frontiers in Sociology

HYPOTHESIS AND THEORY
published: 04 June 2018
doi: 10.3389/fsoc.2018.00014

# How Individual Habits Fit/Unfit Social Norms: From the Historical Perspective to a Neurobiological Repositioning of an Unresolved Problem

*Giuseppe Lorini*[1]* *and Francesco Marrosu*[2]

[1] *Dipartimento di Giurisprudenza, Università degli Studi di Cagliari, Cagliari, Italy,* [2] *Dipartimento di Scienze Mediche e Sanità Pubblica, Università degli Studi di Cagliari, Cagliari, Italy*

Human beings are "rule-following animals" or "nomic animals" whose behavior is strictly supported by social norms that reflect shared expectations on what a particular social context considers as appropriate behavior. Yet, little is known about the biological processes that determine how we learn to accept a particular social behavior as the most appropriate, and even less is available to highlight the deepest levels/reasons that motivate the trade-off between the daily practice of individual habits and the processes involved in conforming them to such codified social expectations as norms are. In this essay, the authors set out to investigate this particular human ability which is indeed fundamental to understanding our social world: in doing so we hypothesize that the biologically hardwired structural organization and the phenotype expressed by individual habits are the benchmark where social norms are challenged. However, by suggesting that the manifold modalities of observing or violating the norms can be subtly constrained by the individual neurobiological milieu, one can only add another problem as the essay cannot account for the very essence of this trade-off which remains largely unexplored and open to new exciting questions.

**Keywords: habits, social norms, nomic animal, neural network, normativity**

OPEN ACCESS

***Edited by:***
*Nicola Lettieri,*
*Istituto Nazionale per l'Analisi delle Politiche Pubbliche (INAPP), Italy*

***Reviewed by:***
*Stephen Turner,*
*University of South Florida, United States*
*Paolo Livrea,*
*University of Bari Medical School, Italy*

***Correspondence:**
*Giuseppe Lorini*
*lorini@unica.it*

***Specialty section:***
*This article was submitted to Evolutionary Sociology and Biosociology, a section of the journal Frontiers in Sociology*

***Received:*** *31 October 2017*
***Accepted:*** *15 May 2018*
***Published:*** *04 June 2018*

***Citation:***
*Lorini G and Marrosu F (2018) How Individual Habits Fit/Unfit Social Norms: From the Historical Perspective to a Neurobiological Repositioning of an Unresolved Problem. Front. Sociol. 3:14. doi: 10.3389/fsoc.2018.00014*

## MAN AS A "NOMIC ANIMAL"

In the last Century, four philosophers who deeply investigated our economic, social, legal and political reality—Wilfrid Sellars, Friederich August von Hayek, Robert Nozick, and John Searle—drew a new image of human beings that went against the two classical images of man as social animal and purpose-seeking, teleological, economical being. According to them, human beings are not only animals that pursue an aim, but also (or primarily) animals whose psycho-biological structure is a nomic ability, i.e., the ability to follow rules (Sellars, 1949; Hayek, 1982; Nozik, 2001; Searle, 2003). In other terms, they conceive men not only as social or political animals, but also as "rule-following animals" or "nomic animals," i.e., animals that can act in light of rules (Conte, 2000). However, in considering that social norms represent the operative constraint of individual behavior, as well as the main obstacle in the process of fulfilling the expectations of a social group (Xiang et al., 2013), the main question indeed debated since Locke (1975) revolves around the

drive that society needs to set out in order to force any individual to abandon his particular norms (or habits) in favor of a general, sometimes more abstract social norm. Following this line, we attempt to investigate the biological bases of such a construct represented by individual habits in the attempt to better understand whether such a corporeal, individualistic representation expressed by habits can establish his/her personal trade-off with the collective counterpart represented by social collective accepted norms.

