

Forthcoming in Marcus Rossberg (ed.), *The Cambridge Handbook of Analytic Philosophy*.
Cambridge: Cambridge University Press.

Philosophical Naturalism and Empirical Approaches to Philosophy

Jonathan Y. Tsou

Iowa State University

This chapter examines the influence of the empirical sciences (e.g., physics, biology, psychology) in contemporary analytic philosophy, with focus on philosophical theories that are guided by findings from the empirical sciences. Scientific approaches to philosophy follow a tradition of philosophical naturalism associated with Quine, which strives to ally philosophical methods and theories more closely with the empirical sciences and away from *a priori* theorizing and conceptual analysis.

In contemporary analytic philosophy, ‘naturalism’ is an ambiguous and equivocal term (Papineau, 2020) that can be distinguished into weaker and stronger methodological commitments:

- N1.** Philosophy should be *constrained* by scientific results. Philosophical theories should not be inconsistent with the findings of empirical science (e.g., the positing of supernatural entities).
- N2.** Philosophy is *continuous* with science. Philosophical standards (e.g., the assumption that knowledge is fallible) and methods (e.g., empirical and experimental methods) should not be different in kind from those adopted in the natural sciences. Moreover, genuine philosophical problems should be tractable with naturalistic empirical methods.
- N3.** Philosophy should be empirically driven. Philosophical theorizing should be *guided* by the results of science and empirical science provides the most promising route to formulating sound philosophical theories.

N1 implies that philosophical theories should be consistent with scientific theories. *N2* implies that philosophical standards and methods should be continuous with those adopted in science. *N3* implies that the empirical scientific findings should be utilized to direct philosophical inquiry. Whereas *N1* is a platitude among many contemporary analytic philosophers, fewer are

committed to *N2* or *N3*. This chapter examines philosophical theories (e.g., theories of mind and ethics) that are committed to *N2* and *N3*, with particular emphasis on *N3*.¹

Quine's Naturalism and its Legacy

In the twentieth century, the logical empiricists (e.g., Carnap, Reichenbach, Hempel), who founded the field of analytic philosophy of science, aimed to replace traditional (i.e., metaphysical) philosophy with *scientific philosophy*.² While Quine's rejection of the analytic-synthetic distinction (Quine, 1951) is regarded as one of the decisive criticisms that led to the demise of logical empiricism, Quine should also be regarded as one of the most influential logical empiricists and most fervent advocates of scientific philosophy. One of Quine's most lasting influences on 20th century analytic philosophy is his advocacy of a naturalistic methodological approach to philosophy.

Quine's argument for naturalized epistemology recommends *replacing* traditional (i.e., foundationalist) epistemology with a scientific project (Quine, 1969a).³ He describes this project as follows:

Epistemology . . . is science self-applied. It is the scientific study of the scientific process. It explores the logical connections between the stimulation of the scientist's sensory receptors and the scientist's output of scientific theory. (Quine, cited in Pyle, 1999, p. 20)

In this formulation, 'stimulation of sensory receptors' is shorthand for 'evidence,' while the 'output' is shorthand for 'theory.' On Quine's understanding, the most central goal of epistemology is to explain how evidence relates to theory. He argues that the natural science (i.e., empirical psychology) is the most promising method to achieve this goal. In a frequently cited passage, Quine (1969a) writes:

Epistemology, or something like it, simply falls into place as a chapter of psychology and hence of natural science. It studies a natural phenomena, viz., a psychological human subject. This human subject is accorded a certain experimentally controlled input ['evidence'] . . . and in the fullness of time the subject delivers as output a description of the three-dimensional external world ['theory'] . . . The relation between the meager input and the torrential output is a relation that we are prompted to study for somewhat

¹ In the history of philosophical naturalism, Dewey is one of the earliest and most prominent advocates of *N2* in his rejection of lofty philosophical ideals of knowledge (e.g., Descartes' ideal of knowledge *that is known with absolute certainty*) in favor of fallibilist ideals that are continuous with those adopted in natural science (Dewey, 1929, 1944). For a broader discussion of Dewey's pragmatic naturalism, see Godfrey-Smith (2002, 2014), Howard (2003), Kim (2003), Brown (2012), and Pearce (2020).

² For discussion of the history of scientific philosophy and logical empiricism, see Cartwright et al. (1996), Giere & Richardson (1996), Richardson (1997, 2008), Friedman (1999, 2001), Hardcastle & Richardson (2003), Reisch (2005), Richardson & Uebel (2007), Friedman & Creath (2007), Uebel (2007), Stadler (2015); and Creath (2021).

³ It is worth distinguishing Quine's "naturalized epistemology" from his "naturalized empiricism." The former is Quine's preferred *methodology* for studying knowledge, while the latter is his resulting holistic ('web-of-belief') view wherein all *knowledge* is empirical and revisable (Gibson, 1995). This chapter focuses on the former. For critical discussion of Quine's holistic and empiricist theory of knowledge, which denies the existence of *a priori* knowledge, see Putnam (1962a, 1962b), Boghossian (1996), Devitt (1996), Friedman (1997, 1999, 2001), Rey (1998), Harman (1999), and Tsou (2003a, 2010).

the same reasons that always prompted epistemology . . . to see how evidence relates to theory. (Quine, 1969a, p. 83)

Quine contends that the *explananda* of naturalized and traditional epistemology are unified in the relationship between evidence (i.e., sensory input) and theory (i.e., linguistic output). Quine's naturalism breaks from traditional epistemology in attempting to understand this relationship from the perspective of natural science, rather than from an Archimedean perspective. On this view: "Epistemology . . . is not logically prior somehow to . . . the refined common sense which is science; it is part of the overall scientific enterprise, . . . which Neurath has likened to that of rebuilding a ship while staying afloat in it" (Quine, 1960, p. 253). Quine (1974) contends that this naturalistic shift is "no gratuitous change of subject matter but an enlightened persistence in the original epistemological problem" (p. 3).

Quine's advocacy of naturalized epistemology is motivated by the failure of foundationalist epistemologies. For Quine, foundationalist epistemology is divided into: (1) a doctrinal tradition that is concerned with truth, and (2) a conceptual tradition that is concerned with meaning. These two forms of reductionism are linked insofar as they aim to *maximize epistemic certainty*. Quine identifies Descartes as exemplifying (1), and he contends that Hume has destroyed the prospects of identifying an indubitable foundation for knowledge. Quine identifies Carnap's epistemological project in the *Aufbau* (Carnap, [1928] 1967) as exemplifying (2), and he argues that Carnap's attempts to translate (or 'rationally reconstruct') sense experience into logical terms fail and Carnap's more liberal reconstructions (Carnap, 1936, 1937) fare no better.⁴ Quine (1969a) writes:

If all we hope for is a reconstruction that links science to experience in explicit ways short of translation, then it would seem more sensible to settle for psychology. Better to discover how science is in fact developed and learned than to fabricate a fictitious structure to a similar effect. (p. 78)

Based on the failure of foundationalist approaches, Quine suggests that we should give up on attempts to reduce knowledge to a (rationalist or empiricist) foundation of certainty and revise the methods of epistemology. If epistemology aims to clarify the relationship between theory and evidence, empirical psychology—and behaviorist psychology in particular (Gibson, 2004a)—offers a more promising method for achieving this goal.

Quine's naturalized epistemology corresponds to a naturalistic approach to ontology, which recommends a realist and fallibilist stance towards posited entities (Glock, 2003). For Quine, ontology is the theory of what there is, and his naturalistic approach involves an investigation into the *ontological commitments* of an accepted body of knowledge (Quine, 1948, 1969b, 1976). Quine recommends that we accept the existence and reality of whatever is *indispensable* (e.g., 'atoms,' 'molecules,' 'tables') in our best scientific theories (i.e., whatever entities the bound variables of the theory must refer to in order to make the theory true)—this criterion being subject to the constraint of *simplicity*, which is meant to prohibit ontological extravagance. Hence, Quine's approach to ontology is inflationary insofar as it recommends a realist stance towards posited entities; however, it is deflationary insofar as commitment to the reality of ontological posits is revisable in the face of new data.

⁴ Quine (1969a) maintains that Carnap's translational projects are unfeasible due to indeterminacy of translation and underdetermination considerations (pp. 79-82). Friedman (1999) and Richardson (1998) argue that Quine misrepresents Carnap's *Aufbau* as a defense of empiricist foundationalism (cf. Tsou, 2003b).

