He argues that an account of development needs a richer explanation of people's ability to reflect on and revise their norms. This criticism concerns Kitcher's *The Ethical Project*, rather than a blanket assessment of Kitcher's later writings. Buchanan's article

For deliberate moral change, the philosophical importance of this argument is that morality can make its own inherited functions objects of criticism. A rule may coordinate a group effectively while distributing burdens unjustly. A norm may promote solidarity while excluding people who deserve consideration. Understanding what a norm accomplishes is therefore indispensable, but it leaves open whether that accomplishment justifies the norm.

Human rights provide Buchanan and Powell with a major case through which to develop their account. Their analysis identifies several conceptual changes:

Change	Significance
Expanding the subjects of justice	More people are recognized as having claims, rather than merely receiving benevolence.
Expanding the matters assessed by justice	Institutions and structural arrangements become objects of moral scrutiny.
Moving some concerns from charity to justice	Assistance may become something owed.
Recognizing positive rights	Some claims require provision, as well as restraint.
Limiting forfeiture	Serious wrongdoing need not eliminate every basic protection.
Recognizing individuals independently of strategic usefulness	Moral status does not rest entirely on reciprocal advantage.

These are the authors' judgments about advances, including judgments whose scope remains contested. Their chapter on moral concepts and human rights

Their historical discussion treats human rights as a complex practice involving movements, organizations, public argument, and legal institutions. It also considers interruptions and reversals, rather than presenting abolition and twentieth-century human rights as an uninterrupted ascent. An important distinction concerns the conditions needed for an innovation's initial emergence and those needed for its preservation or revival. Cultural achievements can survive circumstances in which they might not first have developed. "Human Rights Naturalized"

Analytically, this suggests a reciprocal relationship between moral ideas and institutions. A moral claim can motivate the creation of an institution; the institution can then make the claim more familiar, more practicable, and more resistant to neglect. Yet institutional recognition is an imperfect indicator of moral acceptance. People can comply opportunistically, and organizations can adopt ambitious language without changing their conduct.

A convincing evaluation therefore needs several kinds of evidence: public endorsement, behavior, institutional enforcement, and the actual position of the people whose claims are supposedly recognized. Treating declarations alone as proof of deep moral transformation would obscure the problem the theory seeks to explain.

What, then, does their joint theory actually propose for **deliberately promoting change**? Their 2016 article identifies three broad tasks: reduce harsh conditions conducive to exclusion; prevent deterioration, including distorted perceptions of deterioration; and protect cultural achievements against efforts to dismantle them through manipulation of belief. Concrete directions include reducing infectious disease, improving physical security, increasing material resources, and building institutions for peaceful cooperation. They explicitly avoid authorizing an elite to direct humanity simply because it possesses a naturalistic theory of moral progress. Their practical discussion, pp. 1012–1014

To translate that framework into the language of institutional reform, I would reconstruct the task as follows. This is an analytical application, rather than a procedure they themselves present in this form.

First, specify the defect precisely. "Society should become more moral" is too indeterminate. Is the problem an unjustified prohibition, a failure to apply an accepted principle, a mistaken description of a group, or an institutional arrangement that makes justified conduct excessively costly? Different diagnoses require different interventions.

Second, identify what maintains the existing pattern. The obstacle may be conviction, but it may also be fear, dependence, professional incentives, informational isolation, or an expectation that others will punish deviation. A person can privately reject a norm while continuing to enforce it because unilateral departure is dangerous.

Third, distinguish persuasion from changes in practical opportunity. An argument may persuade someone that a practice is wrong without enabling that person to leave it. Conversely, an institutional reform may make better conduct possible before the person fully accepts its moral rationale. Investigating both processes helps avoid attributing every failure to deficient character.

Fourth, provide opportunities for correction. Reformers need ways to discover whether their diagnosis was mistaken, whether burdens have shifted to another group, or whether apparent success reflects concealed noncompliance. Including affected people can improve information, but their participation must have a meaningful connection to decisions.

Fifth, assess durability and moral costs. A reform that depends on exceptional prosperity, a charismatic advocate, or continuous coercion may be vulnerable. Its evaluation should include what happens during disagreement and stress, and whether its implementation damages other justified commitments.

This reconstruction brings out the theory's relevance to the transition from customary morality toward consciously revised norms. Nevertheless, the analogy with legislation has limits.

A legal system can specify procedures for enacting and repealing rules. Moral authority is more dispersed. A state can repeal a penalty, an association can amend its code, and a university can change disciplinary rules. None of those actions guarantees that the corresponding moral condemnation disappears from families, religious communities, friendships, or conscience.

It is consequently useful to distinguish three dimensions:

Dimension	What deliberate action can address
Formal regulation	Written rules, sanctions, institutional responsibilities
Social expectations	Approval, blame, reputation, inclusion and exclusion
Personal conviction	Conscience, reasons, emotions, moral identity

The institutionalization of a moral reform can affect all three, but they need not move together. A written norm may acquire official force without widespread acceptance. A widely accepted moral claim may lack an effective institutional expression. Internalized shame may outlast both formal and informal sanctions.

For this reason, I would interpret Buchanan and Powell as contributing to a theory of **reflexive moral revision**, in which communities examine and alter the conditions and content of their

morality. Describing their work as a complete theory of moral legislation would overstate what they establish.

There is also a fundamental difference between making a rule and making it morally valid. An institution's decision can establish that its members are subject to a rule within that organization. Whether imposing the rule is justified remains a further question. Conscious revision makes that question more visible; it does not eliminate it.

Their position on **biomedical moral enhancement** requires similar precision. They criticize arguments that move too quickly from evolutionary limitations to the supposed necessity of altering human biology. Their preferred emphasis is on cultural and institutional innovation informed by a realistic account of psychology. Their joint chapter on the evolution of moral enhancement

The 2018 book develops doubts about whether biomedical enhancement would be a necessary or effective principal route to progress. This is compatible with accepting some interventions in principle. The central issue is comparative: which interventions can reliably improve the relevant capacities and conduct, under what circumstances? The book's biomedical enhancement chapter

Ingmar Persson and Julian Savulescu challenge their assessment. Among their objections are that reducing hostility need not generate intrinsic concern for outsiders, that institutional commitments do not necessarily reveal adequate moral motivation, and that biomedical and social interventions may complement each other. They also stress the difficulty of securing the motivation needed to create better institutions in the first place. Persson and Savulescu's critique

The underlying analytical problem is that "more altruism" is an incomplete specification of improvement. One must ask toward whom altruism is directed, whether it responds to accurate beliefs, and how it relates to fairness, judgment, and competing responsibilities. Strengthening an emotion cannot settle those questions by itself.

Buchanan's later solo work takes the issue of deliberate design further. In his account of *Our Moral Fate*, he explicitly proposes that sufficiently developed knowledge of interactions between environments and the moral mind could enable people to shape predominant moralities and their own moral identities. He describes this ambition as moral institutional design. He also presents the later book as developing the earlier account's treatment of social knowledge, psychology, and ideology. Buchanan's précis

In a subsequent interview, he offers examples involving political arrangements that encourage compromise, resistance to informational isolation, and scrutiny of communications that activate perceptions of group threat. He also recognizes the difficulty of obtaining cooperation for reform under conditions of polarization. These are proposals for investigation, not established causal laws or a jointly authored constitutional blueprint. MIT Press interview

The theory's strengths become clearer when placed beside its unresolved problems.

One strength is its separation of moral aspiration from assumptions about feasibility. An attractive principle does not tell us whether people can sustain the conduct it requires.

Conversely, an inherited disposition does not establish impossibility. The space between those claims is where empirical investigation matters.

A second strength is its treatment of institutions as conditions of moral agency. Moral thought occurs within arrangements that distribute information, security, influence, and opportunities for action. This helps explain why intelligent and conscientious people can reproduce defective practices, and why reforms aimed exclusively at individual conviction can disappoint.

A third strength is its inclusion of regression and the removal of mistaken prohibitions. These prevent progress from being equated with the accumulation of ever more demanding norms. They also make maintenance and correction legitimate tasks of moral reform.

The principal difficulties can be stated more precisely.

First, **causal naturalism does not settle moral justification**. William J. FitzPatrick questions whether a scientific explanation of moral development warrants the stronger philosophical characterization of the overall theory as naturalistic. Kim Sterelny raises a related difficulty concerning the grounding of inclusive obligations. Their criticisms expose a gap between explaining why a commitment develops and establishing its authority. The 2019 symposium

My assessment is that this gap need not defeat the project, provided its scope is kept clear. A causal theory can be valuable while leaving its normative premises open to independent defense. Problems arise when scientific credibility is allowed to make contested moral judgments appear already settled.

Second, **favorable conditions may be hardest to create where inclusion is most urgently needed**. Hanno Sauer presses this objection and proposes greater attention to institutions that can secure morally desirable outcomes without requiring extensive changes in individual benevolence. His argument concerns whether institutions should mainly cultivate more inclusive psychology or reduce reliance on demanding psychological capacities. “Butchering Benevolence”

The challenge suggests a question of sequencing. Limited cooperation motivated by ordinary interests might sometimes create safer conditions, which then permit broader moral learning. But that sequence cannot simply be assumed. Particular reforms require evidence about how the initial institutional change can occur.

Third, **the evolutionary reconstruction remains contestable**. Andrés Segovia-Cuéllar and Lorenzo Del Savio argue that Buchanan and Powell retain problematic elements of a narrative that treats contemporary inclusive cooperation as a departure from a more fundamentally small-group human nature. Their criticism concerns the historical and developmental assumptions underlying that contrast. Their analysis of evolutionary mismatch theories

A distinction follows: evidence that present attitudes are sensitive to circumstances does not automatically confirm a particular account of prehistoric selection. The practical case for improving institutions may survive revisions to the evolutionary explanation.

Fourth, **influence and legitimacy remain different questions**. If education, media, and institutional procedures shape moral judgment, those controlling them possess significant

power. Calling their objective “progress” supplies no independent justification for exercising that power. Deliberate moral reform therefore needs an account of participation, accountability, dissent, and limits on authority.

Fifth, **partial improvements cannot automatically establish the superiority of an entire age or civilization**. A society may improve one dimension while worsening another. Buchanan and Powell themselves distinguish local judgments of progress from global assessments, whose justification is considerably harder. Their concluding discussion

For research on conscious moral change, I would therefore use their theory as a framework for investigating specific transformations: what was criticized, what reasons supported the criticism, which institutions maintained the earlier practice, which interventions changed it, and whether the resulting improvement endured. The most demanding extension would be to establish institutions that can promote such learning while allowing their own moral assumptions and authority to be challenged. That is where their explanatory achievement meets the still unfinished political theory of who may legitimately guide moral change.

****

Sally Haslanger

Sally Haslanger's central contribution to the problem of conscious moral change is a theory of how people can deliberately transform the social practices that shape their moral perceptions, identities, and possibilities for action. In her account, lasting reform involves changes in public meanings, everyday habits, material conditions, and institutional arrangements. Moral argument contributes to this process, but its effectiveness depends on the social world in which people encounter and act upon it.

Her approach is particularly relevant to a difficult question: **How can people recognize and change an unjust way of life when participation in that way of life has helped form their understanding of what is normal, reasonable, and good?**

Haslanger addresses this question through several connected projects: the critical reconstruction of concepts; an account of social practices and structures; a theory of ideology; an epistemology of collective criticism; and an increasingly developed account of practical social transformation. What follows reconstructs these connections and then assesses their strengths and unresolved problems.

Her theory developed through a movement from the analysis of social categories toward the practical conditions of transformation. Her early work examines how gender and race can be understood as socially constituted positions. The essays collected in *Resisting Reality: Social Construction and Social Critique* (2012) establish much of this framework. During the following decade, her attention increasingly turns to social meanings, structural explanation, ideology, and collective agency. More recent work considers complex social systems, participatory design, and the relationship between legal intervention and cultural change.

These developments should be read together. Her proposals about the meaning of social categories are one component of a larger inquiry into how societies reproduce themselves and how their reproduction can be redirected. Her own account of this research emphasizes the neglected domain between individual persons and the state: culture, practices, and social organization. Haslanger's research overview

The first methodological move is to ask what our concepts should enable us to understand and accomplish. Haslanger calls this an *ameliorative* approach. A concept can be evaluated partly by examining whether it helps us identify significant phenomena, explain them adequately, and pursue defensible practical purposes.

Three questions must therefore be distinguished:

Approach	Governing question
Conceptual inquiry	What do people ordinarily understand by this concept?
Descriptive inquiry	What kind or pattern does our classificatory practice actually track?
Ameliorative inquiry	What concept should we employ, given critically examined purposes?

These inquiries can inform one another. Existing usage supplies evidence, but does not automatically settle what conceptual resources a critical inquiry needs.

Haslanger also distinguishes **manifest concepts**, which express what we take ourselves to be classifying; **operative concepts**, reflected in how our practices actually classify; and **target concepts**, which we have reason to adopt. The discrepancies between them create opportunities for criticism. An institution may profess one understanding of membership while acting according to another. "What Good Are Our Intuitions? Philosophical Analysis and Social Kinds"

The significance for moral change is substantial. Consider, as an illustrative application, an organization that claims to reward "merit." Its official understanding might emphasize excellent work, while its decisions actually reward constant availability, personal familiarity, and conformity to an established professional style. A reformer must investigate both the decisions and the evaluative category organizing them. Otherwise, a demand for "more meritocracy" might reinforce precisely the practice under criticism.

This makes conceptual inquiry a possible part of moral reform. It also makes the reformer responsible for explaining why a proposed classification serves legitimate purposes.

Her early analyses of gender and race demonstrate this method. In "Gender and Race: (What) Are They? (What) Do We Want Them To Be?" (2000), Haslanger develops categories intended to illuminate hierarchical social relations. Her proposed account of women concerns systematic subordination associated with observed or imagined bodily characteristics interpreted as indicating a female reproductive role. Her account of racialization similarly concerns social positioning through bodily markers associated with presumed ancestry.

These are explanatory proposals directed toward particular critical purposes. They should not be reduced to the claim that every ordinary use of "woman" or "race" must have this meaning.

The intended practical consequence is to make relations of domination visible and contestable. Moreover, ending hierarchical gender relations does not straightforwardly entail eliminating every possible gender distinction. The early paper leaves open whether a just society might retain nonhierarchical genders and recognizes that reproductive differences can remain socially significant. "Gender and Race"

A general lesson can be drawn from this strategy: moral change sometimes requires reconsidering what appears to be a description of a person's intrinsic nature. A classification may actually describe a relationship produced and maintained by social arrangements. Recognizing that difference changes the questions we ask about responsibility, necessity, and possible alternatives.

Nevertheless, showing that something is socially constructed establishes neither that it is unjust nor that it is easily changed. Both conclusions require further argument.

Haslanger's later work recognizes that a concept useful for diagnosing injustice may be inadequate for organizing a better society. In "Going On, Not in the Same Way," she distinguishes the disruptive work of exposing deficiencies from the constructive work of making better possibilities available.

She also acknowledges limitations in her earlier proposals: their capacity to reveal hierarchy did not supply adequate replacement concepts, and she had insufficiently anticipated some linguistic and political harms. A conceptual tool directed against one form of oppression can reinforce another.

This acknowledgment matters for understanding her approach to deliberate moral change. Ameliorative inquiry is itself revisable. Its proposals must be evaluated in relation to their consequences for the people whose lives they organize. "Going On, Not in the Same Way"

A documented example of criticism within this project comes from Katharine Jenkins. Jenkins argues that Haslanger's account can marginalize some trans women and proposes distinguishing gender as social class from gender as identity. This illustrates how a shared emancipatory purpose can support conceptual plurality rather than agreement on one privileged definition. Jenkins, "Amelioration and Inclusion"

Conceptual improvement can involve either correcting our understanding or changing the conceptual content itself. Haslanger distinguishes *epistemic amelioration* from *semantic amelioration*. In the former, we come to understand the subject matter better; in the latter, the classificatory content changes.