## THE EXPRESSION OF HABITS AS BASIC ELEMENTS OF INDIVIDUAL BEHAVIOR: DEFINITION AND HISTORICAL PERSPECTIVE

As an exquisitely human prerogative, the process of building and following norms presumes an evolution from self-satisfactory rules to a socially concordant behavior. Moreover, during this process, the crucial step in successfully introducing a norm depends on the appeal, rather than the power, that is represented by the social shaping of individual habits, as human behavior is the result of the stratification of "acquired" habits. In order to define a habit, it seems necessary, at first sight, to follow a particular point of view in human science as the definition of habit, which in humanistic parlance sounds rather different from the one proposed by "hard" sciences (i.e., neurophysiology/neuropsychology, computer sciences, economy, etc.). However, upon closer examination, these differences are only a matter of steps or layers of abstract/experimental procedures subtended by the epistemological approach applied in habit interpretation and habit formation, and are based on a common ground which attempts to interpret these notions in a changing social context (humanistic disciplines) or ("hard" sciences) in a constitutional disposition or in a dispositional plasticity (Güell, 2014).

In a recent essay, Barandiaran and Di Paolo (2014) offer their contribution to better understanding the "genealogy" of habits. They define the positions of philosophers and psychologists regarding habits in the perspective of the dispute between associationism, which postulates that mental phenomena are formed by combination or association of simple elements, and organicism which maintains that habits are molded by a multifactorial involvement of the body, the outside world and their interactions with other habits. However, there is no doubt that these positions are now less defined and that several reasons contribute to assuming a more holistic point of view. Among these reasons, a place of prominence can be attributed to the development of neurosciences which is nowadays expressed by means of a coherent lexicon which encompasses new semantic and conceptual considerations of the notion of habits, by moving their boundaries from a pure humanistic parlance toward more articulated procedures which now also include computer sciences, neuroeconomics, neurophysiology and neuropsychology, while other disciplines have delayed supporting this field of investigation.

## SOCIAL ONTOGENESIS OF HABITS: FROM SOCIAL NORMS TO INDIVIDUAL HABITS

Among the first attempts to naturalize the origin of norm-oriented behaviors from a habit perspective, it is worth mentioning Margaret Mead's position. This author's original view is represented by her highlights in the underpinnings of the perception of the natural world in the formation of social conventions. In particular, Mead considers the practice of social rules as stratified by previous generations "either passively adopted or biologically innated." However, this hypothesis deals with an initial set of concatenated actions selected as habits that are particularly adapted to a given social contest. When the community, mostly represented in this situation by the family, together accepts or approves actions as appropriate in a particular context, the outcome of this process is represented by the formation of a basic habit. According to Mead and based on her studies in the social context of the Samoan islands, a child learns from the social community how and when it is appropriate to behave in a given situation according to a peculiar habit, though in the first steps of this process it might inappropriately "overinclude" several other situations in a primitive economic procedure "one size fits all-like" in attempting to face the first aspects of social life, such as those pertaining to family bonds.

Building a habit has several advantages at this stage. Among them, for a child, the approval of a habit can be considered a real step toward the learning process of a norm, particularly when a habit contains some "prosocial" aspects of behavior (for example: sharing a toy with other children). A second aspect, which can help raise a primitive concept of norm, is in the fact that a habit can be transmitted by imitation. Finally, a third advantage lies in the fact that, in children, forming a habit requires relatively simple conscious mechanisms (Mead, 2001).

## PHILOSOPHICAL REMARKS ON HABITS: FROM AN ORIGINAL UNITARIAN THEORY TO THE ASSOCIATIONIST/ORGANICIST DEBATE

There is widespread agreement that among the first theories aimed at investigating the concept of habit Aristotle's contribution is of crucial importance. Indeed, since then (IVth century BC) though the "*habitus*" has been investigated by other philosophers, the deep implications of Aristotle's concepts are hard to dismiss as they are surprisingly modern. Aristotle's role has a peculiar aspect in interpreting the genesis of habits in modern science as its description is in accordance with the most recent achievements of neurophysiology (see Graybiel and Smith, 2014, for good and bad habits) and because it represents a reconciliation *ante litteram* of the associationist and organicist positions vying for first place priority in the habit formation dispute. For our purpose, it is relevant that Aristotle argues that the nature of a habit revolves around its etymological roots, as this word contains both the idea of a place where one lives and an interior part with which each individual exposes himself to human society (where the term "*habitus*" is handled according

to the Latin translation of the Greek term ἦθος "*ethos*"). From back then until recent times this point of view was no longer acknowledged in a unitarian theory. Indeed, in a better attempt to describe the origins of habits, the essence of the Aristotelian view diverges according to the principles of the associationist and the organicistic schools of thought. The crucial point of this different view, is represented by the accent that the associationist theory places on the contribution of relationships between the conditions of where and when a habit become established as a rooted personal trait, while the organicist theory is more focused on its explanation of the presence or absence of structural conditions which assume a deterministic role in building a habit.

