

The Pragmatist Reappraisal of Habit in Contemporary Cognitive Science, Neuroscience, and Social Theory: Introductory Essay

Italo Testa and Fausto Caruana

The aim of this book is to evaluate the contribution that the notion of habit could make to current debate at the crossroads between philosophy, cognitive sciences, neurosciences, and social theory. This topic is addressed in a broad sense, dealing with the different aspects of the pragmatic turn involved by 4E (embodied, embedded, extended, and enactive) cognitive science, and tracing back the roots of such a pragmatic turn to both classical and contemporary pragmatism. Its aim is to explore the many facets of the notion of habit and to use it as the guiding thread for the theoretical reconstruction and critical reassessment of pragmatist arguments that are of great relevance to contemporary thought. In addressing such questions, the book gathers original contributions from philosophers, cognitive scientists, neuroscientists, and social theorists, aiming to offer an interdisciplinary account of “habit,” a notion whose importance is today receiving growing attention in different fields of research but whose different theoretical and historical aspects still need to be connected systematically. Notably, the common reference to the pragmatist approach to this concept is also crucial to ensure a consistent and coherent outcome, as it links together the single chapters in which the systematic project of the book is articulated.

The notion of habit, and an understanding of experience as a process of habit formation, constitute a major aspect of classical pragmatist approaches to cognition, social action, and aesthetics (Kilpinen 2012), and have played an important role in classical social sciences (Camic 1986). Still, since the middle of the last century, intentionalist and representational models have hegemonized cognitive sciences, action theory, and social ontology. More recently, the importance of the notion of habit as a viable alternative to current paradigms in some of these fields is being rediscovered. For instance, Pollard (2006) has argued against neglecting habit in action theory, Turner (2002) has emphasized the role of habit in the foundation of social theory, whereas Barandiaran and Di Paolo (2014) have pleaded for the reappraisal of habit as

a theoretical alternative to the notion of “mental representation” adopted by computational and informational cognitive science. More generally, pragmatism nowadays increasingly plays an important role as the main alternative to classical cognitive science (Fodor 2008). The rediscovery in recent years of the contemporary relevance of classical pragmatist theories such as Dewey’s, Peirce’s, James’s, and Mead’s (Rockwell 2005; Menary 2010, 2012; Johnson 2010; Paolucci 2011; Hutto and Myin 2013; Solymosi and Shook 2013; Madzia and Jung 2016; Dreon 2019) deals with the different aspects of the pragmatist turn involved in 4E cognition.

In the book we explore how the continuity from motor routines to intelligent behavior described by pragmatist accounts of habit enables 4E cognitive science to overcome the dualism between low-level aspects of cognition, on the one hand, and creative and intelligent thinking, on the other hand, and to finally offer a unifying view of cognition. We investigate the role that pragmatist insights could play in overcoming mind/body, mind/world and perception/action dualisms, inherited from the Cartesian and Kantian tradition, which have dominated internalist research programs of the last few decades and which have started to be tackled only recently. In particular, we explore how a habit-based notion of experience inspired by pragmatism could make it possible to appreciate the continuity between sensory, motor, and social aspects of action. This could offer useful theoretical tools for studying embodied, enactive, embedded, and extended approaches to cognition that have recently reevaluated the notion of habit (see Noë 2009). What is more, it could also be helpful in order to reframe neuroscientific research in interactive social contexts, to implement experimental approaches to aesthetic perception, and to offer an alternative to socio-ontological models which are based on the internalist and representational notion of collective intentionality and neglect the role of habit formation in the constitution of social entities.