The difficulty is continuity: if reform changes the meaning of a concept, are reformers still addressing the same question as their predecessors? Haslanger explores how the continuing purposes, history, and coordinating role of a concept can support continuity through change. Moral learning need not be understood as abandoning an old subject whenever an inherited classification is challenged. "How Not to Change the Subject"

Her subsequent work with Stephen Yablo develops the possibility that amelioration sometimes corrects an interpretation of a term rather than simply assigning it a new meaning. Their account also recognizes constraints: inherited usage, continuity of inquiry, and the practical usability of a proposed interpretation matter. Consequently, conceptual reform cannot be justified merely by announcing that a preferred definition would have desirable political effects. Haslanger and Yablo, "Amelioration as Course Correction"

The second major component of her theory concerns social practices. A practice, in her account, involves learned patterns of mutually responsive behavior through which people coordinate around things regarded as valuable or disvaluable. Practices include much more than consciously accepted rules. They involve skills, expectations, interpretations, bodily responses, and engagement with a material environment.

Teaching, parenting, hiring, cooking, and distributing care are examples. Each involves people responding to one another through socially acquired understandings of what they are doing.

Practices also constitute roles. A teacher's authority, a student's responsibilities, and the significance of classroom activities depend on an organized context. This means that moral

agency is partly formed within practices: people acquire the capacities through which they recognize situations and respond to them. "What Is a Social Practice?"

An implication is that moral reform must examine the organization of action itself. Someone may sincerely endorse an egalitarian principle while competently performing a role whose ordinary requirements reproduce inequality. The discrepancy can arise without conscious hypocrisy.

For example, a hypothetical supervisor may support equal opportunity while routinely assigning valuable projects to employees available for late evening meetings. The supervisor's personal conviction matters, but so do working hours, care obligations, promotion criteria, and the meaning attached to "commitment."

Material conditions are integral to this account. Haslanger challenges explanations that try to understand social systems entirely through individuals' attitudes or collective intentions. Transport networks, food systems, schools, and workplaces depend on physical arrangements whose properties affect what people can do.

Changing people's beliefs about accessibility does not by itself make a staircase accessible. Equally, rebuilding an entrance can alter who participates, who becomes visible, and what others come to expect. Cultural and material changes can reinforce one another.

Her materialism therefore involves attention to the physical conditions of social organization, alongside interpretation and agency. It does not reduce culture to an economically determined appearance. "Failures of Methodological Individualism: The Materiality of Social Systems"

Her later distinction between structures and systems explains why reforms can be absorbed without ending injustice. Structures concern networks of relations and positions; systems are concrete, historically situated processes in which those relations operate.

A social system can adapt to interventions while preserving an underlying pattern of advantage. Changing one rule may leave other mechanisms capable of reproducing a similar outcome. Conversely, because interactions are nonlinear, a comparatively small intervention can sometimes have extensive effects.

This account gives particular importance to feedback: practices shape people and environments that, in turn, sustain those practices. Structural injustice concerns objectionable relations and arrangements; systemic analysis additionally investigates the processes through which they are maintained and reproduced. "Systemic and Structural Injustice: Is There a Difference?"

The following diagram is a schematic reconstruction of that feedback relationship:

This helps clarify the meaning of *conscious* change in her framework. Reformers can investigate these relationships and intervene deliberately, but they cannot assume complete knowledge or control of the resulting process.

The third component is her account of ideology as something embedded in ordinary social competence. Haslanger uses *cultural technē* for the publicly available repertoire of meanings, symbols, scripts, and interpretive resources through which people organize their lives.

Such a repertoire becomes ideological, in her critical usage, when it helps sustain unjust social relations. Culture as such is indispensable; the task is to criticize particular ways in which it organizes perception and action.

This explains why ideology cannot be addressed solely by correcting explicitly false propositions. It may operate through what appears respectable, embarrassing, admirable, threatening, or simply obvious. A reform that leaves these meanings intact can encounter resistance even after formal rules have changed. "Culture and Critique"

In "Cognition as a Social Skill," she develops the related idea that people learn how to perceive and respond fluently within a social environment. They acquire practical orientations involving thought, emotion, and action.

This creates a tension: the very capacities that enable successful participation can make someone an effective participant in unjust practices. Ideology can also help produce the circumstances that seem to confirm it. Consequently, a criticism confined to whether people accurately describe existing circumstances may miss how those circumstances were made.

Yet culture is not completely unified or closed. People move among different practices, encounter competing interpretations, and improvise. These openings help explain how criticism remains possible despite social conditioning. "Cognition as a Social Skill"

Consider a constructed example. If a community systematically denies some members opportunities to acquire a skill, the resulting difference in performance may become evidence, within that community, that those members are unsuited to the activity. The initial exclusion produces a pattern that subsequently legitimizes exclusion.

The moral criticism must distinguish the observed pattern from the explanation offered for it. Otherwise, even accurate observations can be incorporated into a misleading account of what people are capable of becoming.

This leads to the central epistemological problem: how can criticism acquire justification from within an ideologically shaped society? Haslanger gives a significant role to collective consciousness raising. People compare experiences, challenge familiar explanations, develop alternative descriptions, and investigate whether apparent personal failures reveal recurring structural problems.

In "Political Epistemology and Social Critique," this is a process of inquiry. Testimony and dissatisfaction can initiate it, but claims must be examined through critical engagement and empirical understanding. Oppositional consciousness is not automatically warranted.

Its achievement can be a justified *pro tanto* claim: a morally significant claim that deserves consideration, without necessarily overriding every competing claim. Haslanger explicitly leaves unresolved the complete political problem of adjudicating conflicts among such claims.

She also questions whether reflective equilibrium conducted within an insufficiently challenged framework can escape the distortions already built into its starting judgments. This criticism concerns the conditions of inquiry, rather than the mere act of reasoning carefully. "Political Epistemology and Social Critique"

The distinction between experience and a critical standpoint is essential. A person may know something about an injustice through undergoing it while lacking a satisfactory account of its causes. Conversely, someone outside that position may possess relevant explanatory knowledge while overlooking aspects visible to those affected. Collective inquiry can bring these resources into contact.

This gives consciousness raising a demanding function. It must improve the interpretation of experience while remaining open to correction.

Her position does not give suffering, group identity, or political commitment automatic epistemic authority. In "Taking a Stand," Haslanger treats a critical standpoint as an achievement requiring a defensible relationship among experience, explanation, and evaluation.

Dissatisfaction can motivate criticism without validating every conclusion drawn from it. A movement may misidentify causes, suppress inconvenient evidence, or exclude people whose experiences challenge its preferred interpretation. Those are defects in its critical practice.

She also resists making complete reflective awareness the goal of emancipation. Everyday life necessarily depends on skills and routines that are not continually examined. The substantive concern is whether practices are harmful or unjust, and whether people have adequate ways to criticize and change them. "Taking a Stand: Second-Order Pathologies or First-Order Critique?"

Her normative position combines the possibility of moral knowledge with openness about how values become determinate. It would be misleading to describe her simply as a relativist, or to assign her without qualification to a strong metaphysical realism.

In "Culture and Critique," her commitment to moral truth is deliberately compatible with several metaethical approaches. Its principal purpose is to preserve the possibility of warranted judgments against injustice; it does not select one fully developed account of what makes moral judgments true. "Culture and Critique"

Her later discussion of situated values adds that the concrete content of justice and flourishing develops through historically particular practices. Multiple acceptable ways of living can exist. New circumstances can require new articulations of values, and practical experience can reveal possibilities unavailable to abstract reflection alone. This is a stronger claim than merely admitting that philosophers sometimes make mistakes about an already complete blueprint. "Situated Knowledge and Situated Values"

The philosophical tension is productive but difficult. If values acquire specificity through social life, criticism must explain how their development can be evaluated. Existing acceptance cannot be decisive, because accepted practices may be oppressive. Yet simply invoking a future moral order does not justify the reformer's authority either.

Haslanger's approach locates the work of justification in continuing inquiry into harms, possibilities, relationships, and competing claims. The resulting judgments can be objective without every detail of a just society being settled in advance.

Her recent engagement with Rahel Jaeggi makes the role of substantive moral judgment especially clear. In an English text accompanying a 2026 German publication, Haslanger resists restricting criticism to internal contradictions, dysfunction, or blocked learning.

An oppressive arrangement can warrant resistance because of its wrongness, even where its organization is relatively stable. Claims about how transformation occurs also require historical and empirical assessment: crisis cannot simply be assumed to be a necessary condition of progress.

She further distinguishes the correctness of criticism from permission to impose a remedy. Someone can identify a wrong accurately while intervening in an objectionable manner. Alliances between affected people and others complicate a simple division between criticism from "inside" and "outside."

These points strengthen her account of moral change by separating moral warrant, causal effectiveness, and legitimate intervention. None automatically supplies the others. "Social Transformation and Moral Critique," English version

The practical agents of transformation are often organized groups and social movements. Haslanger's attention to movements follows from the collective organization of the problem. Individual refusal can be costly, unintelligible to others, or easily neutralized.

Movements can alter public meanings, make private experiences communicable, establish new expectations, and coordinate departures from established conduct. They can also change material conditions. Her discussion includes both disruptive action, such as strikes, and ordinary interventions, such as reorganizing spaces or creating opportunities for groups to meet.

A further contribution is practical demonstration. A different arrangement becomes more credible when people have experienced it working. Collective activity can therefore create knowledge of alternatives while building the capacity to sustain them. "Social Movements and Social Change," lecture handout

This yields an important analytical distinction between **persuading people that a norm is wrong** and **making it possible for them to live under a different norm**. The two processes can reinforce each other, but neither should be assumed to follow automatically from the other.

For instance, a person may agree that an employment practice is unfair while remaining dependent on it for income. Organized reform can change the options available to that person. In this respect, moral change concerns the conditions under which convictions can become effective agency.

Participatory design gives this practical orientation a concrete form. In her account of co-design, affected communities contribute to identifying problems, determining priorities, and developing responses.

Her Amherst lecture discusses work with girls in Kenya involving reproductive health information and the production of reusable menstrual products. The intervention combines knowledge, material resources, practical skills, and altered possibilities for participation. She treats it as a potential contribution to broader transformation, while expressly acknowledging that the case does not establish that a local intervention will produce systemic change.

The distinction between a promising intervention and demonstrated transformation is part of the argument. "Situated Knowledge and Situated Values: Co-Designing Social Justice"

The broader implication is that reformers should expect to learn from the people with whom they work. Local participants are also interpreters of the problem and potential designers of its solution. This affects both the quality of the intervention and the distribution of authority within the reform process.

Education has a similarly double role. In "Reproducing Social Hierarchy (or Not!)," Haslanger argues that education transmits social know-how as well as explicit knowledge. It can prepare people to navigate established arrangements successfully, thereby helping reproduce those arrangements.

Her proposed response includes opportunities to develop critical standpoints through collective inquiry. Merely introducing a more diverse group of participants does not necessarily transform the assumptions under which everyone operates. The educational task includes developing capacities for recognizing and investigating those assumptions throughout life. "Reproducing Social Hierarchy (or Not!)"

Her earlier study of professional philosophy provides an institutional application. She examined representation, publication practices, disciplinary expectations, and professional climate, and recommended better evidence, greater accountability, challenges to exclusionary schemas, and organization. The proposed changes concern how a profession evaluates and reproduces itself.

Those historical observations should be read as an analysis of the period studied, rather than as current statistics about philosophy. Their continuing theoretical significance lies in the connection she draws between institutional procedures and the judgments through which professional merit is recognized. "Changing the Ideology and Culture of Philosophy: Not by Reason (Alone)"

Law remains an important instrument, but its effects depend on its interaction with these other processes. In "How Social Theory Identifies Levers for Social Change" (2025), Haslanger presents law as one force within the reproduction and transformation of complex social systems.

Law can alter material constraints and public meanings, as well as rights and obligations. Its coercive backing gives it substantial power. However, its effects depend on practices that interpret, implement, accommodate, or resist it.

The significance for conscious moral reform is that legislation requires a social account of how it will operate. A formally appropriate intervention can be ineffective when it leaves the supporting organization of conduct intact. "How Social Theory Identifies Levers for Social Change"

Her earlier discussion of dueling, drawing on Lawrence Lessig, illustrates the interaction. A prohibition could fail when refusing a challenge remained dishonorable. A sanction connecting participation in dueling with exclusion from public office could instead create a conflict within the honor system itself: the duty of public service supplied a culturally intelligible reason to refuse.

The example concerns legal influence through social meaning. It also suggests why reform requires attention to the reasons that participants can recognize within their existing relationships. "How to Change a Social Structure," author's draft

Her treatment of intersectionality complicates any proposal to improve the position of a group as though its members occupied the same situation. In "Social Systems and Intersectional Oppression," Haslanger emphasizes the interdependent production of social positions differentiated through race, gender, class, and other relations.

A reform that benefits some women can shift burdens onto other women. The analysis therefore has to follow concrete relations and consequences, rather than assume a uniform experience of gender oppression.

Her account also questions whether racism, patriarchy, and capitalism should always be modeled as separate interacting systems. A single integrated formation can exhibit several distinguishable patterns of domination. The practical consequence is attentiveness to effects elsewhere in the system when one relationship changes. "Social Systems and Intersectional Oppression"

The implications for individual responsibility are demanding, but they do not require an impossible exit from social life. Haslanger recognizes that people often cannot simply stop participating in compromised practices. Their responsibilities arise within relationships that constrain their options and provide essential goods.

Her discussion of practical reason therefore asks how people should respond when ordinary coordination conflicts with justified criticism. A practice may generate reasons for participation while also giving rise to reasons for its reform. Movement among different social contexts can help people recognize that arrangements taken for granted in one setting are open to challenge. "Practical Reason and Social Practices"

Elsewhere she argues that correcting one's own biases is an incomplete understanding of responsibility. The moral question also concerns what people can do together. Social explanation contributes by changing their understanding of their circumstances and the forms of collective action available to them. "Social Structure, Narrative, and Explanation"

A worked example can clarify the overall theory. The following is my reconstruction of how her approach could address an expectation that women should bear primary responsibility for unpaid care; it is not a program she presents in this exact form.

Dimension of reform	Question to investigate	Possible intervention
Moral interpretation	Is unequal responsibility understood as natural, obligatory, or freely chosen?	Develop public discussion of its costs and justifications.
Collective knowledge	Which recurring burdens remain individually experienced?	Enable caregivers to compare experiences and investigate patterns.
Material conditions	What makes alternative arrangements difficult?	Improve access to care services, time, and financial support.

Dimension of reform	Question to investigate	Possible intervention
Organizational practice	How do workplaces and institutions assume care will be provided?	Redesign schedules, leave arrangements, and expectations of availability.
Social meaning	What counts as a responsible parent or committed employee?	Make alternative performances of these roles recognizable and respected.
Distribution of consequences	Does reform transfer burdens onto people with less power?	Examine care workers' conditions and the effects across households and classes.

This reconstruction shows why changing an explicit moral proposition is only one part of the process. People may publicly accept shared responsibility while retaining arrangements that make it impracticable. Conversely, an institutional change can permit new conduct that gradually alters expectations.

It also shows why the direction of causation can vary. Moral criticism may initiate institutional reform; altered institutions may make a moral alternative credible; organized experiments may change both simultaneously. Any concrete strategy must investigate which relationships are causally important in the particular setting.

The standard of success should consequently include what happens to burdens, opportunities, standing, and agency. A change in vocabulary is relevant insofar as it contributes to changes that can be justified.

My assessment is that Haslanger's strongest achievement is connecting the justification of reform with an account of its social conditions. Moral philosophy frequently needs an explanation of how people acquire the capacities, opportunities, and collective power to act on a justified judgment. Her framework makes these questions central.

Several further strengths follow.

First, it explains how injustice can persist among people who reject it explicitly. That explanation permits criticism to move beyond choosing between allegations of insincerity and resignation about human nature.

Second, it directs attention to the intermediate organizations and practices through which much ordinary conduct is shaped. Families, workplaces, schools, professions, and movements become major sites of moral inquiry.

Third, it makes conceptual reform answerable to practical purposes and consequences. Theorists must consider what their categories enable people to see and do.