Many philosophers have been unconvinced by Quine's argument for naturalized epistemology (Rysiew, 2020). A common complaint is that Quine appears to be recommending a shift away from *normative epistemology* towards *descriptive scientific theories* (Putnam, 1982). For example, van Fraassen (1995) argues that traditional epistemology is concerned primarily with *rationality and justification* (not merely the relationship between theory and evidence), and Quine has not provided persuasive reasons for abandoning these *normative ideals*. Moreover, Quine's presentation of naturalized epistemology as a *replacement* for traditional epistemology appears to be unmotivated since each respective program addresses fundamentally different questions (Rorty, 1979, ch. 5; Stroud, 1984, ch. 6), despite Quine's claims to the contrary. For example, Kim (1988) writes:

None of us . . . would want to quarrel with Quine about the interest or importance of the psychological study of how our sensory input causes our epistemic output . . . What is mysterious is why this recommendation has to be coupled with the rejection of normative epistemology . . . [I]t is difficult to see how an "epistemology" that has been purged of . . . an appropriate normative concept of justification or evidence, can have anything to do with the concerns of traditional epistemology. And unless naturalized epistemology and classical epistemology share some of their central concerns, it's difficult to see how one could *replace* the other, or be a . . . better way . . . of doing the other. (p. 391, emphasis in original)

These considerations suggest that Quine has provided poor reasons for adopting naturalized epistemology, and his conclusion that epistemologists ought to 'settle for psychology' is confused. More sympathetic commentators interpret Quine's 'replacement' claim as the argument that normative philosophical ideals (e.g., epistemic justification) cannot be articulated without consulting scientific theories (Kornblith, 1994a) or that abstract philosophical normative ideals should be replaced with scientific normative ideals (Gibson, 2004c).

While few philosophers have followed Quine's specific recommendation of providing a psychological (i.e., behaviorist) explanation of how scientists' sensory inputs (i.e., 'evidence') relate to their linguistic outputs (i.e., 'theories'), many have been sympathetic with Quine's more general call for naturalistic methodological approaches to philosophy. Some philosophers endorse Quine's argument—encapsulated in *N2*—that philosophy is continuous with science in terms of its standards and methods (Haack, 1990, 1993; Kitcher, 1992; Kornblith, 1994b; Rouse, 2002; Maddy, 2007; Roth, 2007, 2008; Wimsatt, 2007). This stance rejects the ideal of philosophy—associated with Frege ([1884] 1950) and Wittgenstein (1921, §4.111-4.11)—as an *autonomous* discipline operating independent of science (Kitcher, 1992). Accordingly, naturalists maintain that the *a priori* ('armchair') methods employed in much of traditional philosophy should be replaced (or at least supplemented) with the findings of empirical science. If philosophy is not distinguished by some infallible *a priori* philosophical method (e.g., the Socratic method, the Cartesian method of doubt, logical analysis, conceptual analysis, reflective equilibrium), then philosophers should place greater priority on scientific and empirical evidence in formulating philosophical theories.

While naturalism has been most influential in epistemology and philosophy of science,⁵ scientific approaches to philosophy have been adopted in other fields. Some naturalists follow

⁵ For discussion of Quine's influence in epistemology, see Kornblith (1994b) and Rysiew (2020). The 'naturalistic turn' in philosophy of science (e.g., see Giere, 1985; Boyd, 1980; Callebaut, 1993)—which accompanied the decline of logical empiricism—was as much due to Kuhn (1962) as it was to Quine. Compared to Carnap's preferred method of logical analyzing scientific languages ('logic of science'), Kuhn argues that science should be analyzed as

Quine's recommendation—encapsulated in *N3*—that philosophical inquiry should be guided by empirical science. For example, some philosophers appeal to theories of physics to defend metaphysical views. In a polemical book, Ladyman et al. (2007) defend a naturalistic approach that demands that genuine metaphysical claims are supported by well-confirmed scientific theories (including at least one fundamental theory of physics). Dowe (2000) defends a 'conserved quantity' theory of causation derived from conservation laws (e.g., energy, momentum) in physics. Ismael (2013, 2016) defends a compatibilist account of free will that accommodates the results of physical theories (e.g., Newtonian mechanics, Quantum mechanics). Other philosophers appeal to the results of biology and cognitive science to formulate metaphysical and epistemological positions. Wimsatt (forthcoming) defends a scientific approach to metaphysics based on biological practice. Bickle (2003) articulates materialist positions on issues in philosophy of mind (e.g., multiple realization, mental causation) through a close examination of findings from cellular and molecular neuroscience (e.g., the role of long-term potentiation in memory). Millikan (1984) and Neander (2017) defend theories of mental content and intentionality derived from evolutionary theory. Others appeal to findings from social sciences (e.g., psychology, economics) and other sciences to articulate theories in various philosophical fields. Findings from cognitive science (e.g., connectionism) and computer science have been central in the development of computational theories of mind (Rescorla, 2020). Pietroski (2018) draws on findings from linguistics to articulate an internalist theory of meaning. Gauthier (1986) derives a contractarian theory of ethics ('morals by agreement') based on economic theories (e.g., decision theory, game theory). Based on scientific findings from a wide range of disciplines (e.g., cognitive science, social psychology, anthropology, psychiatry), Prinz (2007) argues that human moral judgments are culturally conditioned emotional responses, which is a perspective he appeals to in defending moral relativism.

Empirical Approaches to Philosophy

Three salient examples of empirical approaches to philosophy that are guided by empirical science are: (1) eliminativist approaches to philosophy of mind, which are guided by neuroscience, (2) theories of evolutionary ethics, which are guided by Darwin's theory of natural selection, and (3) empirical approaches to moral psychology, which are guided by social and cognitive psychology. Examination of these views elucidates the motivations of *N3* and the challenges facing empirical approaches to philosophy.

Eliminative Materialism

Eliminative materialism is a naturalistic view on mind—most closely associated with the Churchlands (Churchland, 1981, 2013; Churchland, 1986)—that follows Quine's radical recommendation of *replacing* philosophical theories formulated from the armchair with empirical theories informed by the results of science. Eliminativists argue that traditional philosophical analyses are impoverished insofar as they are guided and informed by *commonsense* psychological theories ('folk psychology') rather than the empirical results of neuroscience. From a methodological perspective, eliminativists argue that folk psychological

a *natural entity or processes* (Bird, 2002; Tsou, 2015). In post-positivist philosophy of science, the naturalistic turn manifested itself in a return to realist accounts of science, the development of naturalistic and causal accounts of scientific knowledge, and the requirement that philosophical analyses of science must be closely engaged with the science under analysis, including its history (Boyd, Gasper, & Trout, 1999. introduction).

concepts should be replaced by neuroscientific concepts in an adequate theory of mind: “our commonsense conception of psychological phenomena constitutes a radically false theory, a theory so fundamentally defective that both the principles and the ontology of that theory will eventually be displaced . . . by completed neuroscience” (Churchland, 1981, p. 67).

While the eliminative materialism defended by the Churchlands is often interpreted as a metaphysical theory about which mental states (e.g., ‘beliefs,’ ‘desires’) exist (e.g., see Ramsey, 2021), it is more charitably understood as a *methodological argument*: philosophical theories of mind should be guided by the results of empirical science (viz., neuroscience), rather than commonsense theories (viz., folk psychology). This methodological stance assumes that: (1) folk psychology is a misleading and likely false *theory*, and (2) in an adequate theory of psychology, folk psychological concepts will be “substantially revised or replaced outright” (Churchland, 1986, p. 396). As Paul Churchland (2013) puts it:

Modern theories of mental illness led to the elimination of witches from our serious ontology. . . . The concepts of folk psychology—belief, desire, pain, joy, and so on—await a similar fate. . . . [W]hen neuroscience has matured to the point where the poverty of our current conceptions is apparent to everyone . . . , we shall then . . . set about reconceiving our internal states and activities . . . Our explanations of one another’s behavior will appeal to such things as our neuropharmacological states, our high-dimensional prototype representations, and the activation-patterns across specialized brain areas. (p. 76, emphasis in original)

From a methodological point of view, eliminative materialism amounts to the objection that traditional philosophical accounts of mind violate *NI* insofar as they are constrained by commonsense (i.e., folk psychology), rather than science. Reasons for thinking folk psychology is false include its poor historical record, its impoverished explanatory power, and the falsity of other commonsense theories (Churchland, 2013, ch. 2). In contrast to productive scientific paradigms, folk psychology is a theoretical framework that has not been elaborated nor revised for over 2000 years. Although folk psychology postulates the existence of distinctive mental states (e.g., ‘beliefs,’ ‘sleep,’ ‘memory’), it only provides superficial explanations of these concepts and has very limited predictive power, especially when compared to the complex explanations and precise predictions found in neuroscience. Moreover, folk theories in other domains have turned out to be false. For example, folk physics suggests that heavier objects fall at a faster speed than lighter objects; however, classical mechanics indicates that this commonsense belief is false. Folk biology suggests that whales are fish; however, modern theories of biological taxonomy indicate that whales are mammals. The falsity of these commonsense theories suggest that folk psychology will also turn out to be false.