## TOWARD A NEUROPHYSIOLOGICAL THEORY OF HABIT

By setting aside the theoretical issue between the concept of habit according to the associationist vs. the organicistic theory, which can still make for an interesting debate between different philosophical schools for their general issues, we attempt to move toward the foundation of the concept of habit as intended in the frame of the main achievements of modern neurophysiology, bearing in mind that it is simply impossible to lump together the bulk of knowledge which daily builds up daily in scientific literature. With these limitations in mind we can draw the first account through the neurophysiological orientation in William James' (1890) studies. The modern aspect of James' work in describing habit formation is represented by the hypothesis that a habit is characterized by a kind of concatenated series of actions determined by the sequential activation of peculiar cerebral areas which possess an affinity of intent. Though James was influenced by the flourishing theories of conditioned reflexes and applied some ideas derived by these theories, his modern intuition revolves around the extrapolation of mostly scientifically proved theories in performances involving lower cerebral activities, and his hypothesis that concatenated events can also work in the domain of the higher functions. This intuition, then broadly unsupported by experimental data, nonetheless introduces the concept that in motor performance, habit allows for a kind of freedom that is extremely useful for "higher" integration ("habit [. . . ] diminishes the conscious attention with which our acts are performed"). The net implication of James' theory is now very different from a canonic associationist view, given the implicit possibility that a habit allows for the personalization of a certain performance: the example of a musician whose practice frees him from technical details so that the artistic interpretation of the melody can be "released" is an elegant example of the concept. Indeed, a century after James' description, thanks to advancements in neurophysiological research, we can now prove the anatomical and physiological basis of habit formation with a greater degree of detail. These acquisitions can help us in the attempt to answer such questions as the "what" (ontologic aspects) and "where" (genetic aspects) of habits and contain, to some extent, the theoretical bases of the concept of plasticity of brain activity.

## HABIT AS EXPRESSION OF "EMBODIED MIND": BETWEEN PHENOMENOLOGY AND NEUROPHYSIOLOGY

The concept of habit has recently been revisited in view of its role in the debate concerning the relationship between world and body (Varela et al., 1991). This interesting position has been developed by integrating neurophysiological notions with previously acquired philosophical investigations. The works of Gallagher and Zahavi (2008) are particularly worthy of mention in this debate as they make a fruitful effort to re-interpret the phenomenological positions of Maurice Merleau-Ponty (2012), who considers the habit as a "pre-reflective" attitude of the body situated (or "thrown" in phenomenological parlance) in the world. According to Gallagher and Zahavi, this original position needs a specific "corporeal style" which is represented by the strategy used by the diadic couple world-brain in building a habit. From the neurophysiological point of view, this theory is also an attempt to give some value to the flow of proprioceptive activity which is only partially recorded by the conscience, but that can in some way be detected and expressed by habits. In addition, it is worth noting that, in their studies, Gallagher and Zahavi mention the term "neuroplasticity" when they discuss habits, with the specific intent of incorporating Merleau-Ponty's concepts related to the perception of the motoric/situational consciousness (i.e., the internal acknowledgment of the corporeal possibility to perform an act: "I can, then I am") more into neurophysiological than philosophical language, by introducing the concept of neural plasticity. Independently of Gallagher and Zahavi, but in the same theoretical line, Barrett (2014) suggests that habits express a dynamic scale-free neuronal performance which is determined by a plastic vectorial attractor. In essence, this theory suggests a semi-deterministic genesis of a particular habit as the outcome of the previous history of neuronal trajectories, which is involved in the roles of personalized complex conditions encoded (and in some way constrained) by a given individual neuronal networking: a concept which recalls the role of disequilibrium in the construction of multiple singularities that, however, as part of complex systems "carry their history on their back" (Prigogine and Stengers, 1984).