Fourth, it connects emancipation with the development of collective capacities. Reform involves people becoming better able to investigate, decide, organize, and sustain alternatives.

Fifth, her acknowledgment of uncertainty guards against assuming that justified intentions establish effective methods. A moral judgment, a causal diagnosis, and an intervention strategy require related but distinct forms of support.

The framework nevertheless leaves substantial problems unresolved. The following criticisms are my assessment of the theory, except where an acknowledged limitation has already been identified.

One concerns **the specification of justice**. Haslanger has good reasons to resist a fully predetermined social blueprint. Yet reformers still need sufficiently determinate criteria for comparing proposals. Equality, autonomy, care, solidarity, and security can conflict. Recognizing their plurality does not settle which sacrifices are justified in a particular intervention.

The theory is strongest when a serious wrong can be identified with substantial agreement and the issue is how to understand and overcome it. It is less complete as a procedure for deciding among rival moral projects whose basic evaluations remain opposed.

A second problem concerns **representation and authority**. Participatory inquiry can reduce paternalism, but participation does not guarantee equal influence. A movement can have internal hierarchies; its most articulate members may dominate its account of shared interests. Those unable to attend meetings, master its vocabulary, or bear the costs of activism may remain underrepresented.

Consequently, the fact that a proposal emerges through a social movement cannot alone establish that it represents everyone affected. A fuller account needs ways to examine internal power, preserve dissent, and correct mistaken claims to representation.

A third concerns **the evidence required for structural explanations**. A persistent disparity may justify investigation, but the diagnosis of its mechanisms requires more than describing it as systemic. Different mechanisms can produce similar outcomes, and similar interventions can have different effects.

The practical force of Haslanger's approach depends on identifying which expectations, resources, incentives, and organizational relationships actually sustain the pattern under examination. Otherwise, the vocabulary of structures can become too general to guide action or to be tested against alternatives.

A fourth concerns **the evaluation of disruptive methods**. Showing that an established practice is unjust does not settle the permissibility of every means used against it. Disruption can expose hidden dependencies and compel attention, but it can also burden vulnerable third parties, intensify conflict, or close opportunities for correction.

Her distinction between warranted criticism and warranted intervention is therefore indispensable. It also points toward further work on proportionality, responsibility for consequences, and the comparative merits of available strategies.

A fifth concerns **the durability and scale of local experiments**. Co-design can generate useful knowledge and strengthen participants' agency. Nevertheless, a successful initiative may

depend on unusual commitment, temporary funding, or favorable local relationships. Broader adoption can alter the conditions that made it successful.

The central empirical questions are whether the practice survives, whether it spreads without losing essential features, and whether it changes the larger processes that reproduce the original problem. Haslanger recognizes this difficulty, but her framework supplies an orientation for investigating it rather than a general solution.

A sixth concerns **conceptual power itself**. Concepts help people identify wrongs, but redefining a public category can redistribute credibility and standing. A reform can clarify one experience while making another difficult to express. It can also obscure disagreement by changing the terms in which disagreement must be stated.

This is why her later self-correction is theoretically significant. Ameliorative concepts should remain open to challenge from those whom they classify. Their political purpose makes scrutiny more necessary.

For comparison with conscious lawmaking, her work supports a differentiated account of deliberate norm change. A legislature can establish a formally valid rule through an authorized procedure. The social effectiveness of that rule is a further matter. Moral norms also depend on recognition, expectation, habit, and the relationships through which people hold one another accountable.

Drawing on Haslanger's framework, I would distinguish three questions:

1. **Justification:** Is the proposed norm defensible?
2. **Authorization:** Who may introduce or enforce the relevant institutional changes?
3. **Social establishment:** How does the norm become intelligible, practicable, and sufficiently sustained in ordinary life?

These questions can have different answers. An institution may validly adopt a rule that remains morally objectionable. A morally persuasive norm may lack widespread acceptance. A widespread expectation may persist without defensible justification.

Haslanger contributes especially to understanding the relationships among these dimensions. Her account also indicates why the repeal analogy has limits: removing an official rule does not remove the social expectations and material dependencies that gave an older morality its force.

The most consequential claim in her work is thus that **deliberate moral improvement requires the capacity to reshape the shared conditions under which people perceive, judge, and act**. Its unresolved challenge is to show, in contested cases, how that transformative capacity can be exercised with adequate evidence, legitimate authority, and continuing openness to correction.

****

Ingmar Persson and Julian Savulescu

Ingmar Persson and Julian Savulescu propose one of the most radical contemporary approaches to deliberate moral change: **human beings should investigate ways of improving the biological foundations of moral motivation, alongside education, reasoning, and institutional reform.** Their central concern is that technological power has expanded much faster than our willingness to use it responsibly. Humanity can increasingly affect distant populations and future generations, while its emotional dispositions remain strongly oriented toward personal interests, familiar people, and immediate circumstances. Their principal joint statement is *Unfit for the Future: The Need for Moral Enhancement* (2012). Oxford's account of the book.

For the question of consciously creating moral change, their distinctive contribution lies in distinguishing **formulating better moral principles from producing people sufficiently motivated to follow them**. They address both, but concentrate on the second. Understanding their position requires separating its philosophical justification, its empirical feasibility, and its political implications.

Their programme developed through several stages, and its qualifications matter. The following works provide the main framework:

Publication	Place in the development of their position
The Perils of Cognitive Enhancement and the Urgent Imperative to Enhance the Moral Character of Humanity (2008)	Establishes the problem of increasing human capabilities without corresponding moral improvement.
Unfit for the Future (2012)	Develops the comprehensive argument concerning technological danger and inadequate moral motivation.
Moral Enhancement, Freedom, and the God Machine (2012)	Examines enhancement, freedom, and the prevention of serious wrongdoing.
Getting Moral Enhancement Right (2013; online 2011)	Defends the programme against John Harris's objections.
Should Moral Bioenhancement Be Compulsory? (2014)	Qualifies the argument concerning compulsory implementation.
Moral Hard-Wiring and Moral Enhancement (2017)	Clarifies the relationship between biological dispositions and cultural change.
The Moral Importance of Reflective Empathy (2018; online 2017)	Explains how reflection can redirect emotional concern.

Publication	Place in the development of their position
The Duty to Be Morally Enhanced (2019; online 2017)	Develops a conditional moral obligation to improve one's moral motivation.
The Evolution of Moral Progress and Biomedical Moral Enhancement (2019)	Responds to Buchanan and Powell and clarifies their account of moral progress.

The starting point is a mismatch between the reach of human action and the reach of human concern. In their account, dispositions that developed under conditions of relatively small communities and limited technological power are inadequate for a world in which ordinary consumption, scientific research, and political decisions can have consequences across continents and centuries.

This is a relational claim. A moral psychology can become inadequate because the circumstances in which it operates have changed, even without people becoming more selfish or cruel. Their worry is therefore compatible with some historical moral improvement: modestly better people may nevertheless become much more dangerous if their capabilities grow enormously. *Unfit for the Future*.

An illustrative reconstruction makes the point clearer. A person might have considerable concern for neighbours, moderate concern for fellow citizens, and very little concern for strangers living a century later. That motivational distribution may support respectable conduct within a local community. It becomes problematic when the same person participates in decisions whose largest costs fall on those distant strangers. The problem concerns the distribution of moral attention as much as the overall quantity of benevolence.

Their diagnosis highlights several recurrent tendencies:

- **Partial altruism:** stronger concern for relatives, friends, and familiar groups.
- **Temporal bias:** greater motivational weight attached to immediate outcomes.
- **Sensitivity to identifiable suffering:** a visible individual often moves us more than anonymous populations.
- **Diminished responsibility in collective action:** people feel less responsible when their contribution is one among millions.
- **The asymmetry between action and omission:** directly causing harm typically troubles people more than failing to prevent comparable suffering.

They connect these tendencies to environmental collective-action problems and the difficulty of sustaining sacrifices for distant beneficiaries. Persson and Savulescu, "Moral Enhancement".

The philosophical significance is that a moral code can become increasingly universal in its language while remaining selective in its motivational operation. Someone may sincerely affirm that all human lives matter and nevertheless consistently organize decisions around a much narrower circle.

Their concern about catastrophe depends on an asymmetry between destruction and preservation. They argue that damaging complex systems is often easier than sustaining or improving them. A destructive act can remove the conditions for a worthwhile future, whereas preventing one catastrophe leaves other threats outstanding.

Their expression *ultimate harm* concerns the permanent destruction of the possibility of worthwhile life on Earth. It is a philosophical category for an exceptionally grave outcome; the definition itself does not establish how likely particular technologies are to produce it. Their argument gives special significance to threats that could eliminate future value rather than merely produce a temporary setback. *The Turn for Ultimate Harm: A Reply to Fenton*.

This produces two different problems, which should be distinguished analytically.

The first is **catastrophic misuse by a small minority**. If very few people possess sufficient destructive power, an improvement in average morality might leave the main danger largely intact.

The second is **cumulative harm produced by ordinary conduct**. Environmental degradation can arise without any participant intending the total result. Here, changes among large numbers of reasonably decent people could matter greatly.

These problems require different evidence and potentially different interventions. Reducing the risk from a determined malicious actor is not the same task as increasing millions of people's willingness to support collective restraint. A weakness in the programme is that arguments for these two objectives sometimes appear under the same general demand for moral enhancement, although their practical requirements differ substantially.

Their criticism of cognitive enhancement follows from this concern about power. Increasing intelligence, memory, planning, or scientific competence does not necessarily improve the ends toward which those capacities are directed. An individual can become more effective at pursuing destructive purposes.

In the 2008 paper, however, their conclusion is qualified. They identify a reason against cognitive enhancement when it is unaccompanied by moral enhancement, but explicitly decline to settle the overall balance of reasons. Their stated practical conclusion is that continuing research into cognitive enhancement should be accompanied by research into moral enhancement. Describing this as an unconditional demand to stop scientific progress would be inaccurate. *The Perils of Cognitive Enhancement*.

A substantial objection comes from J. Adam Carter and Emma Gordon: cognitive and moral development may be too interconnected for this separation to bear all the argumentative weight placed upon it. Improved understanding, intellectual responsibility, and recognition of relevant evidence can contribute to moral improvement; sound moral conduct also requires cognitive capacities. Their objection challenges a sharp division between morally neutral intelligence and independently adjustable moral motivation. Carter and Gordon, *On Cognitive and Moral Enhancement*.

The strongest interpretation of Persson and Savulescu's argument consequently concerns **technical competence without adequate concern**, rather than intelligence understood in the broadest sense. Scientific ingenuity and practical cleverness clearly do not guarantee moral judgment. Whether a richer conception of rationality includes ethical improvement remains a further philosophical question.

Their theory of change needs to be understood at three distinct levels. The following is an analytical reconstruction of the programme:

Level	What changes?	Example
Moral judgment	Beliefs about what ought to be done	Recognizing duties toward future generations
Moral motivation	The strength and scope of concern	Becoming more willing to bear costs for those generations
Social implementation	Institutions and coordinated behaviour	Establishing arrangements that make collective protection effective

These levels can move independently. People can acquire better moral beliefs without changing their conduct. Their concern can increase without their beliefs becoming more accurate. Institutions can improve behaviour through incentives without changing either beliefs or intrinsic motives.

This distinction prevents an overly broad use of "moral enhancement." A device that physically prevents an assault, an incentive that makes honesty profitable, and an intervention that increases concern for another person are three different achievements. Each might be valuable, but they raise different questions about character, responsibility, and freedom.

The principal motivational targets are altruism and a sense of justice. In their response to commentators, Persson and Savulescu explain altruism through imaginative understanding of another conscious subject and concern for that subject's welfare for its own sake. Justice supplies a further consideration, because greater benevolence toward favoured individuals can remain unfair.

They also clarify that their underlying outlook is consequentialist in broad structure, combining concern for welfare with justice. This need not entail replacing every familiar rule in daily life: they allow a distinction between underlying justification and the practical principles it is best for people to follow. Their initial aspiration can be improvement toward levels of motivation already found in unusually committed people, rather than the immediate creation of morally perfect beings. *Reply to Commentators on Unfit for the Future*.

Several implications follow.

First, their objective cannot be identified with obedience. A compliant population may tolerate grave injustice.

Second, it cannot be identified with cooperation in general. Cooperation can serve exclusionary or oppressive purposes.

Third, it cannot be identified with generosity alone. Someone can be generous within a family while supporting policies that severely disadvantage outsiders.

The relevant question is therefore always: **concern for whom, directed toward which ends, under which conception of justice?**

This exposes a difficulty in translating the theory into an intervention. "Increase concern" sounds comparatively straightforward. "Increase appropriately directed concern while correcting mistaken judgments about its scope" is a much more demanding objective.

Their account of reflective empathy shows that conscious moral change remains central to the programme. Responding to Paul Bloom and Jesse Prinz, they accept that spontaneous empathy is selective and poorly responsive to large numbers. Their reply is that imagining another person's experience can be voluntary and guided by reflection.

Reason can direct attention toward people who would otherwise remain emotionally remote. Imagination can then help generate concern that an abstract proposition alone does not produce. Their aim is an interaction between reflection and emotion: neither unregulated emotional identification nor detached calculation is sufficient. They do not claim that this process removes every bias or makes people feel equally strongly about everyone. *The Moral Importance of Reflective Empathy*.

Consider an illustrative case. Someone recognizes that a distant population faces preventable suffering but experiences this knowledge as an impersonal statistic. Deliberately attending to what that situation involves for particular lives may strengthen motivation. Reflection must then correct the opposite danger: allowing the most vivid case to displace others with greater needs.

This reveals two dimensions of improvement:

- **Intensity:** how strongly someone cares.
- **Scope and distribution:** whose interests receive that concern, and with what relative weight.

Increasing intensity without correcting scope may deepen favouritism. Expanding stated concern without increasing its motivational force may produce little behavioural change. A successful programme must address both.

Their argument for a duty of enhancement begins from ordinary moral obligations. If someone ought to perform an action, and a safe, proportionate means would make fulfilling that obligation more likely, there can be a further obligation to use that means. Where suitable means do not exist, there may be reason to develop them.

They apply this reasoning to biomedical methods as well as other forms of improvement. In the narrow sense, enhancement strengthens intrinsic moral motivation rather than inducing conformity through rewards and punishments. Its desirability remains conditional on safety, effectiveness, costs, and competing considerations. *The Duty to Be Morally Enhanced*.

The argument is stronger than a claim that self-improvement is merely admirable. It treats attention to one's foreseeable moral weaknesses as part of responsible agency. If I repeatedly fail others in a predictable way, recognizing that failure may give me a reason to alter the conditions under which I act.

Nevertheless, two distinctions are essential.

A duty to improve oneself does not automatically give another person authority to impose improvement. Obligations and enforcement powers require separate justification.

Likewise, **a duty to use an effective means does not establish that a particular proposed means is effective**. The normative argument cannot substitute for evidence about drugs, neurological interventions, or educational programmes.

The duty argument therefore supports a general orientation toward responsible self-development more securely than it supports any particular biomedical policy.

Their biological argument establishes a research possibility, not a completed technology. They propose investigating pharmaceutical, neurological, and genetic means of affecting dispositions relevant to morality. Their defence is that biological influence on motivation should not be treated as morally disqualifying in itself. They also address the objection that such interventions cannot precisely target the relevant dispositions, treating this as a serious empirical challenge rather than a demonstrated impossibility. *Getting Moral Enhancement Right*.

Here the distinction between three claims is decisive:

1. Moral behaviour depends partly on biological processes.
2. Those processes can sometimes be altered.
3. They can be altered safely and reliably so as to produce moral improvement.

The first two claims do not entail the third. A capacity can have biological foundations while remaining difficult to modify selectively. An intervention might improve one response, impair another, and operate differently across situations.

Nor does a genetic contribution to a trait establish a simple genetic route to improving it. The programme should therefore be evaluated through specific interventions and outcomes, rather than through a general opposition between "biological" and "social" explanations.