In the first section of this introductory essay we first sketch the role that the notion of habit has played in the work of pragmatist authors such as James, Peirce, and Dewey, and give an account of its ambivalent role in the development of psychology and of cognitive sciences from James’s introspectionism, through behaviorism and computationalism, up to 4E cognition and its rediscovery of a pragmatist action-oriented stance to cognition. We then investigate in the second section how the abandonment of the notion of habit in cognitive sciences in the second half of the twentieth century was paralleled by the adoption of a dualism between automatic routine and intelligent action and by an approach to cognition based on the notion of mental representation. This notion was subsequently put under pressure by the emerging paradigm of 4E cognition, whose push toward an antirepresentationalist turn is leading to a reassessment of the notion of habit. In the third section we explore how habit formation has been investigated within contemporary neuroscience in a dynamic perspective based

on the interplay between automatism and goal-oriented behavior. This section highlights the role that the pragmatist ideomotor principle plays in autoptic and pragmatic coding approaches to cortical motor systems, and how recent research on mirror neurons pragmatically links action with social cognition and cultural practices. In the fourth section we see how the adoption of the dualism between rational action and mechanical routines also influenced the development of twentieth-century sociological thought, and is nowadays being reconsidered by social theory. Finally, in the fifth section of this introductory essay we provide an overview of the book and a chapter-by-chapter summary.

Chapters have been organized around a number of precise theoretical focal points corresponding to the parts and sections of the book, with each section designed to verify the incidence of a pragmatist approach on the genesis and structure of specific aspects of the contemporary research on social cognition that have emerged in differentiated and often noncommunicating disciplines. The book is divided into three parts devoted respectively to: *The Sensorimotor Embodiment of Habits* (Part I), dealing with neuroscientific approaches to habit formation (Section 1), emotional habits (Section 2), and performative skills (Section 3); *The Enactment of Habits in Mind and World* (Part II), dealing with the deployment of habits in the background of agency (Section 4), in intentionality and language (Section 5), and in moral life (Section 6); and finally, *Socially Embedded and Culturally Extended Habits* (Part III), dealing with the sociopsychological development of habits in social practices (Section 7), their embedment in cultural and aesthetic artifacts (Section 8), and their ontological extension in social structures and political institutions (Section 9).

From Pragmatism to 4E Cognitive Science and Back: An Historical Overview

The notion of habit was a fundamental psychological concept well before psychology was established as a well-defined discipline. A systematic review of this concept is outside the scope of this introduction, but a quick overview of the history of philosophy would be sufficient to show to what extent, from Aristotle onwards, authors such as Thomas Aquinas, Montaigne, Pascal, Hume, Hegel, Ravaisson, Kierkegaard, Nietzsche, and others have conceived habits as a key concept to understanding the human mind and behavior. But a comprehensive reconstruction of the history of the notion of habit is a chapter of the history of thought that still needs to be written and only a few works have recently started to undertake this task (Sparrow and Hutchinson 2013; Carlisle 2014; Piazza 2018). This is perhaps due to the fact that the philosophical tradition has showed an ambivalent attitude toward habits, which have been seen both as an indispensable part of life that structures our experience and allows free agency, but also as an “obstacle to reflection and a threat to freedom”

(Carlisle 2014: 3). This ambivalence, into which the philosophical tradition tends to divide itself, somehow reflects the intimate dynamic structure of habit formation, and has led to the development of dualist, intellectualist accounts – to be found for instance in Descartes, Kant, and Bergson – that oppose habitual and intelligent behavior, habit automatism, and reflective rationality, which as we will see has somewhat influenced the history of psychology. Still at the end of the nineteenth century, the notion of habit also played a crucial role in the work of naturalists of the caliber of Darwin, Lamarck, and Lloyd Morgan. However, its official debut in the history of psychology must be traced back to the publication of *Principles of Psychology*, in which the American pragmatist and psychologist William James placed this notion as a cornerstone of his functionalist psychology.