For the Churchlands, traditional metaphysical theories of mind are misguided insofar as they aim to account for *distinctive psychological (or ‘mental’) states* implied by our ordinary psychological terms (e.g., ‘perception,’ ‘belief,’ ‘desire,’ ‘pain,’ ‘memory,’ ‘fear’). Dualists (e.g., Chalmers, 1996) argue that that these ‘mental states’ possess special metaphysical properties (e.g., they are non-physical, they are only accessible in introspection) that differentiate them from physical states. The main challenge facing dualist accounts (e.g., causal interactionism, epiphenomenalism) is the mind-body problem: how are (non-physical) mental states causally related to (physical) brain states? For identity theorists (e.g., Place, 1956; Smart, 1959), the mind-body problem is a pseudo-problem since psychological states are assumed to be *identical* to brain states. However, identity theorists (and materialists more generally) face the challenge of explaining how apparently non-physical psychological states (e.g., the perception of red, the

belief that it is raining) are *actually* brain states. This problem is addressed by showing how psychological states (e.g., ‘pain’) can be *reduced* to brain states (e.g., c-fiber firings). This framework suggests a generic method for evaluating the correctness of dualism against identity theory. If different folk psychological states (e.g., ‘pain,’ ‘memory,’ ‘beliefs’ ‘intention’) are reducible to brain states, then identity theory is correct; if these states are irreducible, then dualism is correct. Eliminativists reject this framework for its assumption that folk psychology provides an accurate taxonomy of psychological states. On behalf of materialism, eliminativists argue that irreducibility of a folk psychological concept should not count as evidence in favor of dualism. If folk psychology provides a misleading or false theory of our inner psychological states, then questions about reducibility should be framed with reference to targets of reduction (i.e., distinct psychological states) individuated by neuroscience, not folk psychology. Against the identity theorists’ strategy of *reducing* folk psychological concepts to brain states, eliminative materialists recommend ignoring folk psychology altogether. In the correct theory of mind, commonsense concepts of mentality (e.g., ‘pain,’ ‘memory’) may need to be radically revised, and some folk psychological concepts (e.g., ‘beliefs,’ ‘intentions’) may be eliminated entirely. Along these lines, Stich (1983) argues that the folk concept of ‘belief’ has no place in the scientific study of psychology, Griffiths (1997) argues that folk concepts of ‘emotions’ should be replaced with neuroscientific concepts, and Hardcastle (1999) argues that the folk concept of ‘pain’ fails to capture the neurobiological complexity of pain sensations.

The Churchlands are also suspicious of theories of mental content (e.g., propositional attitudes) that are derived from folk psychological concepts. Besides postulating the existence of distinctive psychological states, folk psychology suggests that some psychological states (e.g., ‘beliefs,’ ‘desires’) have special philosophical significance for addressing questions concerning how the mind represents the world. These psychological states or ‘propositional attitudes’ are special because they exhibit *intentionality* (i.e., they represent or are *about* a particular state-of-affairs), and thus have semantic content. Because of their representational capacity, propositional attitudes play a central role in folk psychological theories that explain and predict the behavior of individuals in terms beliefs and desires (Dennett, 1987). The Churchlands regard the propositional attitudes paradigm as a dressed-up version of folk psychology. Paul Churchland (1981) argues that if folk psychology is a false theory, propositional attitudes might not exist at all, at least not in the way assumed in folk psychology. Patricia Churchland (1986, ch. 9) argues that sentential attitudes do not appear to play any crucial role in cognition, and she eschews the modeling of cognitive structures on linguistic structures (e.g., Fodor, 1975). Others argue that contemporary models in cognitive science explain cognition without appealing to or postulating the sorts of representational entities assumed in traditional theories of mental content (Ramsey, Stich, & Garon, 1990; Chemero, 2009). These analyses indicate how folk psychological concepts like ‘belief’ and ‘desire’ could be eliminated in a scientific theory.

As a naturalistic methodology, eliminative materialism places more trust in empirical science than commonsense. More generally, eliminativists are opposed to the reliance on commonsense intuitions in philosophy. For example, metaphysical dualism is driven by the Cartesian *intuition* that psychological states (e.g., beliefs, pains) are *fundamentally different* from physical objects or properties (e.g., tables, chairs). Whereas introspection intuitively suggests that psychological states are non-physical, science often reveals that commonsense intuitions and introspection are wrong (Churchland, 2013, pp. 24-25). Accordingly, philosophers should place much less evidential weight on commonsense intuitions and more on scientific findings. This naturalistic stance implies that several influential problems in philosophy of mind are misguided

(Churchland, 1996). For example, the ‘hard problem’ of consciousness (i.e., explaining why and how experience arises from a physical basis) defended by Chalmers (1995) may be a pseudo-problem that ultimately rests on the intuition that experience is an irreducible fundamental property of the world (Dennett, 1995; Tsou, 2013). The related argument (Nagel, 1974; Chalmers, 1996) that a satisfactory explanation of consciousness must explain its qualitative and subjective features violates *N2* by presupposing an unreasonably lofty philosophical standard (which transcends scientific standards of explanation) and rests on the intuition that the most important features of consciousness are its qualitative aspects. Against this view, some (Dennett, 1988; Frankish, 2016) argue that the qualitative features of consciousness (i.e., ‘qualia’) should be eliminated, rather than explained.

Evolutionary Ethics

In contemporary ethics, few canonical normative theories (e.g., utilitarianism, Kantianism, virtue theory, contractarianism) are formulated based on scientific findings. Theories of evolutionary ethics are an exception to this rule. Whereas eliminative materialists recommend that theories of mind should be informed, generally, by the results of neuroscience (rather than commonsense), evolutionary ethicists formulate normative theories and meta-ethical positions based on Darwin’s theory of the evolution of human morality.

In *Descent of Man*, Darwin (1871) argues that the mental faculties or ‘powers’ (e.g., intelligence, emotions, imitation, memory, learning) of humans and other animals differ in degree, not in kind (ch. 2); and he devotes a chapter to the mental faculty of *moral sense* (ch. 3). ‘Moral sense’ (or ‘conscience’) is a naturally selected *social instinct* that compels individuals to act for the general welfare of one’s community (or social group). Social instincts are impulses to form social bonds with members of one’s species, to develop feelings of sympathy for them, and to perform various services for them. Darwin (1871) emphasizes that these “feelings and services” are not extended to all species-members, but “only to those of the same association” (Darwin, 1871, p. 72).⁶ In contemporary terms (Sober & Wilson, 1998, ch. 1), moral sense impels individuals to perform *biologically altruistic behaviors* that benefit the biological fitness of a recipient at a potential cost to the donor’s fitness (e.g., donating money to a charity). Darwin (1871) notes that community level cooperative behaviors are observed among other animals, and he contends that: “[A]ny animal . . . endowed with social instincts, would inevitably acquire a moral sense or conscience, as soon as its intellectual powers had become as developed . . . as in man” (p. 72).

Darwin presents natural selection (viz., group selection) as the main mechanism responsible for the presence and increasing prevalence of moral sense in humans.⁷ Whereas natural selection at the level of individuals may favor selfish behavior over altruistic behavior, natural selection *at the level of groups* favors communities composed of cooperative individuals:

When two tribes of primeval man . . . came into competition, if the one tribe included . . . a greater number of courageous, sympathetic, and faithful members, who were always

⁶ Darwin (1871) suggests that social instincts represent an *extension* of more basic instincts (e.g., sympathy, love, affection) that humans have towards their own offspring and family members (p. 80).

⁷ Since the modern synthesis of Darwin’s theory of natural selection and Mendel’s theory of genetics, group selection has been controversial (see Lloyd, 2020). Whereas some (e.g., William, 1966; Dawkins, 1989) argue that natural selection only operates at the level of genes, others (e.g., Sober & Wilson, 1998) argue that natural selection operates on multiple levels, including groups.

ready to warn each other of danger, to aid and defend each other, this tribe would without doubt succeed best and conquer the other. (Darwin, 1871, p. 162)

Group selection also explains the improvement of moral standards and increased prevalence of moral sense among humans:

[A]n advancement in the standard of morality and an increase in the number of well-endowed men will certainly give an immense advantage to one tribe over another. . . . [A] tribe including many members who, from possessing in a high degree the spirit of patriotism, fidelity, obedience, courage, and sympathy, were always ready to give aid to each other and sacrifice themselves for the common good, would be victorious over most other tribes; and this would be natural selection (Darwin, 1871, p. 166).