The experimental evidence demonstrates influence over particular responses, with much more limited implications for moral character. Three examples illustrate both the interest and the limitations of this research.

In a 2010 experiment, Molly Crockett and colleagues found that citalopram influenced judgments and behaviour associated with aversion to harming others. Effects varied with participants' existing empathy. This provides evidence that pharmacological intervention can affect processes relevant to moral decisions. It does not show that increasing such aversion always produces the ethically correct decision: some dilemmas concern whether preventing greater harm justifies an immediately harmful act. Crockett and colleagues, *Serotonin Selectively Influences Moral Judgment and Behavior*.

Oxytocin became prominent because of its associations with social attachment and prosocial responses. Yet experiments by Carsten de Dreu and colleagues reported increased favouritism toward one's own group. This illustrates precisely why stronger social attachment cannot simply be equated with universal benevolence. It is also important to describe these as findings under particular experimental conditions, rather than a universal law about the substance. De Dreu and colleagues, *Oxytocin Promotes Human Ethnocentrism*.

A 2012 study by Sylvia Terbeck and colleagues, including Savulescu, examined propranolol in 36 participants. It reported reduced implicit racial bias on a laboratory measure, without a corresponding change in the explicit prejudice measure. That difference is philosophically relevant: an automatic response may change while consciously endorsed attitudes do not. But neither measure by itself establishes durable improvement in real social conduct. Terbeck and colleagues, *Propranolol Reduces Implicit Negative Racial Bias*.

These examples leave a substantial distance between **altering a morally relevant psychological process** and **making an individual reliably more just across contexts**.

A 2025 scoping review of the debate likewise identified a general shortage of empirical evidence capable of settling the broader claims for or against moral bioenhancement. The field remains much more developed as a philosophical controversy than as a demonstrated programme of population-level moral transformation. Bissette and colleagues, *What's Left of Moral Bioenhancement?*.

Their disagreement with Buchanan and Powell concerns the relationship between cultural progress and motivational limitations. In *Moral Hard-Wiring and Moral Enhancement*, Persson and Savulescu distinguish articulated principles and public commitments from actual motivations and conduct. They do not claim that inherited dispositions make cultural learning impossible. Their concern is that changes in official morality can run ahead of changes in what people are prepared to do. *Moral Hard-Wiring and Moral Enhancement*.

Their 2019 response accepts that favourable social conditions can support more inclusive morality. It nevertheless questions whether diminished hostility and increased mutually advantageous cooperation necessarily produce concern for outsiders for their own sake. They also reject the interpretation that their position denies moral progress altogether. Their claim concerns its adequacy relative to technological expansion, and they favour exploring social, institutional, psychological, and biological avenues together. *The Evolution of Moral Progress and Biomedical Moral Enhancement*.

An analytical example helps locate the disagreement. A society may welcome outsiders because their participation brings prosperity. This can be a valuable development. A stronger form of inclusion would extend concern to outsiders who cannot contribute economically and may require substantial assistance. Explaining the first development does not automatically explain the second.

Their distinction is valuable, although it creates a demanding empirical task. To establish that motivations have failed to improve sufficiently, one needs more than examples of hypocrisy or continued wrongdoing. One needs an account of relevant baselines, changing circumstances, institutional constraints, and the degree of improvement required.

The freedom objection is central because deliberately changing motivation can affect the authorship of conduct. John Harris argues that moral agency involves understanding reasons and retaining a meaningful possibility of acting wrongly. A system that makes wrongdoing impossible may secure acceptable behaviour while removing something important about moral choice. Harris, *Moral Enhancement and Freedom*.

Persson and Savulescu reply that strengthening moral motivation need not bypass reasoning. Their intended improvement makes a person more responsive to reasons for concern. Someone who more readily acts on a considered judgment need not thereby be less free. They also draw on arguments questioning whether moral responsibility always requires the availability of an alternative action. *Moral Bioenhancement, Freedom and Reason*.

The disagreement becomes clearer if three cases are separated.

A person may voluntarily reduce an impulsive tendency that repeatedly defeats their own commitments. This could increase practical self-government.

A second person may receive an intervention that produces a new motivation they subsequently understand and endorse. Whether this supports autonomy depends partly on how the change occurred and whether it remains open to examination.

A third person may be manipulated into endorsing the purposes of an external authority. Behaviour may become more orderly while autonomy deteriorates.

The mere fact that motivation has a biological cause does not distinguish these cases. Consent, understanding, control, responsiveness to reasons, and opportunities for revision matter.

The "God Machine" explores a different issue from ordinary moral bioenhancement. Their thought experiment imagines an extraordinarily capable system that monitors mental states and intervenes to prevent gravely immoral conduct. It leaves considerable room for other choices.

Crucially, they distinguish such prevention from improvement of moral motivation. Preventing someone from carrying out an atrocity does not necessarily make that person more benevolent or just. The thought experiment tests whether the loss of certain possibilities must be unacceptable when it protects potential victims; it is not an available technological programme. *Moral Enhancement, Freedom, and the God Machine*.

Its philosophical force comes partly from shifting attention toward victims. The value of an aggressor's opportunity to choose wrongdoing cannot be assessed without considering what exercising that opportunity does to others.

Its principal limitation is that the imagined system is given capacities that actual authorities lack. Real institutions make mistakes, disagree about morality, respond to political incentives, and can expand their purposes. An idealized case may establish that prevention is not objectionable under every conceivable circumstance. It does not establish that a real system of mental surveillance and intervention would be justified.

Robert Sparrow develops the political objection through the idea of **freedom as non-domination**. A person can be subject to another's controlling power even when that power is not currently exercised. Consequently, showing that some enhanced actions remain psychologically voluntary does not settle whether the relationship between enhancers and enhanced is politically free. Sparrow, *Better Living Through Chemistry?*.

Their position on compulsion is conditional and developed over time. Their early argument entertained compulsory enhancement if safe and effective methods became available, partly because those most in need might refuse.

In their 2014 reply to Vojin Rakić, they clarify that their book had left the compulsory-versus-voluntary question unresolved. They acknowledge that incentives might suffice and that the magnitude of the danger would matter. They do not exclude justified restrictions of freedom, but postpone practical judgment pending suitable technologies and circumstances. *Should Moral Bioenhancement Be Compulsory?*.

It would therefore misrepresent them either to describe their programme as exclusively voluntary or to present it as an immediate demand for compulsory medication of entire populations.

Rakić's alternative is to defend voluntary enhancement while resisting compulsion. His criticism emphasizes the moral importance of the decision to undergo enhancement and disputes whether their treatment of freedom adequately protects it. Rakić, *Moral Bioenhancement and Free Will: Continuing the Debate*.

The unresolved issue is especially sharp because the refusal problem works in two directions. Refusal might limit a beneficial programme's effectiveness. But treating refusal as evidence of moral deficiency could make opposition impossible to express without becoming a target of intervention. Any proposed implementation would have to avoid that circularity.

Their analysis of democracy exposes a tension between desirable decisions and legitimate political agency. They argue that electoral institutions can reproduce citizens' short-term and partial preferences. Aggregating those preferences does not automatically correct them. Governments may therefore fail to protect interests whose beneficiaries are distant, politically weak, or not yet born. Persson and Savulescu, "Moral Enhancement".

My assessment is that this diagnosis identifies a real theoretical possibility but does not establish the superiority of any biomedical response. Institutional design offers several different objects of intervention: incentives, information, representation, decision procedures, and concentrations of power. Improving citizens' motivations is one possible route among these.

Moreover, changing the motivations from which political preferences arise introduces an additional problem of legitimacy. Democratic authority ordinarily depends partly on citizens being able to evaluate and challenge government. If government gains power to alter the evaluative dispositions of citizens, it may influence the very source of its authorization.

Sparrow's further egalitarian criticism is relevant here. He argues that claims about biologically superior moral dispositions can threaten equal political standing and encourage hierarchies between enhanced and unenhanced people. This is a criticism of the programme's possible implications, not evidence that Persson and Savulescu endorse government by an enhanced elite. Sparrow, *Egalitarianism and Moral Bioenhancement*.

The problem of who improves the improvers is not solved by the proposal itself. The authors acknowledge that morally imperfect people would have to develop, finance, select, and administer the relevant technologies. Their later discussion allows an incremental response: people can have enough motivation during calm reflection to support measures that help them during subsequent temptation. Partial improvement might make further improvement easier. *The Evolution of Moral Progress and Biomedical Moral Enhancement*.

That response is plausible for some forms of voluntary self-development. A person need not already possess perfect self-control to adopt a practice that improves it.

The collective case is harder. Researchers, corporations, public agencies, and political leaders may disagree about both the target and the distribution of benefits. An intervention could be selected because it produces compliance, commercial advantage, or political stability. Calling it "moral" would not resolve those interests.

The programme consequently requires an institutional theory of accountability in addition to a psychological theory of motivation. Its moral aspirations cannot themselves supply that institutional architecture.

A further difficulty is that virtues depend on context and on their relations to other capacities. Greater trust can support cooperation but also increase vulnerability to exploitation. Reduced anger can prevent cruelty but weaken resistance to injustice. Stronger loyalty can sustain mutual care but obstruct impartial judgment.

Harris Wiseman's *The Myth of the Moral Brain* emphasizes the biological, psychological, social, and environmental interactions involved in moral functioning. His criticism challenges simplified assumptions about moving from a neural change to a reliable moral improvement. It should not be read as showing that Persson and Savulescu literally posit a single brain centre for morality. Wiseman, *The Myth of the Moral Brain*.

The broader implication is that morality may require **better regulation and integration of dispositions**, rather than a uniform increase or decrease in individual traits.

For example, courage does not become morally beneficial simply by becoming stronger. Its value depends on understanding which risks are worth taking and which purposes deserve commitment. The same applies to empathy, persistence, and willingness to cooperate.

This also raises a measurement problem. A convincing demonstration of improvement would need to distinguish stable character change from temporary states, genuine concern from conformity, and better judgment from agreement with the investigator's preferred answer.

Their work raises a demandingness problem as well. If morally relevant concern extends to distant strangers, future people, and nonhuman animals, the demands of morality may become much greater than those ordinarily accepted. The authors explicitly connect the possibility of more demanding obligations with the need for stronger motivation. *The Duty to Be Morally Enhanced*.

There are two ways to interpret this development.

On one interpretation, enhancement helps people fulfil obligations that were already justified but difficult to meet.

On another, expanding the ability to sacrifice creates pressure to require still greater sacrifice. Improved capacity then changes expectations and potentially raises the standard of blame.

This does not refute enhancement. It does mean that the programme needs an account of legitimate personal projects, attachments, and limits to required self-sacrifice. A society of individuals more willing to serve others could be admirable; a society in which everyone is expected to suppress every personal commitment would raise a different set of moral questions.

Recent work involving Savulescu increasingly considers technologies that assist active moral development. A 2025 paper with Emma Gordon, Katherine Cheung, and Brian Earp distinguishes speculative interventions that directly produce improvement from technologies that facilitate the person's own moral learning. It discusses psychedelic-assisted contexts, virtual reality, and "Socratic" AI while leaving their effectiveness an open empirical question.

The paper argues that assistance need not deprive improvement of its value when the individual's understanding and agency remain central. This is a development in Savulescu's wider collaborative work; Persson is not a coauthor. *Moral Enhancement and Cheapened Achievement: Psychedelics, Virtual Reality and AI*.

Conceptually, such approaches bring the debate closer to established questions about education. A technology might support attention, imaginative perspective-taking, or reflection without specifying the final moral judgment. Nevertheless, the choice of experiences, questions, and feedback still shapes the learner. Facilitation does not eliminate the need to examine influence and authority.

For the broader theory of conscious moral change, the comparison with deliberate lawmaking reveals both the originality and the limits of their approach. The following is my interpretation of its significance.

A legislature can alter publicly recognized rules. Persson and Savulescu direct attention toward the psychological capacities through which rules and principles acquire motivational force. Their programme therefore adds a further dimension to conscious moral reform: deliberate intervention in the dispositions that support moral action.

But this does not amount to the positivization of morality in the same sense as legislation.

First, **a successful intervention cannot establish the validity of its own moral target**. Producing agreement with a principle does not show that the principle is correct.

Second, **the authority to influence motives is distinct from the authority to promulgate rules**. Even if a community accepts a prohibition, it may disagree about whether authorities should modify citizens so that violating it becomes psychologically difficult or impossible.

Third, **moral change requires opportunities for correction**. Future generations may reject present assumptions. An intervention that strengthens reflection and responsiveness to reasons can support that possibility. One that fixes particular commitments beyond criticism could obstruct it.

Fourth, **changes in behaviour, character, and moral understanding must remain distinguishable**. A system could become more orderly while its members become less capable of independent moral judgment. Conversely, stronger moral judgment might increase disagreement and resistance to existing institutions.

These distinctions are especially important when evaluating claims that a new morality is more advanced. The capacity to produce a disposition is a causal achievement. Whether producing it constitutes progress remains a normative question.

My assessment is that their most powerful contribution concerns the gap between justified moral insight and effective moral motivation. They force theories of deliberate moral change to address why knowing, publicly affirming, and even sincerely accepting a principle so often fail to produce action. They also challenge the assumption that expanding human capabilities is beneficial irrespective of the purposes governing their use.

Their argument is less conclusive when it moves from that diagnosis to the necessity of biomedical intervention. Several additional propositions would have to be established: that relevant motivational limitations cannot be adequately addressed through other means; that a specific intervention produces durable improvement; that benefits exceed harms; and that its development and use can be governed legitimately. The gravity of the problem gives reasons to investigate these possibilities, but does not itself establish them.

Their theory thus leaves a fundamental task for any consciously directed transformation of morality: **to improve people's capacity to respond to moral reasons while preserving their capacity to question the proposed improvement itself.**

****

Markus Gabriel

Markus Gabriel's theory of conscious moral change centres on the claim that human beings can discover moral truths, correct inherited moral errors, and deliberately organize social life around the resulting knowledge. Its distinctive feature is the connection between moral realism and institutional reform: moral progress requires both better understanding and arrangements that make morally better conduct possible, durable, and widespread.

"Creating moral change" therefore has a particular meaning for Gabriel. People can change their convictions, institutions, and practices; their decisions do not, merely by being accepted or enforced, make something morally right. This distinction separates his project from a conception of moral reform as the authoritative invention of new values. His principal statement is *Moralischer Fortschritt in dunklen Zeiten* (2020), translated as *Moral Progress in Dark Times: Universal Values for the 21st Century* (2022). Publisher's description.

His subsequent work makes the practical side increasingly significant. The jointly authored *Towards a New Enlightenment* (2022) addresses the organization of knowledge; *Gutes tun* (2024) develops ethical capitalism; *Moralische Tatsachen* (2025) gives a more systematic account of moral reality and knowledge; and *Ethische Intelligenz* (2026) considers AI's possible contribution to moral improvement. These developments make it insufficient to interpret him solely as a philosopher asserting the existence of timeless values.

The following analysis distinguishes Gabriel's stated positions from the further implications and objections that can be developed from them.

Gabriel's starting point is a distinction between **the existence of socially accepted morality and the correctness of that morality**. A sociological description can establish that a community approves a practice, teaches it to children, rewards conformity, and punishes dissent. Whether that practice deserves approval remains a separate question. His 2025 account presents moral facts as features of reality concerning what persons should do or refrain from doing, connecting ethical judgment with the human capacity to understand and direct one's life. Gabriel's presentation of *Moralische Tatsachen*.

The importance of this distinction can be illustrated by considering an accepted hierarchy. Its defenders might explain its historical origins, its contribution to stability, or its compatibility with religious tradition. Those explanations establish neither its justice nor its injustice. Conversely, discovering that a reform movement has economic supporters or political ambitions does not settle whether its demands are morally justified. Causal explanation and moral justification answer different questions.

For a theory of deliberate moral change, this means that neither antiquity nor novelty provides a sufficient warrant. A traditional norm may preserve a valuable insight; a fashionable norm may express prejudice. The reformer must show what is wrong with the existing practice and why the proposed alternative is better. Calling an alternative "modern" does not perform that work.