James's notion of habit is indebted to two thinkers: the English physician and physiologist William Benjamin Carpenter (1813–1885), who suggested that thought and behavior are largely unconscious, and the French psychologist Léon Dumont (1837–1877), who compared human habits with the laws of inanimate nature (Blanco 2014). This last intuition, provocatively supported by the suggestion that the philosophy of habit is, in the first instance, a chapter of physics, paved the way for the contemporary perspective that “habits are due to the plasticity of materials to outward agents” and that “the brain-matter is plastic” (James 1890). The Jamesian synthesis between the perspectives of the above-mentioned scholars has contributed to making James's theory of habit an incredibly modern explanation and, to some extent, a forerunner theory of brain plasticity. At the same time, however, the associationist flavor underpinning some of James's formulations contributed to the common understanding of James's theory of habit as a mechanistic explanation, with habits reduced to blind routines, with little or no contribution from our cognitive processes (see Seger and Spiering 2011; Smith and Graybiel 2016). This reading underestimates the fact that even James, while speaking of habit as the “enormous flywheel of society, its most precious conservative agent,” was not only attributing a passive and inertial side to it. He was, at the same time, underlining the plastic and teleological structure of habits (James 1890). Indeed, James's view is that humans are “bundles of habit” not only so far as actions are concerned, but also in their moral character, which also consists of “an organized set of habits of reaction” (James 1899). More generally, James's perspective is that “all our life, so far as it has definite form, is but a mass of habits – practical, emotional, and intellectual – systematically organized for our weal or woe, and bearing us irresistibly toward our destiny, whatever the latter may be” (James 1899), hence providing evidence for a much wider concept of habit. His suggestion that “the philosophical student had to get into the *habit* of thinking *unhabitually*” (Leary 2013) is in line with this wider conception. Of note, merging this wider interpretation of habit with the idea that we are subject to the law of habit in

consequence of the fact that we have bodies (James 1899) leads to an embodied, and evolutionarily sound, perspective on human and animal cognition.

The hypothesis that James's theory of habit was limited to a mechanistic explanation is also at odds with the fact that a third, and possibly more important, influence on James's thinking is his friend and colleague Charles Sanders Peirce, who developed a strongly nonmechanistic, externalist, theory of habits (see West and Anderson 2016; Fabbrichesi 2019). In Peirce's view, "multiple reiterated behavior of the same kind, under similar combinations of percepts and fancies, produces a tendency – the habit – actually to behave in a similar way under similar circumstances in the future" (Peirce 1931–60; CP 5.487). The emphasis on the role of "similar circumstances" shows that, in Peirce's hands, habits are not bodily states but, rather, forms of ongoing interaction between living organisms and the environment, and it would be a logical category error to reduce the relational nature of habits to an individual property. In other words, habits are "schematically structured sequences of acts in the world [and] formed during actual behavior when action accommodates to the objective conditions" (Määttänen 2015: 33). Hence, they are different from automatic routines.

Another important element of Peirce's conception of habit is that habits are not only limited to motor behavior, but are vehicles of cognition, insofar as cognition is per se anticipation of action. Such a perspective leads to the assumption that what a thing means is simply which habits it involves, and places habits at the root of his pragmatic maxim: "consider what effects, that might conceivably have practical bearings, we conceive the object of our conception to have. Then, our conception of these effects is the whole of our conception of the object" (CP 5.402). Summing up, not only are habits something more than mere automatic routines: according to Peirce, habits are beliefs, meanings.

The same emphasis on the relational and interactive nature of habits is expressed by John Dewey's theory of habit formation, which appears as an integration of James's psychological approach and Peirce's relational view of habits. His best conceptualization of habits is available in *Human Nature and Conduct*, first published in 1922, where we can find a sharp criticism of the intellectualist identification of habits with dead routines, leading Dewey to distinguish between "routine habits" and "intelligent habits," both understood as stabilized patterns of behavior (Dewey 1983: 32 and 48: see on this Testa 2017a, 2017b and 2017d). Dewey's approach to habit formation as a process of stabilization of patterns of interaction provides here a unified account of the social character of both the individual and the collective level of habituation which drives an interdisciplinary program of research. Dewey's approach in *Human Nature and Conduct* is firstly meant to pave the way for the foundation of psychology as a social discipline, under the premise that "an understanding of habit and of different types of habit is the key to social psychology" (1983: 3).