While Darwin (1871) acknowledges that progress is not an invariable rule of natural selection, he adopts the ‘cheerful view’ that progress has been more general than retrogression in the intellectual and moral traits of humans (ch. 5).

In addition to explaining the evolution of moral instincts in humans, Darwin articulates a normative theory, wherein the *improvement* of human morality consists of extending our social instincts to members outside our putative communities:

As man advances . . . and small tribes are united into larger communities, the simplest reason would tell each individual that he *ought to extend his social instincts and sympathies to all members of the same nation . . . This point being once reached, there is only an artificial barrier to prevent his sympathies extending to the men of all nations and races.* If . . . such men are separated from him by great differences in appearance or habits, experience unfortunately shews us how long it is before we look at them as our fellow-creatures. *Sympathy beyond the confines of man, that is humanity to the lower animals, seems to be one of the latest moral acquisitions . . . This virtue, one of our noblest with which man is endowed, seems to arise incidentally from our sympathies becoming more tender and more widely diffused, until they are extended to all sentient beings.* (Darwin, 1871, pp. 100-101, emphasis added)

While natural selection has endowed humans with a natural disposition to act for the good of their community, expanding one’s social community beyond artificial in-group/ out-group distinctions is a sign of moral progress. Darwin’s stance implies that humans’ natural tribalist instincts *impede* moral progress by limiting who is regarded as a member of one’s community and by justifying prejudicial behavior (e.g., racism, nationalism, slavery, colonialism) towards out-group members. Humans reach a higher stage of morality when they extend their social instincts and sympathies more broadly to include all humans and even members of other species. Darwin’s view implies that the highest moral behaviors will be altruistic behaviors that *conflict with one’s biological interest*, e.g., adopting a nonbiologically related child, nursing an injured animal. Darwin’s contention that the extension of our sympathies to other species is one of the “noblest” human moral virtues is represented in contemporary ethical arguments for the humane treatment of animals (Singer, 1975; Rachels, 1990).

Some philosophers draw skeptical ethical and meta-ethical conclusions based on the evolution of morality. Ruse (1986) argues that the evolution of moral instincts undermines moral realism (i.e., the view that there are objective mind-independent moral truths) and supports ethical skepticism. He endorses the neo-Darwinian, sociobiological view defended by E. O. Wilson (1975, 1978) that moral sentiments that predispose humans towards cooperation are explained by *kin selection* (i.e., helping behavior towards biological relatives) and *reciprocal altruism* (i.e., helping behavior towards non-relatives premised on the expectation of

reciprocation). While evolutionary theory can *explain* the presence of moral sentiments, Ruse contends that evolutionary facts cannot *justify* objective moral values. On this issue, Ruse assumes that attempts to infer moral truths from natural facts would: (1) violate Hume's injunction that 'ought'-statements cannot be inferred from 'is'-statements, and (2) commit the naturalistic fallacy of inferring something is 'good' because of natural facts about that thing (Moore, 1903). For Ruse, the fact that moral sentiments are adaptive (i.e., evolutionarily stable strategies that optimize biological fitness) implies nothing about how humans *should* act. Moreover, Ruse (1986) argues that in order for moral sentiments to *effectively motivate action*, humans have been *programmed to believe* that their moral sentiments track objective moral truths (cf. Churchland, 2019). This stance cuts against moral realism in two ways. If cooperative and prosocial tendencies in humans are adaptations that optimize inclusive fitness, then apparently 'moral' behavior can be explained as (genetically) self-interested behavior. Moreover, if humans have evolved to believe their moral sentiments correspond to objective moral truths, then moral judgments that find robust consensus across cultures (e.g., do not murder) are not evidence of the truth of these moral statements, but "a collective illusion foisted upon us by our genes" (Ruse, 1986, p. 253). Similarly, Joyce (2006) and Street (2006) argue that if our moral dispositions evolved because they contributed to biological fitness, we should have low confidence that our moral judgments (e.g., do not exploit, do not steal) correspond to *mind independent moral truths* presupposed by moral realists (cf. Mackie, 1977).⁸

In contrast to ethical skeptics, other philosophers argue that objective moral values can be derived from (and justified by) evolutionary facts. Richards (1986) argues that acting for community welfare is the highest moral good and this moral imperative is justified by evolutionary facts. He assumes, contra-Ruse, that actions are *genuinely altruistic* (as opposed to actions that can be reduced to genetic selfishness). Richards' argument is formulated from the perspective of practical reason, and it only applies to creatures who *act intentionally* (i.e., act consciously based on explicit reasons). If we assume that evolution constituted humans to act for the community good, then rational humans will recognize this goal and conclude that they *ought to* act in ways to effectively achieve it. The 'ought' in this formulation refers to an 'expectation' (or 'prediction') analogous to the 'oughts' found in scientific contexts (e.g., 'if lightning strikes, then it ought to thunder') or practical contexts (e.g., 'if one is a good student, then one ought to study for their classes'). What makes this a 'moral ought' for Richards is the assumption that evolution provides the basis for human morality: If (as assumed by Ruse and others) evolution furnishes the natural basis for the *moral part of human nature* (viz., the tendency to act altruistically), then normally-constituted rational humans ought to act altruistically. Collier and Stingl (1993) argue that this sense of 'moral ought' (i.e., moral expectation) fails to establish the *moral obligation* sought in deontological ethics (p. 51). Richards (1986) argues that his account implies moral obligation because humans *cannot avoid* the structured context of evolution, which imposes a *practical necessity* to follow moral norms (cf. Gewirth, 1978). Accordingly, the moral imperative that one ought to act altruistically is derived from and justified by evolutionary facts, and Richards argues that this shows that G. E. Moore's naturalistic fallacy is not a fallacy (cf. Campbell, 1993).⁹

⁸ For a more comprehensive and critical discussion of debunking arguments, see Machery and Mallon (2010), Ruse and Richards (2017), and FitzPatrick (2020).

⁹ Curiously, evolutionary ethicists have largely ignored Darwin's own normative suggestion that the highest moral behaviors *conflict* with individuals' biological interest insofar as they involve extending altruistic impulses beyond our putative communities to all sentient life (for an exception, see Wilson, 2010). It is worth noting that Richards'

Implicit Bias Research and Moral Responsibility

Since the 2000s, interest in moral psychology by naturalistic and empirical philosophers has grown rapidly (e.g., see Doris et al., 2010; Doris et al., 2020). Moral psychology is an interdisciplinary field that: (1) investigates the psychological processes involved in moral judgment and behavior, and (2) examines the implications of such findings for traditional ethical and meta-ethical debates. Naturalistic philosophers adopting empirical (and experimental) methodological approaches have provided influential contributions to this second task. For example, based on research from social psychology (e.g., Milgram's obedience experiments, Zimbardo's Stanford prison experiment, the phone-booth experiment) that indicate moral behavior is significantly influenced by (sometimes minor and arbitrary) social and situational factors, some philosophers (e.g., Doris, 1998, 2002; Harman, 1999) argue that virtue ethics is misguided in both its emphasis on identifying *stable character traits* ('virtues') and its corollary assumption that character traits are what matter for ethics.

One line of research in moral psychology pursued by empirically oriented philosophers concerns the implications of research on implicit bias for philosophical accounts of moral responsibility (Holroyd et al., 2017). Research on implicit bias indicates that people unintentionally and unconsciously act on the basis of prejudice and stereotypes (Brownstein, 2019). Some philosophers have examined whether moral responsibility requires agents to be *responsible for their implicit biases* (e.g., the implicit attitude that black men are violent and aggressive) and the behaviors that result from such biases (e.g., a police officer using excessive force on a black civilian). Some philosophers argue that moral responsibility does not require agents to be responsible for their implicit biases. For example, Wolf (1987) defends an influential compatibilist account of moral responsibility (i.e., the 'sane deep-self' view), wherein individuals who were molded (i.e., determined) to endorse repugnant values (e.g., a child who was raised to embrace the sadistic and fascist values of his father) should not be held morally responsible for their actions because their deep self is 'insane' (i.e., they lack the requisite cognitive and normative capacity to perceive the world rationally). In the terms of implicit bias, Saul (2013) argues that individuals should not be held morally responsible for their implicit biases, since individuals should not be blamed for biases that they are unaware of and that were determined by forces out of their control. While Saul contends that *awareness* of one's biases is a necessary condition for moral responsibility, Sher (2009) and Holroyd (2012) argue that it is not being aware of one's implicit biases that matters for moral responsibility, but *when one should be aware of these biases*. This implies that individuals are sometimes responsible for their implicit biases. Along these lines, Washington and Kelly (2010) defend an externalist account wherein moral responsibility depends on contextual features of the social environment and individuals are morally responsible for actions stemming from implicit biases that they *should be aware of*. Consider two individuals in different historical contexts (e.g., 1950 versus 2020) who are tasked to hire an employee. Both hirers profess themselves to be egalitarians, but make

account risks committing the naturalistic fallacy precisely because he assumes moral behaviors *coincide* with biological interest.

decisions on the basis of implicit biases. While neither is aware of implicit bias research, Washington and Kelly argue that the hirer in 2020 is morally blameworthy, while the hirer in 1950 is not. In the context of 2020 (but not 1950), individuals in positions to make hiring decisions *should be aware* of implicit bias research. Hence, assessments of moral responsibility depend on what biases individuals *can be expected to be aware of* in a particular social and historical context. This captures the idea that we should not necessarily hold actors in the distant past morally responsible for actions that we now regard as blameworthy.