Gabriel's defence of moral facts belongs to his broader **New Realism**. His ontology recognizes different fields of sense: domains in which things exist and can be understood. Physical objects do not exhaust reality. In an interview with Graham Harman, he explicitly includes moral values among real, nonphysical objects. Their absence from a physical description therefore does not establish their unreality. A moral judgment need not be a disguised statement about particles, brain activity, or economic advantage to be capable of truth. Gabriel's interview with Harman.

An explanatory analogy helps here. A physical description of someone signing a document does not, by itself, establish whether the signature constitutes a promise, whether consent was coerced, or whether the transaction exploits another person. Understanding the action requires concepts appropriate to its interpersonal and institutional setting. Gabriel's approach allows moral understanding to concern something real even though its subject matter cannot be captured by physics alone.

Nevertheless, his realism should not be described as making morality wholly independent of humanity. He explicitly distinguishes independence from private or collective opinion from independence from human mindedness altogether. Moral requirements address beings capable of understanding and responding to them. In his 2022 replies to critics, humanism and the accessibility of moral knowledge are integral to his position. Gabriel, "Repliken auf Klingner, Moledo, Stiening und Pasquarè".

The conceptual point is that *being about human beings* differs from *being made true by whatever human beings believe*. An obligation may depend on a person's vulnerability, capacities, relationships, and available options without depending on whether others approve of that obligation.

Gabriel combines three commitments: moral realism, humanism, and universalism. Moral facts exist; human beings can know them; and their validity does not depend on the cultural identity of the persons concerned. He also distinguishes good, evil, and the morally neutral. That third category matters because recognizing objective morality does not require treating every difference in lifestyle as a moral conflict. Erica Lucast Stonestreet's analysis of the book.

Analytically, this gives moral reform two possible directions. It can establish that an accepted practice is wrong. It can also establish that a condemned practice belongs within the sphere of legitimate freedom. Moral progress need not mean an ever-expanding accumulation of prohibitions. It can involve removing unjustified condemnation and releasing people from requirements that have no adequate moral basis.

Universal validity also needs to be distinguished from uniform prescriptions. A general injunction can be too crude to determine what someone should do in a particular situation. Gabriel's later formulation, **ethical situationism**, makes concrete circumstances central. His "concrete ethics" proceeds through situations and their morally relevant similarities, rather than treating practice simply as the application of a completed inventory of abstract principles. Discussion of this position in the review of *Moralische Tatsachen*.

Consider an example of my own. Two employees disclose confidential information. One does so for personal revenge; the other exposes a serious danger that management has concealed. Describing both actions simply as "violations of confidentiality" omits facts necessary for moral assessment. Yet recognizing those differences does not make judgment arbitrary. Anyone facing the same relevant circumstances should be assessed by the same considerations.

This is a demanding combination: sensitivity to circumstances together with resistance to special pleading. The circumstance "this person is my political ally" cannot automatically justify conduct that would be condemned in an opponent. By contrast, differences in coercion, knowledge, responsibility, or available alternatives may be morally relevant.

The difficult question then becomes how people acquire reliable moral knowledge. Gabriel's 2025 explanation gives conscience a limited but important role. Conscience draws attention to something requiring examination; it does not deliver an infallible command. He also describes moral learning through identifying relevant similarities between situations and combining clearer judgments to address more difficult cases. Gabriel's 2025 interview with Michael Hesse.

That account can be reconstructed as a process involving several different capacities:

- Sensitivity enables someone to notice humiliation, suffering, unfairness, or an unacknowledged responsibility.
- Investigation establishes what is actually happening and which consequences are reasonably foreseeable.
- Reflection examines whether an initial judgment rests on a defensible reason.
- Comparison tests whether similar cases receive similar treatment.
- Dialogue introduces experiences and objections that the original observer may have missed.
- Revision corrects judgments that fail these tests.

This sequence is an analytical reconstruction, rather than a standardized method Gabriel has experimentally validated. Its value is that it identifies why neither spontaneous feeling nor abstract reasoning is sufficient on its own.

Gabriel connects this approach with Iris Murdoch's idea of moral attention: learning to attend properly to persons and situations. He extends that orientation toward the social and ecological relations in which human life is embedded. Introduction to *Moralische Tatsachen*.

The implication for conscious change is substantial. Moral education must affect what people notice, not merely the statements they can repeat. Someone may sincerely endorse equal respect while overlooking how a routine procedure humiliates a subordinate. Another person may condemn exploitation while failing to investigate the conditions under which a familiar product is made. Moral learning concerns the interpretation of ordinary situations as well as the acceptance of general ideals.

At the same time, these examples reveal an epistemological difficulty. How do we distinguish a newly accurate perception from a newly acquired prejudice? A strongly felt conviction may result from attention to neglected suffering, but it may also result from selective exposure, social pressure, or an emotionally persuasive narrative. The existence of moral truth does not automatically provide a reliable test for every claim to possess it.

Stonestreet identifies a related weakness in the earlier book: Gabriel emphasizes the need for an adequate description of a morally complex situation but offers limited guidance for resolving this descriptive problem. That criticism concerns the 2020/2022 book; the later work expressly attempts a fuller treatment of practical knowledge. Stonestreet's review.

A particularly controversial element is Gabriel's denial of genuine moral dilemmas. His 2022 clarification targets situations allegedly imposing incompatible, unconditional obligations. He accepts tragic circumstances in which losses cannot be avoided and distinguishes these from culpable wrongdoing. He also argues that institutional arrangements should reduce the circumstances that force people into such tragic choices. Gabriel's replies to his critics.

This distinction has practical force. A society should examine whether repeated tragic choices arise from remediable organizational failures. For example, a public service that continually forces staff to choose between neglecting one legitimate need and another may require changes in resources, responsibilities, or procedures. Deliberate moral reform can therefore address the production of bad situations, rather than repeatedly demanding heroic decisions from individuals trapped within them.

However, Gabriel's denial of dilemmas is a substantive philosophical commitment. Moral realism alone does not logically establish it. Someone can believe that moral requirements are objective while also holding that certain requirements conflict irreducibly. The disagreement concerns the structure of obligation and the relation between duty and possibility. A detailed account of moral change should keep that issue separate from the more general claim that moral judgments can be true.

Gabriel's diagnosis of contemporary society also distinguishes technological achievement from moral progress. In his reflections on the pandemic, he argued that scientific knowledge needed coordination with the humanities and social sciences because decisions affected multiple dimensions of human life. Improving technical control over one problem could produce harms elsewhere. Gabriel's interview on the pandemic and the humanities.

The general lesson is that increased capability leaves the question of worthwhile purposes open. A technology can make a beneficial activity easier, intensify an injustice, or do both simultaneously. Deliberate moral change must therefore involve evaluating the purposes, distribution, and institutional conditions of innovation.

Gabriel's proposal for a **New Enlightenment** addresses this coordination problem. It seeks to restore confidence in knowledge and moral reasoning while making them effective in social practice. In a 2020 interview, he explicitly described his writing as an attempt to influence society in a morally desirable direction. He stressed self-criticism, global coordination, and the need to introduce ethical education much earlier, including in schools. Gabriel's interview in *Börsenblatt*.

The educational implications can be developed further. Teaching ethics for this purpose would involve examining reasons, identifying inconsistencies, interpreting evidence, and reconsidering one's own conduct. Students would need opportunities to encounter cases in which respectable conventions conceal mistreatment, as well as cases in which moral accusations are insufficiently supported. Otherwise, ethical education risks becoming another mechanism of conformity.

This is also why public morality cannot simply be dismissed as an obsolete collection of errors. Ordinary experience contains moral knowledge: people recognize betrayal, care, loyalty, cruelty, and responsibility long before they study philosophical ethics. The relevant task is to distinguish sound practical understanding from prejudice and to extend insight into situations where it has been inconsistently applied.

The institutional dimension becomes explicit in *Towards a New Enlightenment: The Case for Future-Oriented Humanities*, coauthored by Gabriel, Christoph Horn, Anna Katsman, Wilhelm Krull, Anna Luisa Lippold, Corine Pelluchon, and Ingo Venzke. The text proposes closer cooperation among scholarship and other social sectors, changes in academic institutions, and support for independent, creative inquiry. It also accommodates more than one philosophical approach; its entire programme should not be attributed exclusively to Gabriel's moral realism. The jointly authored text.

Its relevance lies in the organization of intellectual work. Universities and research institutions would contribute to formulating possible futures, examining their values, and assessing their consequences. This requires communication across disciplinary boundaries and with people outside academia. The New Institute's presentation of the programme explicitly calls for expanding the humanities' contribution beyond interpretation and criticism to creative and normative work. The New Institute's account of the programme.

One can see why this appeals to Gabriel. A philosopher may identify a morally important question but lack the empirical knowledge needed to answer it. An engineer may understand a technical possibility without possessing special authority concerning its social desirability. A manager may understand organizational constraints while remaining unaware of experiences that those constraints conceal. Cooperation potentially corrects these partial perspectives.

Yet cooperation itself requires an institutional design. Who selects participants? Which experiences count? Who controls the agenda and funding? Can participants challenge the definition of the problem, or may they only improve implementation of an already chosen policy? A theory of moral learning becomes politically consequential at exactly these points.

Gabriel offers an unusually concrete illustration in his discussion of policing. He proposes forums bringing police officers together with people who have experienced discrimination, with philosophical and sociological participation, to understand situations and develop proposals for political authorities. The intended movement is from experience and dialogue toward institutional change. Gabriel's interview in *taz*.

The strength of this proposal is its attempt to connect those affected by a practice with those responsible for carrying it out. Its limitations are equally instructive. Dialogue may reveal misunderstanding, but it may also encounter intimidation, incompatible interests, or incentives to conceal misconduct. Evaluation would therefore need to ask whether participants can speak freely, whether proposals reach decision makers, and whether subsequent conduct changes.

The clearest formulation of Gabriel's institutional conception appears in his 2025 book: moral progress involves embedding accumulated practical knowledge in institutions. Law helps preserve, extend, and transmit such achievements. Progress consequently includes maintaining successful arrangements and making further improvements; recognizing a moral truth is insufficient without changes in social practice. Introduction to *Moralische Tatsachen*.

For analytical purposes, several distinct achievements must therefore be kept apart:

Dimension of change	What improves	What would remain insufficient
Moral understanding	People recognize a previously overlooked wrong	Recognition without changed conduct
Public justification	A reform is supported by defensible reasons	Widespread approval alone
Practical opportunity	People gain feasible ways to act better	Demanding conduct that institutions obstruct
Organizational practice	Procedures and incentives support improvement	A declaration without implementation
Institutional durability	Gains survive changes in personnel and circumstances	Temporary success dependent on exceptional individuals
Correctability	Institutions can recognize and remedy their own errors	Treating an adopted reform as beyond criticism

This table is my reconstruction of the requirements implicit in the project. It also supplies criteria for evaluating reforms that claim Gabriel's philosophical support.

His later work includes a particularly relevant engagement with Luhmann. Gabriel invokes ethics as a reflection theory of morality, while arguing that reflection itself returns to moral practice through proposals for improvement. He emphasizes cooperation among overlapping normative arrangements. This is Gabriel's characterization and revision of the relationship, rather than a neutral account of Luhmann's entire theory. *Moralische Tatsachen*, introduction, p. 25.

The underlying question is whether reflection on morality can become an organized source of moral transformation. Once ethical inquiry identifies contradictions or neglected responsibilities, can its conclusions enter education, administration, business, and law in a way that changes ordinary conduct? Gabriel's project answers affirmatively. But the different selection mechanisms of these institutions remain a major problem: political advantage, profitability, professional reputation, and moral justification do not always converge.

This brings his theory close to the question of **the conscious institutionalization of morality**, including its partial analogy with legislation. The analogy should be handled carefully. An institution can introduce a code, redefine a professional duty, prohibit a practice, or establish a new remedy. These acts change the social and legal environment. They do not settle the moral correctness of what has been introduced.

Consider an organization that adopts a rule against humiliation. The rule can make expectations explicit, create a complaint procedure, and change sanctions. Its adoption can therefore support moral improvement. But its wording may also be vague, its enforcement selective, or its application hostile to legitimate criticism. The moral reason for reform must continue to govern the evaluation of the reform itself.

For Gabriel's approach, this means that the authority to enact rules and the authority of moral reasons must remain distinguishable. Otherwise, an institution could make itself morally

infallible simply by declaring its preferred standards binding. Equally, a philosopher's confidence in a moral conclusion cannot by itself establish a right to impose that conclusion.

Gabriel explicitly states that democratic education should help people overcome error while the state must not dictate truth to its citizens. He also advocates cooperation among disciplines and social sectors on equal terms. Gabriel's interview in *reformiert*.

A constructive development of his position would therefore require protections for disagreement, public justification, institutional review, and participation by affected people. These are institutional requirements I infer from combining his realism with his acknowledgment of fallibility. They should not be mistaken for a complete constitutional design supplied by Gabriel.

His treatment of identity and universalism raises a related issue. He criticizes political approaches that fix people within stereotyped group identities, while recognizing that constructed categories have real effects. The relevant moral issue is how persons are treated, including experiences of oppression. Stonestreet's discussion of Gabriel's identity critique.

The attraction is clear: condemnation of mistreatment should not depend on the victim's membership in a favoured group. The difficulty is that identifying systematic mistreatment often requires studying group patterns. A defensible universalism must therefore be able to recognize patterned disadvantage without turning its victims into homogeneous representatives of an identity.

Gabriel's criticism of public moralizing also bears directly on conscious change. In his 2020 *taz* interview, he contrasts online expressions of indignation with effective institutional action and criticizes the tendency to experience denunciation as an achievement in itself. The *taz* interview.

His sweeping dismissal of some online activity is open to objection: communication can coordinate action, reveal wrongdoing, and alter institutional incentives. The useful analytical distinction is between expression and demonstrated effect. A campaign should be assessed by what it enables people and institutions to do, alongside the accuracy and fairness of its claims.

The economic extension of Gabriel's programme makes institutional agency more concrete. *Gutes tun* (2024), published in English as *Doing Good: How Ethical Capitalism Can Save Liberal Democracy* (2026), argues for bringing ethical insight into the operation of capitalism. English publisher's presentation.

In a 2024 lecture, Gabriel drew on Colin Mayer's conception of businesses solving problems in morally acceptable ways. This allows economic activity to contribute to improvement through products and services that meet legitimate needs. It also separates the moral assessment of outcomes from the personal virtue of entrepreneurs: beneficial results need not originate in admirable motives. Eurac Research's account of the lecture.

His proposal for a **Chief Philosophy Officer**, supported by an ethics team, gives this orientation an organizational form. Such a role would bring ethical examination into corporate decision making and the assessment of what the company produces. Gabriel reaffirmed the proposal in a 2026 interview on ethical capitalism. Interview in *Filosofie Magazine*.

Its potential contribution is to make ethical inquiry a continuing organizational responsibility. The decisive issue is its actual authority. Can an ethics officer challenge a lucrative product, demand access to internal information, or publish a dissenting judgment? What happens when the officer's conclusion conflicts with the interests of those who control the appointment?

Kohei Saito raises the corresponding objection that ethical capitalism risks becoming "ethics washing" and that corporate philosophers can be absorbed into existing power structures. His criticism appears in the Japanese publisher's discussion of Gabriel's economic work. Saito's critical review.

That objection identifies a structural limit: ethical expertise can improve decisions only if institutions allow it to affect decisions. Adding an advisory position does not itself resolve conflicts among profit, competitive pressure, public welfare, and long-term responsibility.

Gabriel's 2026 *Ethische Intelligenz* extends the project to AI. In his own presentation, AI agents could help articulate social values and support technical and moral progress. He envisages humane AI becoming part of viable business models and improving human ethical intelligence. This is a programme for development, not evidence that an algorithm for moral truth has been established. Gabriel's presentation of *Ethische Intelligenz*.

The proposal creates a particularly sharp problem within his own realism. Data about accepted values record what people say, approve, and do. They do not automatically distinguish moral knowledge from inherited prejudice. An AI system could reproduce a majority's errors with impressive consistency. It could also help expose inconsistencies, compare situations, and make neglected consequences visible. The philosophical and institutional task is to establish how such assistance remains open to correction and accountable to the people affected.