Here “social psychology” is broadly understood as an investigation whose main problem is not “how either individual or collective mind forms social groups and customs, but how different customs, established interacting arrangements, form and nurture different minds” (1983: 56). In this way Dewey captures the social character of cognition and relates it to the constitution and the development of behavior in the process of habituation. Since habits are “formed in the interaction of biological aptitudes with a social environment” (1983: 3), the investigation of their social character is an interdisciplinary matter where physiology, medicine, anthropology, moral psychology, and social theory work closely together in Dewey’s account. Here James’s intuition about the nature of habits is expanded by Dewey, since plasticity in his account concerns both the fact that individuals can be molded by enculturation owing to the plasticity of the brain-matter and their natural impulses, and the fact that social customs can in their turn be changed by individual and collective action.

Within the pragmatist tradition, Dewey’s habit-based approach to functional psychology was also developed by George Herbert Mead in his theory of the act (Mead 1934), which explained cognition as a matter of the active behavior of the organism, whose “attitudes come into existence as neural pathways encoding bodily habits which are responding to certain kinds of environmental stimulation” (Madzia 2013; see also Baggio 2016). Contrary to Watsonian behaviorism, which explains away consciousness, Mead’s so-called “social behaviorism,” while denying the existence of mindedness as an independent substance, was committed to account for it in naturalistic and behavioristic terms by reconceiving mental phenomena functionally (Mead 1934: 10), in accord with the functionalist approach also deployed by Dewey in his 1896 paper on the “Reflex Arc” (Dewey 1981a). Mead’s functionalist psychology, with its idea of the emergence of mind and self out of the social process of significant communication, was also part of a broader interdisciplinary program, whereby social cognition was being investigated at the interplay between evolutionary psychology and social theory, and which led to the foundation of the symbolic interactionist school of sociology and social psychology (Blumer 1969, 2004).

During the first half of the twentieth century, the functionalist program was abandoned and replaced by the rise of Watsonian behaviorism. The ultimate aim of behaviorism was to provide a science of human and animal behavior as a third-person perspective discipline that avoided reference to mentalistic, unobservable objects. In contrast to James’s understanding of psychology, characterized by concepts such as “self-consciousness” and “introspection,” behaviorism was characterized by the idea that “introspection forms no essential part of the [psychological] methods, nor is the scientific value of its data dependent upon the readiness with which they lend themselves to interpretation in terms of consciousness” (Watson 1913). Despite this radically different approach to psychology, there were several contact points between pragmatism and behaviorism, the latter being influenced by William James in

several respects (Baum 2005), while on the other hand Mead had developed within pragmatism his peculiar form of social behaviorism. Indeed, while rejecting the introspectionist Jamesian approach, Watsonian behaviorism maintained the concept of habit as a major technical concept. A concept of habit intended as a reflex-like automatic disposition instantiated by repetition – that is, the most popular interpretation of James’s account – was of course in line with the behaviorist approach which placed the stimulus–response relationship at the core of the psychological enterprise.

In the second half of the twentieth century, the cognitive revolution introduced new concepts with a mentalistic, Cartesian flavor, arguing that complex cognitive activity could be explained only if postulating “cognitive maps,” or “mental representations,” allowing both humans and animals to navigate the space around them and to perform cognitive evaluation and intentional, goal-related, behaviors (Tolman 1948; Miller, Gallanter, and Pribram 1960). In addition, the introduction of computationalism, along with the computer metaphor and the view of cognition as information processing, led to a systematic replacement of the concept of habit with that of “mental representation,” which was originally conceived as a symbolic entity implemented in the brain in an amodal (i.e. unrelated to sensory modalities) fashion (see Barsalou et al. 2003 for criticism of this approach). Summing up, the introduction of mentalistic concepts such as “representation,” “goal” and “evaluation” was associated with a slow but relentless abandonment of the concept of habit.