An empirical issue surrounding implicit bias in relation to moral responsibility concerns what *control* (or power) individuals have in changing their implicit biases (Brownstein, 2019). Incompatibilist accounts of freedom imply that agents are morally responsible for actions that are under their control, wherein ‘controlled actions’ imply that *one could have acted otherwise*. A prima facie reason for thinking that individuals are not morally responsible for actions resulting from implicit attitudes is that *individuals cannot control these biases* insofar as they are caused by factors outside their control and not rationally revisable. If implicit attitudes are uncontrollable insofar as they are ubiquitous, subconscious, and automatic (in contrast to the controllability and reasons-responsiveness of explicit reasoning) and implicit attitudes cause action, then individuals should not be held morally responsible for actions caused by implicit bias. Holroyd and Kelly (2016) argue that the automaticity of implicit biases are distinguishable from their controllability, and they contend that individuals can have some (‘ecological’) control over changing their implicit biases (cf. Suhler & Churchland, 2009). In this regard, Holroyd (2012) distinguishes between cognitive states and behaviors that individuals *directly control* (i.e., direct voluntary control) and *indirectly control* (i.e., control that is fostered over a period of practice and training). Like other activities which we indirectly control (e.g., playing the piano well), Holroyd argues that individuals (aware of their biases) can indirectly control (and be held morally responsible for) their implicit attitudes. Buckwalter (2019) argues that scientific evidence (e.g., psychological research aimed at changing implicit biases) supports the conclusion that implicit biases are flexible and changeable (i.e., controllable) among individuals, which suggests that individuals should be held morally responsible for their implicit biases and behaviors influenced by such biases.

Issues for Empirical Approaches to Philosophy

The survey of empirical approaches to philosophy in this chapter highlights some salient features of such methodological approaches (e.g., the prioritization of empirical evidence over commonsense, the formulation of philosophical theories based on scientific theories, criticism of traditional philosophical theories on empirical grounds); however, it is difficult to draw any systematic generalizations. By way of conclusion, I discuss three related issues facing empirically informed approaches to philosophy: (1) how empirical scientific findings can inform normative philosophical concepts, (2) whether there is a legitimate role for intuitions or commonsense in philosophical inquiry, and (3) what is distinctively ‘philosophical’ in such approaches to philosophy.

A central issue regarding empirical approaches to philosophy is the problem of how descriptive empirical theories can inform normative philosophical concepts (e.g., ‘justification,’ ‘moral goodness’). In defending a thoroughly empirical theory of knowledge, Kornblith (2002) argues that identifying the (descriptive) reliable psychological processes involved in belief formation is relevant for normative (prescriptive) epistemological issues insofar as those processes reveal features of *good* (i.e., reliable) psychological functioning, which have evolved because they helped humans and other animals survive. Similarly, Richards argues that identifying the (descriptive) evolutionary functions of moral instincts (i.e., promoting the survival of human communities) is relevant to normative (prescriptive) issues concerning how humans ought to act. These naturalistic analyses risk committing some version of the naturalistic fallacy by assuming that normative philosophical values can be *directly inferred* from facts about how humans have evolved. Conversely, purely empirical naturalists, such as Quine and Kornblith, argue that shifting focus to (descriptive) natural processes—away from idealized and abstract concepts—provides a more realistic and promising approach for articulating normative philosophical concepts. Other empirical analyses (e.g., empirical approaches to moral psychology) adopt the more deflationary strategy of utilizing empirical scientific results to criticize and revise traditional normative philosophical theories and concepts (e.g., virtue ethics, moral responsibility).

Another methodological issue facing empirical approaches is whether there is any a legitimate role for *a priori* methods (e.g., intuitions or commonsense) in philosophical inquiry, which concerns *the extent* that philosophical methodology should be empirical. Naturalistic philosophers committed to *N2* prioritize scientific and empirical methods over *a priori* methods; however, few accept the extreme position (defended by Quine, Kornblith, and the Churchlands) that genuine philosophical problems should be reconceived as *purely empirical problems* addressable by scientific methods. Recent attention has focused on the viability of conceptual analysis—especially its reliance on intuitions—as a philosophical method (Pust, 2017; Margolis & Laurence, 2021). Much of this discussion has been advanced by findings in experimental philosophy (Knobe & Nichols, 2017) that indicate intuitions about philosophical thought experiments are an unreliable source of information (e.g., intuitions show significant cross-cultural variation, professional philosophers’ intuitions are subject to cognitive biases). Hence, intuitions should not be regarded as a robust source of *evidence* in philosophy (Weingberg, Nichols, & Stich, 2001; Alexander, Mallon, & Weinberg, 2010). This perspective vitiates philosophical methods, such as conceptual analysis (Chalmers, 1996; Jackson, 1998; cf. Bealer, 1998, Bonjour, 1998), that assume that commonsense intuitions provide *a priori* justification for philosophical conclusions.¹⁰ These *a priori* methodologies violate *N1* (by being constrained by commonsense rather than science), *N2* (by assuming philosophy is an autonomous *a priori* discipline), and *N3* (by privileging *a priori* methods over *a posteriori* methods). They are paradigms of the armchair philosophy dismissed by Quine. Other naturalists grant a limited, but

¹⁰ Conceptual analysis has been criticized for its *a priori* nature, reliance on folk intuitions, and presupposition of a misleading theory of meaning (e.g., see Stich, 1992; Hardcastle, 1996; DePaul & Ramsey, 1998; Laurence & Margolis, 2003; Nimtz, 2004; Papineau, 2021). Machery (2017) defends a naturalistic version of conceptual analysis that demands that concepts are empirically validated.

legitimate, evidentiary role for intuitions in philosophy. In contrast to Kornblith's purely empiricist approach to reliabilism, Goldman (1999, 2007) defends a moderate naturalism, wherein intuitions play an indispensable role for clarifying *the (normative) standards of reliability* that belief-forming psychological processes should satisfy. While Goldman maintains that intuitions are fallible and revisable based on empirical considerations, he argues that they provide a defeasible form of philosophical evidence that includes *a priori* and *a posteriori* components (cf. Williamson, 2007). By contrast, Papineau (2013) argues that intuitions are *empirical* in nature and intuitions about thought experiments play a crucial role in *clarifying implicit assumptions* that require critical examination. Moreover, Papineau suggests that *empirically supported* intuitions can provide evidence for philosophical conclusions. In this connection, some experimental philosophers have conducted empirical studies aimed at identifying widely held folk intuitions, e.g., whether folk intuitions support compatibilist or incompatibilist theories of moral responsibility (see Doris et al., 2020). Some research suggests that intuitions about Gettier cases are robust across cultures (Machery et al., 2015), which contradicts earlier findings in experimental philosophy.

Finally, empirical approaches to philosophy face the issue of what is distinctively 'philosophical' about their methods. This objection has been raised against the naturalistic approaches to epistemology defended by Kornblith (Bonjour, 2006) as well as experimental philosophy (Sorrell, 2018): while these empirical projects are interesting, they should not be regarded as philosophy. This type of objection assumes a narrow and monistic vision of philosophy and ignores the wide diversity of methods—including a range of *a priori* and *a posteriori* methods—that have been adopted in contemporary analytic philosophy. A more constructive view of philosophy, which can accommodate the actual historical and contemporary practices of philosophers, tolerates and embraces a plurality of methods as providing defeasible evidence for or against philosophical views. In this metaphilosophical ideal, a plurality of philosophical methods is valuable because it encourages criticism of competing views (including alternative philosophical methods) that would be unavailable if all philosophers adopted the same methods (cf. Feyerabend, 1975). This argument for methodological pluralism need not assume that there will be 'progress' in philosophical methods towards improved methods (cf. Wilson, 2014). Rather, it encourages a plurality of philosophical methods as a means to avoid methodological dogmatism and stagnation. On the other hand, there is a robust tradition of analytic philosophy—following the historical trajectory of logical empiricism and Quine—that has strong methodological affinities towards science and empiricism. These naturalistic and empirical approaches represent an important chapter in the history of scientific philosophy and the broader history of debates between rationalists and empiricists.