Several broader reservations follow from this reconstruction.

First, Gabriel sometimes presents the alternatives to realism too narrowly. In his 2025 interview, he argues as though abandoning objective moral facts leaves only descriptions of behaviour. Yet constructivist and expressivist approaches can also formulate moral criticism and defend universal commitments. One can reject his account of moral reality without abandoning the distinction between better and worse conduct. Gabriel's formulation of the argument. His case for realism therefore needs more than the observation that serious moral criticism is indispensable.

Second, agreement about elementary wrongs does not automatically resolve institutional controversies. People can agree that avoidable suffering matters while disagreeing about causal evidence, responsibility, priorities, and feasible remedies. These disputes require inquiry at a level of detail that general moral confidence cannot replace.

Third, moral knowledge and motivation remain distinct problems. A person may understand a responsibility while facing strong incentives to disregard it. An organization may recognize harm while profiting from its continuation. Consciously producing moral change therefore requires attention to resources, sanctions, dependencies, and opportunities for collective action, alongside improvements in judgment.

Fourth, moral progress requires defensible ways of assessing results. A change may reduce one injustice while producing another, benefit some people while imposing serious costs on others,

or improve attitudes without improving conduct. Gabriel's philosophical orientation helps formulate the questions, but it does not supply an agreed measurement system for answering them.

Finally, his claim to novelty should be located precisely. Moral realism, universalism, ethical education, and the institutional importance of morality all have substantial histories. Gabriel's distinctive contribution lies in connecting them through his ontology, his account of situated practical knowledge, and a programme spanning public institutions, academic inquiry, economic organizations, and technology. His strongest challenge to moral reformers is that they must explain why a change is better and how it becomes effective. The corresponding challenge to his own theory is to show how institutions claiming moral knowledge can reliably recognize and correct their mistakes.

****

The American Moral Majority Movement's analyses and theory

The American Moral Majority developed a practical programme for consciously reshaping public morality through religious mobilization, education, mass communication, elections, and law. Its distinctive aim was to restore the public authority of standards it regarded as permanently valid. It therefore combined a theology of moral continuity with an ambitious strategy of social and institutional change.

This was an explicit objective. In a 1980 television interview, Jerry Falwell described the organization as a means of achieving **"moral change through the legislative process."** That formulation makes the movement particularly relevant to the question of how morality becomes an object of deliberate collective intervention. Falwell, *Politics in the Pulpit*, 1980

An essential qualification is that the Moral Majority did not produce a unified philosophical theory comparable to those of Rawls or Habermas. Its position emerged from sermons, books, campaign documents, political practices, and arguments supplied by associated intellectuals. Consequently, the account below distinguishes its declared convictions from the broader theory of moral transformation that can be reconstructed from them.

The historical organization must first be distinguished from the wider American Christian Right. Founded in 1979 under Falwell's leadership and dissolved in 1989, the Moral Majority was one influential organization within a larger constellation of conservative religious movements. Falwell's Baptist ministry, television audience, and educational institutions provided an established base for political mobilization. Conservative strategists helped connect that base to electoral campaigning. Its constituency was especially associated with white conservative Protestants, although its coalition ambitions extended beyond them. Association of Religion Data Archives, historical overview; Pew Research Center, historical assessment

The distinctions matter intellectually. Falwell was its principal public spokesman; Tim LaHaye helped articulate its diagnosis of cultural conflict; Paul Weyrich and other conservative organizers contributed political strategy. Francis Schaeffer was an influential allied thinker whose arguments helped justify Christian political engagement. Their positions overlapped without constituting a single doctrine. Nor did all evangelicals support the organization: American evangelicalism contained substantial political and theological diversity. National Humanities Center, "The Christian Right"

Its emergence cannot adequately be explained by one moral controversy. The movement arose amid conflicts over abortion, sexuality, feminism, education, religious authority, and federal intervention. These disputes acquired force because religious institutions already

possessed audiences, organizations, money, and leadership networks capable of converting dissatisfaction into collective action.

The role of race and private religious education is especially important. Historian Randall Balmer argues that federal pressure on racially discriminatory Christian schools, particularly disputes over tax exemptions, supplied a decisive organizational stimulus. The relevant federal policy developed before Jimmy Carter's presidency; it was not simply a new Carter initiative. This interpretation reveals that the movement's defence of morality was also a defence of institutional autonomy and established social arrangements. It does not establish that every participant's opposition to abortion was insincere. Balmer, "The Historian's Pickaxe"

Daniel K. Williams challenges explanations that reduce the Christian Right's origins either to segregation or to an immediate reaction against *Roe v. Wade*. He emphasizes the broader development of conservative Sunbelt religion, economic and organizational growth, and opposition to secularization and the sexual revolution. Evangelical Republican voting also preceded the Moral Majority. Its achievement was therefore to intensify and coordinate existing tendencies, rather than introduce religious conservatives to politics for the first time. Williams, "What Were the Real Origins of the Christian Right?"

For understanding conscious moral change, the historical point is this: **a moral movement emerges when leaders connect particular grievances to a comprehensive explanation of social disorder and provide an organization through which people can act.** Moral conviction alone does not explain why a national campaign begins at one moment rather than another.

The movement's foundational distinction was between unchanging moral truth and a changing society that had departed from it. On its own understanding, the central moral standards were not human inventions open to revision whenever social preferences changed. Their ultimate authority came from God, biblical revelation, and an order of human life believed to be objectively binding.

This position produced an asymmetry:

- Social practices could change, but change did not establish their moral legitimacy.
- Laws could authorize conduct that remained morally wrong.
- Majorities could accept practices that violated objective moral standards.
- Institutions could therefore require reform even when they reflected prevailing opinion.

The movement's normative standpoint was consequently more than customary traditionalism. If actual public conduct had become permissive, the movement could condemn that conduct by appealing to a higher standard. Its relationship to public morality was thus double: it claimed to represent ordinary Americans' enduring convictions while also seeking to correct the society in which those Americans lived.

This is crucial to the meaning of *moral change*. The movement generally rejected changing the ultimate criteria of right and wrong. It sought changes in what people believed, what they did, what institutions taught, and what the state permitted. Its desired transformation was **a change in the social effectiveness of morality**, accompanied by a restoration of its claimed foundations.

Falwell's diagnosis of moral decline contained a recognizable account of moral socialization. In *Listen, America!* (1980), he attributed changes in younger generations partly to television, public schooling, weakened discipline, and government policies. Television allegedly normalized permissive conduct; schools allegedly displaced biblical authority with secular humanism; welfare arrangements allegedly weakened responsibility and the work ethic. These were his causal claims, not findings established by the book through systematic research. Falwell, excerpt from *Listen, America!*

The theoretical implication is substantial. If institutions teach people how to evaluate conduct, moral decline cannot be remedied solely by telling individuals to behave better. The environments in which judgments and habits develop must also change.

An analytical reconstruction of this reasoning yields three levels:

Level	Source of moral disorder in the movement's diagnosis	Intended intervention
Individual	Sin, weak discipline, insufficient religious commitment	Conversion, repentance, religious instruction
Social formation	Families, schools, and media transmitting undesirable standards	Strengthen parental and religious authority; challenge cultural messages
Public institutions	Law and government protecting or encouraging objectionable conduct	Electoral mobilization, legislation, administrative change, judicial appointments

These levels were supposed to reinforce one another. Personal renewal would generate responsible citizens; responsible citizens would change institutions; reformed institutions would provide better conditions for the moral formation of subsequent generations.

There was, however, an unresolved tension. The movement stressed personal responsibility while also arguing that powerful social environments shaped individual conduct. A fully developed theory would have needed to explain how responsibility should be allocated among individuals, families, markets, schools, and governments. Its political rhetoric often moved between these explanations without resolving their relationship.

"Secular humanism" supplied the movement with an account of an opposing moral system. Particularly in LaHaye's arguments, secularization meant that human preferences and judgments were displacing divine authority. Cultural institutions were portrayed as promoting this displacement, while Christians had allowed their public influence to diminish through political withdrawal. His *The Battle for the Mind* appeared in 1980; a later interview clearly restated this account of institutional influence and Christian abdication. LaHaye, bibliographic record; LaHaye, *Modern Reformation* interview

This diagnosis made apparently separate disputes intelligible as parts of one struggle. School curricula, abortion, sexual conduct, and religious expression could all be interpreted as consequences of replacing a God-centred worldview with a human-centred one.

Its explanatory weakness was that it could turn diverse actors and arguments into a single adversary. Secular humanists, liberal Christians, constitutional lawyers, feminists, commercial

broadcasters, and public educators did not form one intellectually unified organization. Their reasons for supporting particular changes also differed considerably.

Moreover, secular humanism offered moral principles of its own. The *Humanist Manifesto II* of 1973 grounded ethics in human experience, needs, dignity, and responsibility. The conflict was therefore between rival accounts of moral authority and moral conduct. It cannot adequately be described as a struggle between possessing morality and having none. American Humanist Association, *Humanist Manifesto II*

The movement's rhetoric joined moral diagnosis to a narrative of national decline and possible redemption. Falwell presented America as a country with a special inheritance and responsibility that faced danger because it had departed from its proper foundations. Scholars have situated *Listen, America!* within the American *jeremiad*: a public denunciation of collective wrongdoing that simultaneously promises renewal through repentance. *Journal of Southern Religion*, roundtable on *Listen, America!*

This narrative performed several functions. It made private anxieties nationally significant, connected ordinary participation to a larger historical purpose, and provided a reason to act urgently. It also reconciled two potentially conflicting messages: the country remained worth defending, yet its institutions required major change.

The difficulty was evidential. Once national troubles were interpreted as consequences of moral departure, almost any crisis could confirm the diagnosis. Economic instability, international weakness, family disruption, and cultural conflict could all become signs of the same underlying disease. Such a framework was effective for mobilization but less effective for distinguishing among competing causal explanations.

Its moral programme connected several policy fields under a common understanding of social order. Falwell's contemporaneous description combined protection of unborn life, the traditional family, opposition to pornography and drugs, national defence, and support for Israel. The organization sought cooperation among people who shared these commitments despite theological differences. Falwell's 1980 explanation of the platform

The following table distinguishes the movement's stated purposes from the social mechanisms implicit in them:

Field	Desired moral direction	Mechanism of change
Abortion	Recognition and protection of unborn human life	Restrictions, constitutional change, changes in judicial interpretation
Family and sexuality	Reinforcement of heterosexual marriage and prescribed family responsibilities	Public advocacy, education, resistance to competing legal recognition
Gender relations	Protection of differentiated family roles	Opposition to feminist reforms, including the Equal Rights Amendment
Education	Greater parental influence and public recognition of religion	School-policy campaigns and support for religious education

Field	Desired moral direction	Mechanism of change
Media	Reduced cultural acceptance of sexual permissiveness and pornography	Monitoring, public criticism, pressure on broadcasters and advertisers
Citizenship	Political participation understood as a moral responsibility	Registration, political instruction, voting, and organizing
National policy	Protection of a nation understood to have moral and religious responsibilities	Support for conservative foreign and defence policies

These positions are documented across accounts of the organization and the wider Christian Right. They should not be treated as a complete formal code unanimously accepted by every supporter. National Humanities Center, historical analysis; Moral Majority: organizational overview

The family occupied a special position because it connected moral authority, everyday discipline, sexuality, and generational reproduction. From the movement's standpoint, weakening the family meant weakening the principal institution through which children learned obligation and self-restraint. Family policy therefore appeared to concern the survival of moral society itself.

Yet the family being defended was historically specific. Falwell's preferred allocation of responsibilities placed particular emphasis on male leadership and women's maternal role. Lauren Winner's analysis points out that treating this arrangement as timeless obscured the variety of earlier household economies and women's work. Winner's contribution to the *Listen, America!* roundtable

The theoretical issue is broader than this historical correction. A movement may describe its preferred institution as natural while selecting one version from several arrangements found in its own tradition. Conscious moral restoration therefore involves interpretation and choice, even when its advocates understand themselves as transmitting an unchanged inheritance.

The first population the movement needed to change was its own potential constituency. Religious conservatives had to acquire a stronger conviction that organized political participation was part of their moral responsibility.

This involved changing the relationship between religious identity and citizenship. A person could retain the same beliefs about family, sexuality, and biblical authority while developing a new judgment about what those beliefs required politically. Voting, campaigning, contributing money, and monitoring officials could become duties attached to an existing religious commitment.

This is a distinct form of moral innovation. It changes the *obligations inferred from a belief* without necessarily changing the belief itself. Political disengagement becomes harder to justify once it is represented as permitting harmful institutions to govern children and society.

Susan Friend Harding's *The Book of Jerry Falwell* (2000) is especially useful here. Her account emphasizes the role of religious language in transforming the public identity of fundamentalists and their allies. Speech did more than express an already constituted political constituency: it

helped people understand themselves as members of one. Scholarly discussion of Harding's analysis

The movement's coalition strategy required a selective separation of theological agreement from moral cooperation. Religious groups that disagreed about salvation, religious authority, and doctrine could nevertheless collaborate on particular policies. Falwell explicitly presented the Moral Majority as a political organization rather than a common church. Falwell, 1980 interview

The sociological consequence was the construction of a shared political identity across confessional boundaries. For a specified range of questions, agreement about family or abortion became more organizationally consequential than disagreement about theology.

Such cooperation did not require theological relativism. Participants could continue to believe that their own religion was uniquely true. They instead narrowed the subjects on which agreement was necessary for joint action.

This arrangement also generated a tension. A coalition becomes broader by reducing the number of commitments required for membership. A confessional religious community preserves its identity by maintaining commitments that distinguish it from others. A politically useful shared morality could therefore appear religiously incomplete.

LaHaye's later discussion of school prayer illustrates the problem particularly well. In a 1995 interview, he defended public prayer partly for its capacity to cultivate recognition of God, distinguishing this purpose from conversion. His evangelical interviewer questioned whether a generalized public religiosity might dilute specifically Christian teaching. *Modern Reformation*, interview with LaHaye, March–April 1995

Mass communication made moral intervention continuous and organizationally scalable. Broadcast preaching could connect dispersed listeners to a recognizable leader. Direct mail could turn an audience into donors and participants. Repeated communication could connect a moral interpretation of events with a specific action.

The crucial development was the coordination of different activities:

1. Identify a practice as morally threatening.
2. Explain its significance through a shared worldview.
3. identify institutions or officeholders held responsible.
4. Offer an accessible action—write, register, attend, contribute, or vote.
5. Report developments and renew participation.

This is my reconstruction of the organizational mechanism. It helps explain why a movement equipped with communication networks could produce political effects even without persuading large numbers of outsiders to adopt its theology.

A concrete example appears in Falwell's *Clean Up Television* campaign papers. Supporters were invited to complete questionnaires and use materials for monitoring programmes and contacting sponsors. The documents also record disagreement over tactics: Falwell said the Moral Majority had withdrawn from a related coalition because it opposed that coalition's boycott approach. The precise tactic mattered; activity by allied organizations cannot

automatically be attributed to the Moral Majority. Reagan Presidential Library, campaign correspondence

This example shows how moral reform could target the conditions of cultural visibility. The intended intervention was to influence which conduct received attractive representation, commercial support, and social acceptance.

Education and professional preparation gave the programme a longer horizon than individual elections. Falwell's educational institutions belonged to a wider effort to sustain a religious community's capacity to reproduce itself and participate in public life.

Archival materials include a 1982 programme offering sessions on political journalism, campaign management, congressional staff work, law, lobbying, and political careers. Such training demonstrates an effort to develop practical competence as well as moral conviction. Reagan Presidential Library, political-training programme

Analytically, this addresses a difficulty common to moral movements. Possessing a conviction does not provide the skills required to implement it. Legal drafting, administration, education, campaigning, and media production each have their own demands. Durable influence requires people capable of operating within those institutions. At the same time, professional participation can change the reformers. Lawyers confront precedent and procedure; politicians negotiate coalitions; administrators encounter resource constraints. Institutional entry subjects absolute moral purposes to practical conditions that cannot simply be preached away.