Taken together, these theories hypothesize that a habit is a temporal construct which depends on the accumulation of a certain amount of significant personal data which are articulated in temporal scales and work in a plastic modality in accordance with any possible remodulatory requests of the networks involved.

## CEREBRAL AREAS INVOLVED IN STRUCTURING HABITS: BASIC PLAYERS

As previously mentioned, James' basic concept is strongly influenced by the pavlovian theories that are at the core of his hypothesis: sequences of actions are the basic expression of habits. However, it is important to note that in support of the concatenated actions aimed at achieving a purpose, the pavlovian system assigns "values" only to a small set of behaviors that are

evolutionarily appropriate responses to particular environmental *stimuli*, while habit systems are instead built for learning and, through repeated training, assigning values to a much larger and more differentiated number of actions (Rangel et al., 2008). Thus, an emergent property of concatenated actions, either acted in a pavlovian "small sets" mode or in habits with a more complex and socially interactive performances, is represented by the concept of outcome and by its correlated value. If we consider value as a new agent in describing the significance of habit, we can attempt to articulate a better definition of habit, based on objective experimental data detailed thanks to the most recent advancements in neurophysiology and neuroanatomy. Accordingly, the most relevant players in structuring habit are represented by the involvement of the basal ganglia along with a vast correlated cortical circuitry. In this new view, habits are considered the result of "sequential, repetitive, motor or cognitive behaviors elicited by external or internal triggers that once released can go to completion without constant conscious oversight" (Smith and Graybiel, 2014, 2016). This definition echoes the theoretical approach to habits proposed by James; however, a crucial implementation is that it highlights the role of a hitherto incompletely studied network of structures, represented by the basal ganglia. In addition, this new wave of studies disentangles the pavlovian component of habit which was borrowed by James from his knowledge of cutting edge XIX century neurophysiology. Together, these cognitions now shift the attention of the scientific community toward a new role of anatomical structures once restrained to the ancillary position of areas dedicated to movement refinement and extends their function to a more holistic view by including many other correlated brain areas (Smith and Graybiel, 2013). As an example of this new view, the initial formation of a habit can be detected by deep arrays of electrode recordings, performed when experimental animals begin to learn a particular motor sequence entailed with a certain value which in the future can be eventually maintained and refined. In this particular condition, the sensorial motor part of the striatum shows high neuron activity, while when the sequence is completely mastered by the animal, the same neurons only show an initial and final activity in relationship with the beginning and the end of the motor sequence (Graybiel, 2008). In addition, given the improvements in knowledge of the brain regions involved in habit formation and following experimental observations both in animal models and in humans, several studies now show that these areas are involved in a substantial contribution correlated to habit formation, which is represented by the process of coupling the decision of pursuing a given habit and its correlation with an intrinsic value represented by the social interference of this process (Di Francesco et al., 2007; Rangel et al., 2008). This perspective opens new interpretations and spurs parallel studies in several disciplines, since the consideration of habits according to an intrinsically expressed value correlated with a habit, biases the process of decision making and, from a neurophysiological perspective, brings these anatomical areas correlated to higher functions involved in the motivational drive process into focus: all in all this view paves the way for disciplines more devoted to highlighting the neurophysiology of areas linked to fear, satisfaction, and in hedonic risk-correlated responses (i.e., risk-seeking habits) that characterize any stable practice of habits. In other words, neurosciences are now compelled to design a framework which can outline the state of the art of the neural processes involved in habit formation, keeping in mind the complex implications correlated with the value of actions and their consequences.