Acknowledgements: I am grateful to Stephen Biggs, David Alexander, Trevor Pearce, and Marcus Rossberg for helpful comments and suggestions on this chapter.

References

- Alexander, Joshua, Mallon, Ronald, & Weinberg, Jonathan M. (2010). Accentuate the negative. *Review of Philosophy and Psychology*, 1(2), 297–314.
- Bealer, George (1998). Intuition and the autonomy of philosophy. In DePaul & Ramsey (1998), pp. 201–239.
- Bickle, John (2003). *Philosophy and Neuroscience: A Ruthlessly Reductive Approach*. Dordrecht: Kluwer.
- Bird, Alexander (2002). Kuhn’s wrong turning. *Studies in History and Philosophy of Science*, 33(3), 443–463.
- Boghossian, Paul Artin (1996). Analyticity reconsidered. *Noûs*, 30(3), 360–391.
- Bonjour, Laurence (1998). *In Defense of Pure Reason: A Rationalist Account of A Priori Justification*. Cambridge: Cambridge University Press.
- Bonjour, Laurence (2006). Kornblith on knowledge and epistemology. *Philosophical Studies*, 127(2), 317–335.
- Boyd, Richard (1980). Scientific realism and naturalistic epistemology. In Peter D. Asquith & Ronald N. Giere (eds.), *PSA 1980*, vol. 2, pp. 613–662. East Lansing, MI: Philosophy of Science Association.
- Boyd, Richard, Gasper, Philip, & Trout, J. D. (eds.) (1999). *The Philosophy of Science*. Cambridge, MA: MIT Press.
- Brown, Matthew J. (2012). John Dewey’s logic of science. *HOPOS: The Journal of the International Society for the History of Philosophy of Science*, 2(2), 258–306.
- Brownstein, Michael (2019). Implicit bias. In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2019 edn.).
<https://plato.stanford.edu/archives/fall2019/entries/implicit-bias/>
- Callebaut, Werner (ed.) (1993). *Taking the Naturalistic Turn, or How Real Philosophy of Science is Done*. Chicago: University of Chicago Press.
- Campbell, Richmond (1996). Can biology make ethics objective? *Biology & Philosophy*, 11(1), 21–31.
- Cappelen, Herman (2012). *Philosophy Without Intuitions*. Oxford: Oxford University Press.

- Carnap, Rudolf ([1928] 1967). *The Logical Structure of the World*. Trans. Rolf A. George. Berkeley, CA: University of California Press.
- Carnap, Rudolf (1936). Testability and meaning. *Philosophy of Science*, 3(4), 419–471.
- Carnap, Rudolf (1937). Testability and meaning—continued. *Philosophy of Science*, 4(1), 1–40.
- Cartwright, Nancy, Cat, Jordi, Fleck, Lola, & Uebel, Thomas E. (1996). *Otto Neurath: Between Science and Politics*. Cambridge: Cambridge University Press.
- Chalmers, David J. (1995). Facing up to the problem of consciousness. *Journal of Consciousness Studies*, 2(3), 200–219.
- Chalmers, David J. (1996). *The Conscious Mind: In Search of a Fundamental Theory*. Oxford: Oxford University Press.
- Chalmers, David (2014). Intuitions in philosophy: A minimal defense. *Philosophical Studies*, 171(3), 535–544.
- Chemero, Anthony (2009). *Radical Embodied Cognitive Science*. Cambridge, MA: MIT Press.
- Churchland, Patricia S. (1986). *Neurophilosophy*. Cambridge, MA: MIT Press.
- Churchland, Patricia S. (1996). The Hornswoggle problem. *Journal of Consciousness Studies*, 2(5–6), 402–408.
- Churchland, Patricia S. (2019). *Conscience: The Origins of Moral Intuition*. New York: W. W. Norton.
- Churchland, Paul M. (1981). Eliminative materialism and the propositional attitudes. *Journal of Philosophy*, 78(2), 67–90.
- Churchland, Paul M. (2013). *Matter and Consciousness*, 3rd edn. Cambridge, MA: MIT press.
- Collier, John, & Stingl, Michael (1993). Evolutionary naturalism and the objectivity of morality. *Biology & Philosophy*, 8(1), 47–60.
- Creath, Richard (2021). Logical Empiricism. In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2021 Edition).
<https://plato.stanford.edu/archives/win2021/entries/logical-empiricism/>
- Darwin, Charles (1871). *The Descent of Man, and Selection in Relation to Sex*. London: John Murray.
- Dawkins, Richard (1989). *The Selfish Gene*, 2nd edn. Oxford: Oxford University Press.
- Dennett, Daniel C. (1987). *The Intentional Stance*. Cambridge, MA: MIT Press.
- Dennett, Daniel C. (1988). Quining qualia. In A. J. Marcel & E. Bisiach (eds.), *Consciousness in Contemporary Science* (pp. 42–77). New York: Oxford University Press.

- DePaul, Michael R., & Ramsey, William (eds.) (1998). *Rethinking Intuition: The Psychology of Intuition and Its Role in Philosophical Inquiry*. Lanham, MD: Rowman & Littlefield.
- Devitt, Michael (1996). *Coming to Our Senses: A Naturalistic Program for Semantic Localism*. Cambridge: Cambridge University Press.
- Devitt, Michael (2015). Relying on intuitions: Where Cappelen and Deutsch go wrong. *Inquiry*, 58(7–8), 669–699.
- Dewey, John (1929). *The Quest for Certainty: A Study of the Relation between Knowledge and Action*. New York: Minton, Balch, and Co.
- Dewey, John (1944). By Nature and by Art. *Journal of Philosophy*, 41(11), 281–292.
- Dowe, Phil (2000). *Physical Causation*. Cambridge: Cambridge University Press.
- Doris, John M. (1998). Persons, situations, and virtue ethics. *Noûs*, 32(4), 504–530.
- Doris, John M. (2002). *Lack of Character: Personality and Moral Behavior*. New York: Cambridge University Press.
- Doris, John, Stich, Stephen, Phillips, Jonathan, & Walmsley, Lachlan (2020). Moral psychology: Empirical approaches. In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2020 edn).
<https://plato.stanford.edu/archives/spr2020/entries/moral-psych-emp/>
- Doris, John M., & the Moral Psychology Research Group (eds.) (2010). *The Moral Psychology Handbook*. Oxford: Oxford University Press.
- Feyerabend, Paul K. (1975). *Against Method: Outline of an Anarchistic Theory of Knowledge*. London: New-Left Books.
- FitzPatrick, William (2021). Morality and evolutionary biology. In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2021 edn).
<https://plato.stanford.edu/archives/spr2021/entries/morality-biology/>
- Fodor, Jerry A. (1975). *The Language of Thought*. Cambridge, MA: Harvard University Press.
- Friedman, Michael (1997). Philosophical naturalism. *Proceedings and Addresses of the American Philosophical Association* 71(2), 7–21.
- Friedman, Michael (1999). *Reconsidering Logical Positivism*. Cambridge: Cambridge University Press.
- Friedman, Michael (2001). *Dynamics of Reason*. Stanford, CA: CSLI Publications.
- Friedman, Michael, & Creath, Richard (eds.) (2007). *The Cambridge Companion to Carnap*. Cambridge: Cambridge University Press.
- Gauthier, David (1986). *Morals by Agreement*. Oxford: Oxford University Press.