Law was central because the movement understood it as both an instrument of restraint and a source of public legitimacy. A prohibition can make conduct harder or riskier. A legal permission can also influence how people interpret its acceptability. These are distinct effects.

Schaeffer made the second argument especially clearly. In the 1982 sermon version of *A Christian Manifesto*, he argued that judicial legalization could lead people without settled ethical convictions to treat a practice as morally acceptable. He also connected cultural transformation to schools, broadcasting, courts, and professional activity. His account joined personal conversion, spiritual renewal, and social change. These were influential allied arguments, rather than a formally adopted Moral Majority philosophical system. Schaeffer, *A Christian Manifesto*, sermon of 1982

The general mechanism can be analysed without accepting his particular conclusions. Law can affect morality through sanctions, institutional recognition, the allocation of resources, and the messages public authorities convey. But these effects are contingent. Legal authorization does not automatically produce moral approval, and prohibition does not necessarily produce inward conviction. The movement's project consequently depended on an empirical assumption that needed separate examination: changing public rules would help change the moral environment in the desired direction. Electoral success alone could not establish that assumption. School prayer illustrates the importance of distinguishing political descriptions from legal facts. The movement often represented constitutional decisions as the removal of God from public education. *Engel v. Vitale* concerned state-sponsored prayer, not a general prohibition on individual religious belief or every form of private prayer. Those are significantly different propositions. United States Courts, explanation of *Engel v. Vitale*

Its electoral strategy attempted to convert moral commitment into control over authoritative decisions. Candidate selection became a judgment about which officeholders would defend the movement's priorities. The expected sequence connected voting to legislation, administration, and judicial appointments. The 1980 Republican platform demonstrates how related commitments entered a party programme. It supported a constitutional amendment protecting unborn life, restrictions on public abortion funding, voluntary school prayer, parental authority, and assistance for families choosing private education. These were commitments of a larger party coalition, not proof that the Moral Majority alone authored or secured them. Republican Party Platform of 1980

One important consequence was to change the political meaning of an existing moral belief. A voter's opposition to abortion could become a reason to prioritize party affiliation or judicial appointments over other considerations. The belief itself might remain stable while its weight in political judgment increased. Here conscious moral transformation includes the construction of a *hierarchy of concerns*. Leaders do not merely tell supporters what is right. They also indicate which wrongs should determine political allegiance when candidates offer mixed packages of policies.

The claim to constitute a "Moral Majority" combined three forms of authority that should be kept separate. These were the truth of its standards, their alleged acceptance by ordinary Americans, and the legitimacy of pursuing them through democratic institutions.

Each answers a different question:

Claim	Question it addresses	What it cannot establish by itself
The standard is morally true	What ought people to do?	How many people accept it
Most citizens support it	What is socially prevalent?	Whether it is morally justified
A law is democratically enacted	How was a binding decision made?	Whether its content is morally correct

The movement's title made these claims appear mutually reinforcing. Yet their relationship could become problematic whenever electoral results, opinion, and moral conviction diverged. A numerical majority might reject the movement's position. A constitutional limit might prevent the majority from enforcing it. The movement might win an election without obtaining broad agreement on every moral issue. These possibilities show why claims about divine authority, public opinion, and democratic authorization cannot substitute for one another.

The title also had a mobilizing function. It invited people who felt culturally disregarded to imagine themselves as part of a larger public whose influence had not yet been adequately organized. Its success as a political identity does not establish that its entire programme commanded majority support.

The democratic difficulty concerned the justification of coercion in a society containing rival moral authorities. Religious citizens are entitled to organize, argue, and seek office. Their participation is part of democratic life. The further question is how proposed laws should be justified to citizens who do not share their religious premises.

The strongest version of the movement's argument was that public institutions inevitably reflect judgments about desirable conduct. Education, family policy, criminal law, and public expenditure cannot avoid normative choices. Excluding religiously motivated citizens from these debates would itself require justification. That argument does not, however, settle which choices public institutions should make. The existence of moral content in all law does not establish the correctness of one particular religious interpretation. Nor does recognition of a majority's convictions resolve questions about conscience, minority rights, or equal citizenship.

It is therefore more precise to examine specific proposals than to treat the entire movement as either ordinary interest representation or a fully developed theocratic programme. It combined democratic participation with claims about moral authority that could stand above political decisions. The tension lay in how those claims were translated into rules binding everyone.

Its programme also involved the selective political interpretation of Christianity. Biblical commitments do not automatically generate one position on every economic, foreign-policy, or social question.

Falwell's defence of free enterprise, for example, connected economic organization to a religious understanding of responsibility and national prosperity. Samuel Hill's analysis of *Listen, America!* highlights the elevation of particular social and economic arrangements into apparently self-evident moral truths. Hill's contribution to the *Listen, America!* roundtable

This raises two questions. First, which additional arguments are needed to move from a religious principle to a policy? Second, why are some principles politically prominent while others remain secondary?

The questions apply even where participants are entirely sincere. Moral movements inevitably select priorities from a larger inheritance. Their public programme reveals organizational decisions about urgency, political opportunity, and coalition maintenance, alongside theological conviction.

There was also a potential contradiction between criticizing commercial culture for undermining family life and defending market arrangements with limited scrutiny of their cultural effects. A stronger institutional theory would examine how working conditions, advertising, consumption, and economic pressures affect the same families the movement sought to protect.

Participation in politics changed the movement as well as providing it with influence. Matthew C. Moen's *The Transformation of the Christian Right* traces changes in strategy and public language, including increased emphasis on grassroots organization and the vocabulary of rights and freedom. His study concerns the broader Christian Right, but identifies an important process surrounding the Moral Majority's period of activity. University of Alabama Press, description of Moen's study

The theoretical implication is that conscious moral reform is reciprocal. A movement entering public institutions learns which arguments are persuasive, which demands can be negotiated, and which alliances are viable. Its original convictions may remain, while its language and priorities change. This complicates a simple model in which moral conviction first forms and is then implemented. Implementation feeds back into the movement's understanding of what it seeks and how it should act.

Its achievements must be evaluated separately at the levels of organization, voting, law, and moral belief. These outcomes are related, but evidence for one is not evidence for all.

A 2024 NBER working paper by Giulia Buccione and Brian G. Knight examines the shift of evangelical electoral support from Carter toward Reagan. It finds evidence consistent with a role for the Moral Majority, drawing on surveys, television exposure, and exposure to state headquarters. Its model emphasizes religious leaders' influence on followers' assessments of candidates. This supports an account of political persuasion and coordination; it does not demonstrate a comprehensive transformation of private morality. Buccione and Knight, paper abstract and bibliographic record

Later survey evidence also supports the durability of religious political engagement. Pew's 2007 assessment reported that white evangelicals strongly favoured churches expressing views on social and political questions. However, such later attitudes do not isolate the Moral Majority's causal contribution from other organizations and historical developments. Pew Research Center, assessment of Falwell's legacy

The appropriate distinctions are therefore:

Outcome	Appropriate assessment
A more visible conservative religious constituency	A major achievement
Stronger coordination between moral issues and electoral choices	Substantial evidence of influence
Immediate enactment of the full programme	Much more limited
Broad public acceptance of the movement's moral standards	Cannot be inferred from electoral success
Continuing influence after organizational dissolution	Significant, but shared with many later actors

Subsequent developments also require careful attribution. The Supreme Court's 2022 *Dobbs* decision overturned *Roe v. Wade*, an outcome aligned with a longstanding movement objective. But it occurred decades after the Moral Majority dissolved and resulted from a much wider history of political and legal organization. The decision itself establishes a legal change; it does not establish corresponding agreement throughout society. Supreme Court, *Dobbs v. Jackson Women's Health Organization*

The most revealing criticism came from former participants who questioned the connection between political power and moral renewal. Cal Thomas and Ed Dobson's *Blinded by Might* (1999) challenged the expectation that the desired society would follow from electing sympathetic officials. Their criticism retained conservative moral concerns while questioning reliance on political influence. Contemporary account of *Blinded by Might*

Dobson later acknowledged that the movement had entered politics without sufficiently developing its philosophy of political engagement. He described the attractions of access to power, the difficulty of compromise, and his subsequent emphasis on local service, care for disadvantaged people, and personal spiritual transformation. PBS interview with Ed Dobson, 2009

This disagreement concerned the mechanism of change as much as the moral destination. One approach expected organized political power to reconstruct the conditions of moral life. The later criticism placed greater confidence in example, service, relationships, and conversion.

Neither position can be established merely by observing that legislation sometimes disappoints. Law can protect people and alter institutional conduct even when it does not transform motives. Conversely, political access can become a substitute for moral achievement if proximity to officials is mistaken for social renewal. The internal debate exposes precisely this difference.

For a general theory of consciously changing morality, the Moral Majority's greatest significance lies in the distinction between moral validity and organized moral effectiveness. Its participants denied that human decisions created ultimate moral truth. Yet they deliberately built mechanisms to select priorities, circulate authoritative interpretations, organize adherents, train personnel, and influence binding decisions.

This produces a qualified parallel with the movement from customary law to deliberate legislation. The organization sought to make moral standards more explicit and institutionally effective. It did not accept that a legitimate human authority could simply repeal divine morality and replace it with a new code.

Three meanings of the "positivization" of morality should therefore be distinguished:

- **Explicit formulation:** converting dispersed convictions into a public programme.
- **Organized implementation:** creating institutions responsible for advancing that programme.
- **Decision-created validity:** treating morality itself as valid because an authorized body enacted it.

The Moral Majority strongly exemplified the first two. Its declared foundations rejected the third.

Nevertheless, explicit formulation and organization already introduce human selection. Leaders decide which interpretations represent the tradition, which issues receive resources, which compromises are tolerable, and which candidates qualify as acceptable. These decisions can reshape the operative morality of a community even when every participant professes fidelity to unchanged principles.

That is the movement's deepest theoretical lesson: **a programme of moral restoration can generate substantial moral and institutional innovation.** The Moral Majority helped redefine political participation as a religious responsibility, connect separate controversies into a common identity, and develop organizations intended to make moral convictions publicly consequential. Its history also demonstrates the limits of the enterprise: changes in organization, electoral allegiance, and law require separate evidence before they can be called changes in inward conviction or improvements in moral life.

****

Rudolf von Jhering and Nicolai Hartmann

Neither Jhering nor Hartmann can accurately be described as an opponent of all conscious moral change. The qualification is especially important for Jhering: he explicitly advocates criticizing existing moral norms, deliberately reforming harmful customs, and developing a scientifically informed programme of moral education. Hartmann places stronger limits on the authority of moral reformers, particularly on the idea that philosophers can create binding values as legislators create laws. Yet he also recognizes moral discoveries, ethical revolutions, and the legitimate criticism of established morality.

Their relevance to your question therefore lies in identifying **what can legitimately be changed, how such change occurs, and where deliberate moral reconstruction encounters limits**. Reading them together is illuminating, but their objections rest on substantially different foundations.

It helps first to distinguish four meanings of "conscious moral change":

Meaning	The question involved
Changing conduct	Can education, incentives, institutions, and sanctions make people behave differently?
Changing prevailing morality	Can established judgments, obligations, and customs deliberately be revised?
Changing moral understanding	Can people become conscious of values or obligations they previously overlooked?
Creating moral validity	Can an individual or institution make something morally right simply by deciding that it should count as right?

Jhering strongly supports deliberate intervention in the first two areas. Hartmann supports the third and allows consequences for the first two, while rejecting the fourth. The decisive distinction is between **influencing moral life** and **possessing sovereign authority over moral truth**.

Jhering's mature theory begins with the social formation of morality. Rudolf von Jhering—also spelled Ihering—develops the relevant arguments principally in the second volume of *Der Zweck im Recht*, first published in 1883. The page references below concern its second revised edition of 1886. This volume is particularly important because it extends his inquiry beyond law into morality, custom, social expectations, and the formation of moral motivation.

His starting point reverses an influential way of explaining morality. Instead of beginning with an isolated individual's conscience and deriving social obligations from it, he asks how a

historically existing social order produces the conscience of its members. People encounter established judgments about honesty, fidelity, gratitude, assistance, responsibility, and disgrace. Through upbringing and participation in social life, these expectations become internal convictions.

Consequently, the apparent immediacy of conscience does not establish that its contents are independent of history. An individual may sincerely experience an obligation as self-evident while remaining unaware of the social processes through which that conviction developed. In Jhering's account, the objective moral order helps form subjective moral feeling. Jhering, *Der Zweck im Recht*, II, preface and §15.

This argument creates a difficulty for a philosopher who proposes to replace ordinary morality with a system constructed through abstract reasoning. Such a philosopher cannot assume that presenting a superior argument will immediately replace the dispositions through which people judge and act. Moral convictions are sustained by institutions, habits, relationships, and expectations extending far beyond explicit doctrine.

However, **the social formation of morality does not, for Jhering, prove that morality must remain untouched**. It can support precisely the opposite inference: if moral dispositions develop under identifiable conditions, knowledge of those conditions may enable deliberate improvement. This is the point at which a reading of Jhering as an opponent of conscious moral change becomes misleading.

His fundamental explanatory and evaluative category is the maintenance and flourishing of social life. Moral obligations direct individuals toward purposes that exceed their immediate advantage. The individual may have to keep an inconvenient promise, assist someone without expecting repayment, or accept sacrifices for others. Such obligations cannot adequately be explained by saying that each act simply maximizes the agent's personal benefit.

Jhering therefore distinguishes the **social purpose of a norm** from the **subjective motive of the person following it**. A person need not consciously calculate the social consequences of truthfulness whenever speaking honestly. The social function may operate through conscience, affection, a sense of duty, or concern for reputation. Explaining the social function does not mean that individuals consciously pursue that function in every action. Jhering, *Der Zweck im Recht*, II, §§15–18.

This distinction provides a substantial argument against a purely intellectual approach to moral reform. Replacing a justification is not necessarily the same thing as replacing a motive. A population may understand a new ethical doctrine without acquiring the dispositions required to act upon it. Conversely, people may behave responsibly without possessing a philosophical explanation of their conduct.

A reconstruction of Jhering's position would therefore ask reformers two separate questions: *Is the proposed norm socially defensible?* And *what will give it practical force?* The second question cannot be answered merely by repeating the answer to the first.

His differentiation of law, morality, and custom sharpens this analysis. Although his terminology is not entirely uniform, *Recht* concerns the legal order backed by organized coercion; *Sitte* concerns socially obligatory customs and forms of conduct; and morality

includes the assessment of dispositions, intentions, and duties. These domains overlap, but their mechanisms differ.

A legislature can enact a prohibition through an authorized procedure. It cannot, through that act alone, ensure that people cease admiring the prohibited conduct or begin condemning it. Similarly, a philosopher can publish a new moral doctrine without thereby establishing the social authority that makes it a lived norm.

Jhering's discussion of dueling gives a concrete explanation of this difference. He observes that a harmful custom can persist despite both moral criticism and legal prohibition. Individuals may themselves disapprove of the practice yet feel compelled to comply because refusal exposes them to disgrace. The obstacle is a system of mutual expectations: each person anticipates how others will judge him.

The practical lesson is that moral reform may require coordinated changes in social expectations. Jhering accordingly discusses the role of associations in opposing harmful customs. Organized action can give individuals the support that they lack when resisting alone. He does not treat the persistence or antiquity of a custom as proof of its moral legitimacy. Customs may outlive their useful functions, and objectionable customs can also arise through harmful influences from the beginning. Jhering, *Der Zweck im Recht*, II, pp. 252–256.

This is already a theory of deliberate moral and customary reform. It explains why reform through proclamation may fail and why reform through collective organization may succeed. Its objection is to an inadequate understanding of the mechanisms of change.

An apparently more conservative passage occurs on page 107. Here Jhering argues that the state becomes morally criticizable when its measures contradict the moral convictions of the people. Read in isolation, this might suggest that prevailing morality constitutes an unconditional barrier to innovation.