## HABIT OUTCOME REPRESENTATION IN NEURONAL NETWORKS: ADVANCED PLAYERS

Based on the new perspective, the concept of habit now extends its role beyond the simple sequential motor assembly of movements or gestures that characterize its first definition. Now, the simple semi-automatic motor program appears subsidiary to freeing the individual from repetitive series of acts and allows for a better understanding of the habit, which moves from a subjective frame to refine its outcome in view of a social (i.e., normated) acknowledgment. Together, these peculiarities related to habits targeted to a social context, tend to acquire patterned characteristics, and possess a value-based decision making. From a neurophysiological point of view, recent studies on experimental animals show that the dorsolateral striatum plays an important role in controlling habits (Balleine, 2005; Yin and Knowlton, 2006). Furthermore, a real breakthrough in this issue has been the discovery of the role of neurotransmitters in refining the outcome of a habit. Particularly, dopamine neurotransmitter has been shown to modulate the activity of this neural circuitry in the evaluation of the outcome of a habit, as the projections of dopamine neurons into the striatal areas contribute to learning how a certain value is correlated to habits and how such value is "graded" in taking a risk in order to achieve a result (Mirenowicz and Schultz, 1994). In addition, it has been shown that the cortico-thalamic loops are involved in the representation of the stimulus-response. This complex neural circuitry is completed by an area positioned in the infra limbic cortex which contributes at least in part to the emotional component of the habit (Coutureau and Killcross, 2003; Killcross and Coutureau, 2003). This particular neural circuitry might be anticorrelated with the adequacy of habits to social norm, as emotional components are more involved in individual achievements, which in turn are more correlated to personal hedonistic pursuit and less prone to be entrained by overt or covert "social trading play" between habit and norms. This neuro-anatomo-physiological knowledge highlights the notion that the most significant contribution in the establishment of a habit is represented in the brain by a multifarious set of areas which encompasses the subcortical regions and the cortical circuits that engage limbic and frontal neuronal assemblies (Smith and Graybiel, 2013). The latter structures are crucial in elaborating the previsions correlated with complex decisions but are also involved in suppressing or expressing particular aspects of habits, which includes their role in the evaluation of economical outcome (Padoa-Schioppa and Assad, 2006). The simplified scheme above needs to be completed by the areas

which regulate the levels of agreement, accomplishment and satisfaction that the practice of a habit determines in individuals. This role is played by the nucleus accumbens, a small neuronal agglomerate positioned in the anterior and deep ventral portion of the striatum (Schultz et al., 1997). Though small, this nucleus is of paramount importance in modulating the mechanisms of habit reinforcement and is activated in altered psycho-somatic conditions related to virtually all neuropathological dynamics involved in the process of addiction (Scofield et al., 2016). A crucial role in this neuronal architecture is represented by the amygdala, an almond-shaped neuronal agglomerate (amygdala = almond in Greek) highly connected to cortical (frontal and cingulate areas) as well as to subcortical structures (striatum and accumbens) in relation to its role in detecting fear and all correlated conditions (Ohman, 2005). Although this part of the brain was originally involved in basic behaviors ("fight or flight"), it is now also considered important in decision making, in view of the extensive network it takes part in LeDoux (2003). For the purpose of this study, the role of the amygdala and its extensive connections is represented by the perception of the risk associated to certain habits (Preuschoff et al., 2006). Taken together, it is broadly accepted that the malfunction of this circuit is important in understanding the neurophysiopathology of the alteration of habits. Moreover, it has been observed that the integration between the two hemispheres plays a crucial role in the full control of moral judgment. Though scarcely investigated, this problem has been recently highlighted by a functional study in split-brain subjects (Steckler et al., 2017).