- Gewirth, Alan (1978). *Reason and Morality*. Chicago: University of Chicago Press.
- Gibson, Roger (1995). Quine on the naturalizing of epistemology. In Leonardi & Santambrogio (1995), pp. 89–103.
- Gibson, Roger F. (2004a). Quine’s behaviorism cum empiricism. In Gibson (2004b), pp. 181–199.
- Gibson, Roger F. (ed.) (2004b). *The Cambridge Companion to Quine*. Cambridge: Cambridge University Press.
- Gibson, Roger F. (2004c). Willard Van Orman Quine. In Gibson (2004b), pp. 1–18.
- Giere, Ronald N. (1985). Philosophy of science naturalized. *Philosophy of Science*, 52(3), 331–356.
- Giere, Ronald N., & Richardson, Alan W. (eds.) (1996). *Origins of Logical Empiricism*. Minnesota Studies in the Philosophy of Science, vol. 16. Minneapolis, MN: University of Minnesota Press.
- Glock, Hans-Johann (2003). *Quine and Davidson on Language, Thought and Reality*. Cambridge: Cambridge University Press.
- Godfrey-Smith, Peter (2002). Dewey on naturalism, realism and science. *Philosophy of Science* 69(S3), S25–S35.
- Godfrey-Smith, Peter (2014). Quine and pragmatism. In Gilbert Harman & Ernie Lepore (eds.), *A Companion to W. V. O. Quine* (pp. 54–68). Malden, MA: Wiley Blackwell.
- Goldman, Alvin I. (1986). *Epistemology and Cognition*. Cambridge, MA: Harvard University Press.
- Goldman, Alvin I. (1999). *A priori* warrant and naturalistic epistemology. *Philosophical Perspectives*, 13, 1–28.
- Goldman, Alvin I. (2007). Philosophical intuitions: Their target, source, and their epistemic status. *Grazer Philosophische Studien*, 74(1), 1–26.
- Griffiths, Paul E. (1997). *What Emotions Really Are*. Chicago: University of Chicago Press.
- Haack, Susan (1990). Rebuilding the ship while sailing on the water. In Roger Barrett & Roger Gibson (eds.), *Perspectives on Quine* (pp. 111–127). Oxford: Blackwell.
- Haack, Susan (1993). The two faces of Quine’s naturalism. *Synthese*, 94(3), 335–356.
- Hardcastle, Gary L., & Richardson, Alan W. (2003). *Logical Empiricism in North America*. Minnesota Studies in the Philosophy of Science, vol. 18. Minneapolis, MN: University of Minnesota Press.
- Hardcastle, Valerie Gray (1996). Review of *The Conscious Mind* by David Chalmers. *Journal of Mind and Behavior*, 17(4), 391–398.

- Hardcastle, Valerie Gray (1999). *The Myth of Pain*, Cambridge, MA: MIT Press.
- Harman, Gilbert (1999). *Reason, Meaning, and Mind*. Oxford: Oxford University Press.
- Harman, Gilbert (2000). The nonexistence of character traits. *Proceedings of the Aristotelian Society*, 100(2000), 223–226.
- Haug, Matthew C. (ed.) (2014). *Philosophical Methodology: The Armchair or the Laboratory?* London: Routledge.
- Holroyd, Jules (2012). Responsibility for implicit bias. *Journal of Social Philosophy*, 43(3), 274–306.
- Holroyd, Jules, Scaife, Robin, & Stafford, Tom (2017). Responsibility for implicit bias. *Philosophy Compass*, 12(3), e12410.
<https://doi.org/10.1111/phc3.12410>
- Holroyd, Jules, & Kelly, Daniel (2016). Implicit bias, character, and control. In Alberto Masala & Jonathan Webber (eds.), *From Personality to Virtue: Essays on the Philosophy of Character* (pp. 106–133). Oxford: Oxford University Press.
- Howard, Don (2003). Two left turns make a right: On the curious political career of North American philosophy of science midcentury. In Hardcastle & Richardson (2003), pp. 25–93.
- Ismael, Jenann (2013). Causation, free will, and naturalism. In Don Ross, Harold Kincaid, & James Ladyman (eds.), *Scientific Metaphysics* (pp. 208–236). Oxford: Oxford University Press.
- Ismael, J. T. (2016). *How Physics Makes Us Free*. Oxford: Oxford University Press.
- Jackson, Frank (1998). *From Metaphysics to Ethics: A Defense of Conceptual Analysis*. Oxford: Oxford University Press.
- Joyce, Richard (2006). *The Evolution of Morality*. Cambridge, MA: MIT Press.
- Kim, Jaegwon (1988). What is ‘naturalized epistemology?’ *Philosophical Perspectives*, 2, 381–405.
- Kim, Jaegwon (2003). The American origins of philosophical naturalism. *Journal of Philosophical Research*, 28(Suppl.), 83–98.
- Kitcher, Philip (1992). The naturalists return. *Philosophical Review*, 101(1), 52–114.
- Knobe, Joshua, & Nichols, Shaun (2017). Experimental philosophy. In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2017 edn.).
<https://plato.stanford.edu/archives/win2017/entries/experimental-philosophy/>
- Kornblith, Hilary (1994a). Introduction: What is naturalistic epistemology? In Kornblith (1994b), pp. 1–14.
- Kornblith, Hilary (1994b). *Naturalizing Epistemology*, 2nd edn. Cambridge, MA: MIT Press.

- Kornblith, Hilary (2002). *Knowledge and Its Place in Nature*. Oxford: Oxford University Press.
- Kuhn, Thomas S. (1962). *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.
- Ladyman, James, Ross, Don, Spurrett, David, & Collier, John (2007). *Every Thing Must Go: Metaphysics Naturalized*. Oxford: Oxford University Press.
- Laurence, Stephen, & Margolis, Eric (2003). Concepts and conceptual analysis. *Philosophy and Phenomenological Research*, 67(2), 253–282.
- Leonardi, Paolo, & Santambrogio, Marco (eds.) (1995). *On Quine: New Essays*. Cambridge: Cambridge University Press.
- Lloyd, Elisabeth (2020). Units and levels of selection. In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2020 edn.).
<https://plato.stanford.edu/archives/spr2020/entries/selection-units/>
- Machery, Edouard (2017). *Philosophy Within Its Proper Bounds*. Oxford: Oxford University Press.
- Machery, Edouard, & Mallon, Ron (2010). Evolution of morality. In Doris et al. (2010), pp. 3–46.
- Machery, Edouard, Stich, Stephen, Rose, David, et al. (2015). Gettier across cultures. *Noûs*, 51(3), 645–664.
- Mackie, J. L. (1977). *Ethics: Inventing Right and Wrong*, Harmondsworth: Penguin Books.
- Maddie, Penelope (2007). *Second Philosophy: A Naturalistic Method*. Oxford: Oxford University Press.
- Margolis, Eric, & Laurence, Stephen (2021). Concepts. In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2021 edn.).
<https://plato.stanford.edu/archives/spr2021/entries/concepts/>
- Millikan, Ruth G. (1984). *Language, Thought, and Other Biological Categories: New Foundations for Realism*. Cambridge, MA: MIT Press.
- Moore, G. E. (1903). *Principia Ethica*, Cambridge: Cambridge University Press.
- Nagel, Thomas (1974). What is it like to be a bat? *Philosophical Review*, 83(4), 435–450.
- Neander, Karen (2017). *A Mark of the Mental: A Defense of Informational Teleosemantics*. Cambridge, MA: MIT Press.
- Nimtz, Christian (2004). Two-dimensional and natural kind terms. *Synthese*, 138(1), 125–148.
- Papineau, David (2014). The poverty of conceptual analysis, in Haug (2014), pp. 166–194.

- Papineau, David (2021). Naturalism. In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2021 edn.).
<https://plato.stanford.edu/archives/sum2021/entries/naturalism/>
- Pearce, Trevor (2020). *Pragmatism's Evolution: Organism and Environment in American Philosophy*. Chicago: University of Chicago Press.
- Pietroski, Paul M. (2018). *Conjoining Meanings: Semantics Without Truth Values*. Oxford: Oxford University Press.
- Place, U. T. (1956). Is consciousness a brain process? *British Journal of Psychology*, 47(1), 44–50.
- Prinz, Jesse J. (2007). *The Emotional Construction of Morals*. Oxford: Oxford University Press.
- Pust, Joel (2019). Intuition. In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2019 edn.).
<https://plato.stanford.edu/archives/sum2019/entries/intuition/>
- Putnam, Hilary (1962a) It ain't necessarily so. *Journal of Philosophy*, 59(22), 658–671.
- Putnam, Hilary (1962b). The analytic and synthetic. In Herbert Feigl & Grover Maxwell (eds.), *Scientific Explanation, Space, and Time*. Minnesota Studies in the Philosophy of Science, vol. 3 (pp. 358–397). Minneapolis: Minnesota University Press.
- Putnam, Hilary (1982). Why reason can't be naturalized. *Synthese*, 52(1), 3–23.
- Pyle, Andrew (ed.) (1999). *Key Philosophers in Conversation: The Cogito Interviews*. London: Routledge.
- Quine, W.V. (1948). On what there is. Reprinted in Quine (1980), pp. 1–19.
- Quine, W. V. (1951). Two dogmas of empiricism. Reprinted (with revisions) in Quine (1980), pp. 20–46.
- Quine, W. V. (1960). Posits and Reality. Reprinted in Quine (1976), pp. 246–254.
- Quine, W. V. (1969a). Epistemology naturalized. In Quine (1969b), pp. 69–90.
- Quine, W.V. (1969b). *Ontological Relativity and Other Essays*. New York: Columbia University Press.
- Quine, W. V. (1960). *The Roots of Reference*. LaSalle, IL: Open Court.
- Quine, W. V. (1976). *The Ways of Paradox and Other Essays*, 2nd edn. Cambridge, MA: Harvard University Press.
- Quine, Willard Van Orman (1980). *From a Logical Point of View: Nine Logico-Philosophical Essays*, 2nd edn. Cambridge, MA: Harvard University Press.