The beginning of the new millennium was dominated by the advent of 4E (embodied, embedded, enactive, and extended) cognitive science, which integrated some of the tenets of computationalism with functionalist insights originally advocated by American pragmatism and European phenomenology. In particular, 4E cognitive science holds that cognitive activity cannot be reduced to disembodied information processing but, rather, that cognition is an action-oriented endeavor with high-level cognition emerging from the action domain and largely shaped by sensorimotor, emotional and cultural experience. Notably, the new emerging view that perception, cognition, and action are not temporally and functionally independent processes is reminiscent of the functionalist approach described by John Dewey in his 1896 paper “The Reflex Arc Concept in Psychology” (Dewey 1981a; see also Venturelli 2012). While all 4E approaches ascribe a crucial relevance to perception and action, the role played by perception or by action for cognition has been differently emphasized. There is no doubt that European phenomenology represents the theoretical background that boosted a strong emphasis on the “primacy of perception” for cognition. In contrast, the action-oriented account promoted by 4E cognitive science is the hallmark of the pragmatist hypothesis that “[t]he whole function of thought is to produce habits of action” (Peirce, CP 5.394) and that “the whole function of thinking is but one step in the production of habits of action” (James 1898). This pragmatist stance toward cognition stimulated

the recovery of a pragmatist reading of habit: a relational concept involving the agent's disposition to act, on the one hand, and the natural, social, and cultural environment, on the other hand.

Overcoming the Dualism between Automatic Routine and Intelligent Action

Habits are commonly considered as implicit associations between contexts and responses, developing through repeated reward learning and persisting when the reward is no longer valued (Wood 2017). Albeit this operational account of habit is largely shared among philosophers, psychologists, and neuroscientists, different scholars have emphasized two distinct aspects of this notion, resulting in two main different conceptualizations. On the one hand, some authors have highlighted the automatic nature of habits, conceiving habits as blind behavioral routines. This perspective – mediated by the Cartesian and Kantian tradition, and largely fostered by a mechanistic reading of James's theory of habit – has led to an intellectualist dichotomy between habits and intelligent behavior, the resolution of which is one of the main aims of this book. On the other hand, other authors have interpreted the notion of habit as a relational concept, which allows the overcoming of ancient dualisms at the crossroads between philosophy and psychology, including the dichotomy between routine and intelligent action, the distinction between action and goal-oriented cognitive activity, and the dualism between agents and their natural, social, and cultural environment.

The view that a pragmatist account of habits is a viable strategy to boost further theoretical advancements in philosophy, psychology, and neuroscience is the common ground of the chapters of this book. The distinction between the pragmatist notion of habit and the interpretation of habit as automatic, mechanical routine maps, at least partially, onto the distinction between the associationist and the organicist historical trends, recently highlighted by Barandiaran and Di Paolo (2014) in their genealogical survey of the notion of habit. Following these authors, the associationist tradition understands habits atomistically, as units that result from the association of ideas or between stimulus and response. This interpretation places habits within the realm of reactive sub-personal mechanisms, and opposed to rational, intentional, and personal levels of cognitive processing. The associationist understanding had its most eminent representatives in the English empiricists and the American behaviorists – who substituted the notion of habit with that of “rate of conditioned response.” The organicist tradition, in contrast, understands habits holistically as dynamic, ecological, and self-organizing teleological structures. In this tradition, habits are conceptualized as traversing a continuum from prereflexive to reflexive embodied cognitive processes, rather than in opposition to rational, volitional processes. According to Barandiaran and Di Paolo, the organicist tradition

originated with Aristotle and was developed through German idealism, French spiritualism, phenomenology, and American pragmatism, but surprisingly it is still largely neglected by contemporary cognitive science and neuroscience.