But the passage expressly works in both directions. The state may deserve criticism for introducing institutions that contradict public moral convictions, **or for maintaining institutions that have lost their justification in the people's eyes**. The latter possibility makes moral change a reason for institutional reform. Moreover, this is an argument about the moral assessment of state action; it does not establish that every unpopular law is legally invalid. Jhering, *Der Zweck im Recht*, II, p. 107.

The most decisive evidence appears in his discussion of **"the ethics of the future."** Jhering distinguishes the investigation of existing morality from the subsequent question of whether it should continue unchanged. Once ethics has understood what exists and why, it must ask, in my translation:

"Must it be as it is?" And, if the answer is negative, "How can it be changed?"

He then argues that historical ethics makes possible a criticism of our own moral convictions. If social welfare supplies the relevant standard, inherited convictions and institutions can be examined against it. An affirmative judgment can replace unreflective acceptance with understanding; a negative judgment can disclose their lack of justification and call for ending their continued validity. Jhering, *Der Zweck im Recht*, II, pp. 128–130.

This passage matters because **Jhering is discussing revision of the norms themselves**, not merely stronger obedience to an unchanged morality. His historical approach is intended to make established moral feeling criticizable. The fact that a conviction feels authoritative does not place it beyond examination.

He subsequently extends the programme from criticism to deliberate formation of character. If ethics discovers the processes through which society forms moral willing, it can employ that knowledge practically. He compares this possibility with the way knowledge of the conditions of plant and animal life permits their deliberate cultivation. Ethics would become, in his formulation:

"the most important branch of social policy: national pedagogy."

Here "national pedagogy" means the education and moral formation of a people. Jhering anticipates an increasingly effective practical influence over the development of moral dispositions. He explicitly rejects the expectation that humanity will improve simply by relying on the spontaneous growth of moral life. Improvement requires learning from historical experience and applying those lessons. Jhering, *Der Zweck im Recht*, II, pp. 131–133.

For the history of conscious moral transformation, this places Jhering surprisingly close to programmes of scientifically informed social reform. His sociological account is not merely a warning that morality escapes deliberate control. It is also intended to provide knowledge through which deliberate influence becomes more effective.

What, then, are his genuine limits on conscious reform?

First, reform must understand the social conditions under which moral norms operate. An abstractly attractive rule may be ineffective or harmful if it disregards those conditions.

Second, it must examine what an established norm accomplishes before removing it. Criticism must distinguish between a convention that has become dysfunctional and an obligation that sustains indispensable forms of cooperation.

Third, reform must address motivation and collective expectations. Demonstrating that a custom is objectionable does not automatically free individuals from its social pressure.

Fourth, reform should proceed with historically limited knowledge. On page 134, Jhering says that ethics should help humanity reach the next attainable point; remote destinations remain beyond its present field of vision. This is a restriction on comprehensive utopian planning, combined with support for practical advancement. Jhering, *Der Zweck im Recht*, II, p. 134.

The principal unresolved difficulty in this position concerns **who identifies society's genuine welfare**. Existing moral convictions may command public allegiance, while a reformer claims that they damage society. Jhering's appeal to socially established morality and his appeal to critical assessment through social welfare can then point in different directions.

His account also leaves a potential gap between explaining a norm's social function and justifying it morally. A practice may preserve an existing social order while benefiting some groups at others' expense. A satisfactory evaluation must therefore ask whose welfare counts, how burdens are distributed, and whether the preservation of an institution is itself desirable.

These are analytical criticisms of his framework. They show why a social explanation of morality does not by itself provide an uncontested mandate for either conservative preservation or deliberate reconstruction.

Hartmann's objection goes deeper into the question of moral authority. Nicolai Hartmann's principal treatment appears in *Ethik* (1926), supplemented by *Das Problem des geistigen Seins* (1933). His central distinction is between values, human awareness of values, and the historically existing moralities through which people interpret and pursue them.

For Hartmann, values are not made valid by individual preference, collective acceptance, state enactment, or philosophical invention. Their validity is independent of whether a particular society adequately recognizes them. This does not mean that our knowledge of values is complete or infallible. On the contrary, the difference between values and our knowledge of them explains the possibility of moral error.

That position immediately prevents prevailing morality from becoming the final authority. An established community can overlook a value or distort its significance. A dissenter may therefore recognize something that the surrounding society fails to recognize.

It also prevents the reformer from becoming the final authority. The fact that someone proposes a new morality does not establish that its content is valid. Both inherited morality and innovative doctrine remain answerable to something beyond their social existence.

Hartmann explicitly considers the analogy between moral philosophy and legislation. In the opening chapter of *Ethik*, he examines the proposition that moral commandments are human products. If this were accepted in the relevant sense, philosophical thought might claim the power to establish and overthrow commandments just as political thought establishes laws. Positive law and positive morality would then stand on the same level.

His immediate objection is that philosophical theories are themselves historically variable and mutually conflicting. A philosopher cannot demonstrate the relativity of ordinary morality and then silently exempt his own doctrine from the same problem. The replacement system would risk reproducing the unsupported claim to absolute authority that it criticized in inherited norms. Hartmann, *Ethik*, chapter 1, especially §b.

This is one of Hartmann's strongest arguments against the conscious manufacture of morality. It challenges an asymmetry of authority:

- Existing morality is treated as historically conditioned and therefore revisable.
- The reformer's principles are treated as universally authoritative.
- Yet the reformer has not explained why his own standpoint escapes historical limitation.

The criticism does not show that reformers are necessarily mistaken. It shows that **their authority requires justification beyond their capacity to criticize their predecessors**.

Hartmann then rejects the opposite extreme. If ethics cannot legislate moral validity, it does not follow that philosophy must become practically powerless. Different moralities advance conflicting claims. Philosophy has the task of examining those claims, distinguishing insight from unwarranted authority, and clarifying their grounds.

Accordingly, when he says that ethics cannot be "positive legislation," he is restricting the source of its authority. Philosophical ethics can affect conduct through knowledge and criticism; it cannot make its objects morally valid by promulgating them. Hartmann, *Ethik*, chapters 1c–3.

His account of moral education makes this restriction more precise. Hartmann rejects the simple identification of virtue with knowledge. Knowing what is right does not guarantee doing it. People can understand an obligation and nevertheless violate it.

But the insufficiency of knowledge does not make knowledge irrelevant. Ethical understanding can bring a person before a decision that would otherwise remain unseen. It can disclose another person's claim, reveal the significance of a situation, or expose a conflict that habitual judgment concealed.

Hartmann therefore describes philosophical ethics as a **maieutics of moral consciousness**: an activity analogous to intellectual midwifery. The teacher helps another person arrive at insight, but cannot replace that person's apprehension with obedience to the teacher's authority.

His comparison with mathematical teaching is instructive. A teacher can direct attention to a geometrical relationship, but the student must come to see why it holds. Likewise, moral education can direct attention, challenge blindness, and encourage reflection, but an imposed verbal formula is not equivalent to genuine moral understanding. Hartmann, *Ethik*, chapters 2d–3a.

This gives Hartmann a substantial objection to programmes that identify moral improvement with the successful transmission of officially approved propositions. Such programmes may change what people say, what they publicly endorse, or what they do under supervision. Whether they improve moral judgment remains a separate question.

This application follows from his theory; it should not be mistaken for a historically specific judgment by Hartmann on contemporary educational or political programmes.

His discussion of invented values addresses the psychological limits of moral reconstruction. In chapter 5, Hartmann argues that the problem has not been a shortage of imaginative moral proposals. People can invent numerous systems. The difficulty is that what is invented does not necessarily convince the living sense of value.

He understands *Wertgefühl*, the feeling or apprehension of value, as a primary responsiveness to what matters morally. It is not simply an arbitrary preference, nor is it identical with an intellectual deduction. A person can be moved by the significance of loyalty, cruelty, betrayal, or generosity before possessing an explicit theory of those experiences.

Consequently, abstract invention cannot assume unlimited power over moral consciousness. Moral feeling has a structure and resistance of its own. A doctrine that is formally consistent may fail to disclose anything that people can recognize as genuinely valuable. Hartmann, *Ethik*, chapter 5c.

There is a vulnerability in this argument. Resistance to a new doctrine does not by itself establish the doctrine's falsity. People can resist moral insight because of prejudice, interest, or habituation. Hartmann's broader position must therefore distinguish between resistance arising

from genuine moral perception and resistance arising from moral blindness. The mere fact that a reform meets opposition cannot settle the issue.

His account of ethical revolution supplies the necessary qualification. **Hartmann explicitly recognizes that moral consciousness is continually changing.** New conflicts and situations disclose previously unnoticed aspects of value. Beneath apparently stable moral terminology, people may gradually change what they perceive, admire, condemn, and consider obligatory.

The words can remain the same while their understood content changes. Concepts and formulas often lag behind the finer movement of moral experience. Thus the historical stability of a moral vocabulary does not prove the stability of the ethos expressed through it. Hartmann, *Ethik*, chapter 6a.

He describes this through the limited field of human value-consciousness. At a given time, people apprehend only part of the realm of values. As attention shifts, previously neglected values become visible while others recede.

Three consequences follow. A moral innovation may constitute a genuine discovery. A later morality is not automatically superior simply because it is later. And an advance in one dimension can coincide with diminished awareness in another.

For example, a society might become more sensitive to individual independence while becoming less attentive to obligations of care. Another might deepen solidarity while neglecting personal freedom. These examples illustrate the structure of Hartmann's argument; they are not his verdicts on particular contemporary societies.

The point is not that every advance must entail an equivalent loss. It is that **the changing focus of moral consciousness does not guarantee comprehensive progress**. A responsible reformer must therefore ask both what a proposed change reveals and what it may obscure.

Hartmann's treatment of moral leaders is especially relevant to the possibility of conscious transformation. Prophets, religious founders, and other bearers of ideas can give decisive expression to emerging moral awareness. Their intervention can precipitate a historical crisis and transform the judgments of large populations.

Yet Hartmann denies that their effectiveness ordinarily consists in creating values from nothing. They articulate what is already developing, often obscurely, in the moral experience of others. Their words provide a form in which people recognize concerns that they could not previously express.

The relation is therefore reciprocal. The leader gives conscious expression to an emerging ethos, while the responsiveness of others supplies the conditions under which that expression becomes historically powerful. Hartmann's image is that the spark succeeds because combustible material is already present. Hartmann, *Ethik*, chapter 6b.

This supports a limit on deliberate moral transformation: **the reformer does not possess unilateral command over the conditions of reception**. Intellectual ingenuity, institutional position, and rhetorical power cannot simply substitute for the moral experiences through which a new demand becomes intelligible.

Nevertheless, Hartmann also recognizes the misunderstood innovator whose insight comes before others are ready to receive it. Such a person may remain isolated, and recognition may occur only later. Historical success is therefore not identical with moral truth. His theory of social receptivity cannot be reduced to the proposition that whatever wins mass acceptance must be right.

Nor does he confine philosophy to recording discoveries already made elsewhere. In chapter 6c, he explicitly calls for a forward-looking, searching ethics alongside retrospective explanation. Philosophy should investigate values for whose recognition people may already be becoming ready. What he rejects is the ambition of a single thinker to complete this undertaking through a comprehensive personal construction. He assigns it instead to sustained, cooperative inquiry extending across generations. Hartmann, *Ethik*, chapter 6b–c.

A further limit arises from the plurality of values. Hartmann warns that a value can become "tyrannical" when it dominates moral attention and excludes other legitimate demands. Justice can become insensitive to love; devotion to future humanity can become hostility toward presently living persons. His response is to seek relationships and syntheses among values. He also distinguishes their height from their strength: the pursuit of higher achievement does not simply cancel the requirements that sustain moral life.

Personal freedom imposes another restriction. Moral action requires the person's own determination; values must not function as irresistible forces that mechanically produce compliance. Consequently, a programme's success in producing prescribed conduct would not, by itself, establish that it had created responsible moral agency. That final implication is an application of his account of freedom. Stanford Encyclopedia of Philosophy, "Nicolai Hartmann," section 4.1.

His later account of shared morality adds a social dimension. In *Das Problem des geistigen Seins*, morality belongs to "objective spirit": historically existing cultural life sustained by individuals but exceeding any one individual's knowledge or control. Moral development occurs through participation, experience, and encounters with other people, as well as through explicit instruction.

Crucially, however, collective morality has no personal conscience. Individual persons must experience and judge the consequences of applying its rules. This leaves room for conscience to challenge established conventions. Social embeddedness therefore explains both the authority of inherited morality and the difficulty of reform; it does not make every inherited rule morally conclusive. Moritz von Kalckreuth's recent study develops this aspect of Hartmann through the discussions of moral learning and conscience in the 1933 work. Kalckreuth, "Common Morality as a Part of the Objective Spirit".

The resulting comparison is more revealing than a common classification as opponents of moral change:

Issue	Jhering	Hartmann
Principal foundation	Social purposes and the historical formation of norms and motivation	Objective values and their incomplete human apprehension
Meaning of moral objectivity	A social normative order exceeding individual preference	Values whose validity does not depend on human acceptance

Issue	Jhering	Hartmann
Criticism of established morality	Evaluates convictions and institutions through their contribution to social welfare	Evaluates prevailing morality through a wider and more adequate recognition of values
Deliberate moral education	Explicitly advocates scientifically informed social education	Supports education that awakens understanding and preserves personal judgment
Revision of prevailing norms	Explicitly permissible and sometimes required	Permissible when it corrects blindness, distortion, or inadequate interpretation
Main limit on reform	Ignorance of social conditions, functions, motivation, and attainable possibilities	Confusing invention with discovery, or imposed conformity with moral insight
Status of the reformer	A practically engaged investigator who still must justify his account of social welfare	A fallible interpreter who cannot confer validity on values by decree

Their different meanings of "objective" are particularly important. **Jhering's objective moral order is historically constituted social reality; Hartmann's objective values are not constituted by society.** Treating those two claims as equivalent would conceal the central philosophical disagreement between them.

A Hartmannian criticism of Jhering would ask whether the successful preservation and flourishing of society can adequately determine moral worth. Could an efficiently organized society nevertheless be morally deficient? Could moral obligations require criticism of purposes that a society regards as essential to itself?

A Jheringian criticism of Hartmann would press in another direction. How does an alleged discovery of an objective value acquire reliable social meaning? What produces the dispositions needed to realize it? How can we distinguish genuine insight from the projection of historically formed preferences?

These are reconstructed questions, rather than evidence of a direct controversy between the two thinkers. They show that their theories can correct different weaknesses in a programme of moral reform.

Their arguments also leave unresolved the institutional question of moral authority. Jhering does not supply an uncontested procedure for deciding what society's true welfare requires. Hartmann does not supply an uncontested procedure for deciding between competing claims to have discovered objective values. Neither theory warrants the simple conclusion that moral philosophers, political leaders, or prevailing majorities should possess final authority over all moral disagreement.

For your specific inquiry into morality becoming deliberately revisable in a manner analogous to positive law, the strongest textual contrast is therefore this: **Jhering expressly envisages conscious criticism, termination of unjustified norms, and organized moral education; Hartmann expressly denies that philosophical ethics can acquire the authority of positive legislation.** Hartmann's denial nevertheless leaves considerable room for deliberate criticism, discovery, education, and transformation of moral life.

For direct study, the most relevant passages are:

- **Jhering, *Der Zweck im Recht*, II, second edition, 1886:** p. 107 on state action and public moral convictions; pp. 128–134 on critical and practical "ethics of the future"; pp. 252–256 on harmful customs and organized reform. Primary text.
- **Hartmann, *Ethik*:** chapters 1–3 on the authority and educational role of ethics; chapters 5–6 on invention, discovery, and ethical revolution; chapters 15–16 on values and historical value-consciousness; chapter 61g on the tyranny of values. Chapter references avoid differences in pagination between editions. German text.
- **Hartmann, *Das Problem des geistigen Seins*:** chapter 22 on prevailing morality and individual morality; chapters 35–38 on conscious leadership, limits of power, collective error, and public opinion; chapter 58 on the tension between living ethos and fixed cultural formulations.