- Rachels, James (1990). *Created from Animals: The Moral Implications of Darwinism*. Oxford: Oxford University Press.
- Ramsey, William (2020). Eliminative materialism. In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2020 edn.).
<https://plato.stanford.edu/archives/sum2020/entries/materialism-eliminative/>
- Ramsey, William, Stich, Stephen, & Garon, Joseph (1990). Connectionism, eliminativism and the future of folk psychology. *Philosophical Perspectives*, 4, 499–533.
- Reisch, George A. (2005). *How the Cold War Transformed Philosophy of Science: To the Icy Slopes of Logic*. Cambridge: Cambridge University Press.
- Rescorla, Michael (2020). The computational theory of mind. In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2020 edn.).
<https://plato.stanford.edu/archives/fall2020/entries/computational-mind/>
- Rey, Georges (1998). A naturalistic a priori. *Philosophical Studies*, 92(1–2), 25–43.
- Richards, Robert J. (1986). A defense of evolutionary ethics. *Biology & Philosophy*, 1(3), 265–293.
- Richardson, Alan (1997). Toward a history of scientific philosophy. *Perspectives on Science*, 5(3), 418–451.
- Richardson, Alan W. (1998). *Carnap's Construction of the World: The Aufbau and the Emergence of Logical Empiricism*. Cambridge: Cambridge University Press.
- Richardson, Alan (2008). Scientific philosophy as a topic for history of science. *Isis*, 99(1), 88–96.
- Richardson, Alan, & Uebel, Thomas (eds.) (2007). *The Cambridge Companion to Logical Empiricism*. Cambridge: Cambridge University Press.
- Rorty, Richard (1979). *Philosophy and the Mirror of Nature*. Princeton: Princeton University Press.
- Roth, Paul A. (2007). Naturalism without the fears. In Stephen P. Turner & Mark W. Risjord (eds.), *Philosophy of Anthropology and Sociology* (pp. 683–708). Amsterdam: Elsevier.
- Roth, Paul A. (2008). The epistemology of science after Quine. In Martin Curd & Stathis Psillos (eds.), *The Routledge Companion to Philosophy of Science* (pp. 3–14). London: Routledge.
- Rouse, Joseph (2002). *How Scientific Practices Matter*. Chicago: University of Chicago Press.
- Ruse, Michael (1986). *Taking Darwin Seriously: A Naturalistic Approach to Philosophy*. Oxford: Blackwell.

- Ruse, Michael, & Robert J. Richards (eds.) (2017). *The Cambridge Handbook of Evolutionary Ethics*. Cambridge: Cambridge University Press.
- Rysiew, Patrick (2020). Naturalism in epistemology. In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2020 edn.).
<https://plato.stanford.edu/archives/fall2020/entries/epistemology-naturalized/>
- Saul, Jennifer (2013). Implicit bias, stereotype threat, and women in philosophy. In Katrina Hutchinson & Fiona Jenkins (eds.), *Women in Philosophy: What Needs to Change?* (pp. 39–60). Oxford: Oxford University Press.
- Sher, George (2009). *Who Knew? Responsibility without Awareness*. New York: Oxford University Press.
- Singer, Peter (1975). *Animal Liberation: A New Ethics for Our Treatment of Animals*. New York: New York Review.
- Smart, J. J. C. (1959) Sensations and brain processes. *Philosophical Review*, 68(2), 141–156.
- Sober, Elliot, & Wilson, David Sloan (1998). *Unto Others: The Evolution and Psychology of Unselfish Behavior*. Cambridge, MA: Harvard University Press.
- Sorrell, Tom (2018). Experimental philosophy and the history of philosophy. *British Journal for the History of Philosophy*, 26(5), 829–849.
- Stadler, Friedrich (2015). *The Vienna Circle: Studies in the Origins, Development, and Influence of Logical Empiricism*, 2nd edn. Vienna Circle Institute Library, vol. 4. Cham: Springer.
- Stich, Stephen (1983). *From Folk Psychology to Cognitive Science: The Case Against Belief*. Cambridge, MA: MIT Press.
- Stich, Stephen (1992). What is a theory of mental representation? *Mind*, 101(402), 243–61.
- Street, Sharon (2006). A Darwinian dilemma for realist theories of value. *Philosophical Studies*, 127(1), 109–66.
- Stroud, Barry (1984). *The Significance of Philosophical Scepticism*. Oxford: Clarendon Press.
- Suhler, Christopher, & Patricia S. Churchland (2009). Control: conscious and otherwise. *Trends in Cognitive Sciences*, 13(8), 341–347.
- Tsou, Jonathan Y. (2003a). A role for reason in science. *Dialogue: Canadian Philosophical Review*, 42(3), 573–598.
- Tsou, Jonathan Y. (2003b). The justification of concepts in Carnap's *Aufbau*. *Philosophy of Science*, 70(4), 671–689.

Tsou, Jonathan Y. (2010). Putnam's account of apriority and scientific change: Its historical and contemporary interest. *Synthese*, 176(3), 429–445.

Tsou, Jonathan Y. (2013). Origins of the qualitative aspects of consciousness: Evolutionary answers to Chalmers' hard problem. In Liz Swan (ed.), *Origins of Mind* (pp. 259–269). Dordrecht: Springer.

Tsou, Jonathan Y. (2015). Reconsidering the Carnap-Kuhn connection. In William J. Devlin & Alisa Bokulich (eds.), *Kuhn's Structure of Scientific Revolutions – 50 Years On*. Boston Studies in the Philosophy and History of Science, vol. 311 (pp. 51–69). Cham: Springer.

Uebel, Thomas (2007). *Empiricism at the Crossroads: The Vienna Circle's Protocol-Sentence Debate Revisited*. LaSalle, IL: Open Court.

van Fraassen, Bas C. (1995). Against naturalized epistemology. In Leonardi & Santambrogio (1995), pp. 89–103.

Washington, Natalia, & Kelly, Daniel (2016). Who's responsible for this? Moral responsibility, externalism, and knowledge about implicit bias. In Michael Brownstein & Jennifer Saul (eds.), *Implicit Bias and Philosophy, Volume 2: Moral Responsibility, Structural Injustice, and Ethics* (pp. 11–36). Oxford: Oxford University Press.

Weinberg, Jonathan M., Nichols, Shaun., & Stich, Stephen (2001). Normativity and epistemic intuitions. *Philosophical Topics*, 29(1–2), 429–460.

Williams, George C. (1966). *Adaptation and Natural Selection: A Critique of Some Current Evolutionary Thought*. Princeton: Princeton University Press.

Williamson, Timothy. (2007). *The Philosophy of Philosophy*. Oxford: Blackwell.

Wilson, Catherine (2010). Darwinian morality. *Evolution: Education and Outreach*, 3(2), 275–287.

Wilson, Edward O. (1975). *Sociobiology*. Cambridge, MA: Harvard University Press.

Wilson, Edward O. (1978). *On Human Nature*. Cambridge, MA: Harvard University Press.

Wilson, Jessica (2014). Three dogmas of metaphysical methodology. In Matthew C. Haug (2014), pp. 145–165.

Wimsatt, William C. (2007). *Re-Engineering Philosophy for Limited Beings: Piecewise Approximations to Reality*. Cambridge, MA: Harvard University Press.

Wimsatt, William C. (forthcoming). Evolution and the metabolism of error: Biological practice as a foundation for a scientific metaphysics. In William C. Bausman, Janella Baxter, & Oliver M. Lean (eds.), *From Biological Practice to Scientific Metaphysics*. Minnesota Studies in the Philosophy of Science. Minneapolis: University of Minnesota Press.

Wittgenstein, Ludwig ([1921] 1922). *Tractatus Logico-Philosophicus*. Trans. C. K. Ogden. London: Routledge & Kegan Paul.

Wolf, Susan (1987). Sanity and the metaphysics of responsibility. In Ferdinand Schoeman (ed.), *Responsibility, Character, and the Emotions* (pp. 46–62). Cambridge: Cambridge University Press.