It can be argued that, even putting aside the distinction between associationist versus organicist traditions, the notion of habit per se has been absent from the agenda of cognitive scientists since the original formulation of this field of research in the second half of the twentieth century. This fact can be put down to two reasons, both mentioned above: first, the growing interest in the notion of mental representation and, second, a greater interest in the study of goal-related, intelligent, behavior. Not surprisingly, the corpus of text in Google Books reveals that the frequency of use of the words “habit,” “representation,” and “goal” shows an interesting, opposite, trend (see Figure I.1). Interestingly, the fact that at the beginning of the twenty-first century, the recovery of the notion of habit is accompanied by a decrease in the use of the concepts of “representation” and “goal” indicates that these concepts are still conceived of as being in competition with one another.

The use of amodal, symbolic, mental representations as a core concept of cognitive science was originally justified by methodological reasons, mainly owing to the desire to overcome the pitfalls of behaviorism, but it has unbalanced the study of the mind in favor of an internalist framework. This characterization was subsequently criticized by 4E cognitive science, which argued that classic mental representations are unable to account for cognitive processes – cognitive systems being dynamically adaptive to organism–environment interactions, achieving variable goals in a changing environment (Solymosi and Shook 2013). The second, and most important, reason for the theoretical abandonment of the notion of habit concerns the dualism between automatic routine and goal-directed action, with the latter gaining the greater share of attention among scholars (see for instance Dickinson 1985). Habits being reduced to rigid motor routines – that is, automatic and ateleological stimulus-response



Figure I.1 Frequency of use of the words “habit,” “representation,” and “goal” in all books in Google Books archive published from 1890 to 2008.

pairings lacking cognitive control (Bernacer, Lombo, and Murillo 2015; Seger and Spiering 2011) – motor behavior was somehow separated from the emotional and cognitive dimension. More recent neuroscientific studies are now increasingly overcoming these dichotomist views (Pessoa 2008, 2013; Uddin et al. 2014; Caruana 2017, 2019).

Habits in Contemporary Neuroscience: More than Behavioral Routines

A common reading of William James's seminal chapter of the *Principles of Psychology* is that habits – intended as nonintelligent behavioral routines characterizing the execution of well-trained actions – allow consciousness to focus on events that fall outside habitual routines. Besides the theoretical limitations of this interpretation of habits, a simple dichotomy between low-order behavioral control, intended as well-trained actions whose execution does not require deliberative control or monitoring, and voluntary, deliberative and conscious control over behavior, is also too coarse to make the concept of habit accessible to scientific study. Over time, neuroscientists emphasized further distinctions to distinguish habits from a broad spectrum of other behavioral routines. Indeed, the major characteristics commonly attributed to habits – that is, acquired through experience, performed repeatedly, executed automatically, composed of complex action sequences, and related to both action and cognitive processes (Graybiel 2008) – are shared by many repetitive behaviors: instinctual action sequences (the “fixed action patterns” studied by ethologists), human and animal rituals, purposeless stereotypes (characterized by minimal flexibility, minimal role of triggers and minimal role of reinforcement), and skill learning. Following the neuroscientist Ann Graybiel (2008), these different behaviors share a similar common ground, as they depend on a similar, interacting neural network encompassing the cortex and the basal ganglia. This interpretation is in line with the wide notion of habit, suggested by pragmatists.

From a neuroscientific point of view, habit formation is a complex and dynamic process, characterized by a dynamic shift of recruitment of different interacting networks. In the first phases – when behaviors are not yet automatic – the neural activity is predominant in brain networks involved in goal-related behaviors, with goals driven by reward signals such as food or sex. After extensive training the behavior becomes repeatedly performed, even if the value of the reward decreases. This phase indicates a switch from a goal-oriented behavior to a stimulus–response behavior, with limbic regions mostly contributing to the first phase and sensorimotor regions monitoring the latter. This finding postulates an interplay between automatic and goal-related behavior, rather than conceiving them as in a competitive relation. Notably, understanding the basis of habit learning would be of crucial importance to gain new information about human