## FROM NEUROPHYSIOLOGY TO NEUROPATHOLOGY: THE ESTABLISHMENT OF "BAD" HABITS

By suggesting that for any individual a habit is the possibility of acting according to "a condition either well or ill disposed," Aristotle underpins the concept of bad habit as formally present in culture since classic Greek philosophy (Aristotle, 2007, *Met.* V 1022b). Though not expressed in scientific terminology, in the *Confessions* Augustine (1997) also gives an impressive account of bad habits as related to his personal history. In his view, bad habits are disadvantageous to the owner, practiced despite violation of the most common social norms and are difficult to break (unless by Divine mediation, in this specific case). Modern neurophysiology can now attribute scientific underpinnings that seem to "validate" Augustine's account on bad habits. Indeed, an extensive amount of experimental and psychopathological data highlights the etiopathology of bad habits with clinical and basic ("hard sciences") studies. To summarize the principal achievements which contribute to the understanding of the genesis of bad habits, one can begin with the classic observation of the famous case of Phineas Gage, the railroad worker who suffered a traumatic lesion involving a large portion of the prefrontal areas and turned, from pious and devoted husband into a callous blasphemer and quarrelsome gambler ("Phineas Gage was no longer Phineas Gage" according to his doctor). This case report was detailed by the elegant works of Damasio et al. (1994) whose studies represent one of the most eloquent support on the role played by lesions targeting cerebral areas peculiar to the formation of habits that violate social norms. In particular, Damasio described the crucial importance of the brain's networking correlated with the damage of ventromedial frontal areas in habit shifting (usually toward bad habits). Though Damasio suggested that in the case of Phineas Gage the frontal lesion involved extensive connections which include the previously mentioned sub-cortical regions (striatal areas, accumbens and amygdala, to mention the most relevant), in common practice anatomical damage cannot entirely explain the neuropathology of bad habits. Indeed, most of the common pathogenetic aspects of bad habits reveal a dynamic aspect which eludes the above-mentioned neuroanatomic correlates, as well as those induced by macroscopic lesions (i.e., organic alterations, tumors, post-traumatic scars).

Habit alterations without apparent anatomic lesions represent, by far, the most common problem affecting an unknown number of subjects (usually considered "borderline" or "psychopathic"), and involve hosts of psychologists, psychiatrists and neurologists, not to mention the excruciating correlated social burden. The main focus of modern research on bad habits is on the study of addiction, either including addictions related to "recreational" or compound abuse (cocaine, heroin, "smart drugs," alcohol, to mention the most widespread) or the various forms of "bad" and antisocial behavior (i.e., gambling, sexual deviations) (Hymanand et al., 2016). Within the study of "bad habits," pathological gambling is of particular interest, besides its dramatic social aspects, because it highlights the role of unaltered macroscopic and microscopic brain structures and calls for a microscale dysregulation orchestrated by neurochemical causes as the basic alterations in absence of anatomical lesions. In particular, case studies of these addictions indicate a major player in this pathophysiology as being the alteration in production, release, uptake and proper distribution of the neurotransmitter dopamine (Di Chiara and Imperato, 1988; Redish, 2004) according to a mechanism which is replayed by most forms of addiction (Kauer and Malenka, 2007). Moreover, the altered mechanism related to the neurotransmitter dopamine in the neural circuitry along with its pivotal role in grading the satisfaction of a habit outcome and its eventual reward, is considered the most reliable driving force in determining the turn point from "norm-oriented habits" to the acquisition and maintenance of bad or antisocial normative-deviant habits. However, the same mechanism can activate the pleasurable feeling which accompanies a habit involving being observant in following certain highly common norms (i.e., duties that are related to a sense of belonging in the community). Accordingly, in a relevant contribution to this investigation, Montague highlights the role of dopamine in midbrain structures, in encoding a kind of neurophysiological grammar aimed at predicting value and reward correlated with habits (Montague et al., 1996; see also D'Ardenne et al., 2008). Though these structures are considered as part of an extensive network which encompasses the striatum and the orbito-frontal cortex, it has been shown that when the habit is tagged with a violation of a social norm, it induces the activation of an area in between,

represented by the anterior insula, which is also crucial in the modulation of subjective feelings (Montague and Lohrenz, 2007). This observation is interesting in that the involvement of connections of the prefrontal areas with the insula may be crucial in determining the individual satisfactory prevalence of a deviated habit outcome (immediate reward) in spite of a more "frontalised" interindividual normated purpose (delayed reward) (Killcross and Coutureau, 2003).

Though this process is under a complicated interindividual variable network of control which calls up the anterior part of the frontal cortex to participate in the process, it is likely that the voluntary actions performed in normated modality can contrast a bad habit according to the power of the "shares" owned by the portfolio of the most rationality-driven frontal areas in a given individual at a given moment. Thus, it is probable, for example, that strong educational constraints can help contrast bad habits. Unfortunately, this process belongs to those highly irregular operations that are dictated by the circumstantial, non-linear work of the neuronal assembly involved. In other words, the chances that in an ambiguous situation norm can overcome habit (either bad or good) are unpredictable as in particular circumstances any intrusive and momentaneous thought can dominate, even against the most stringent logical evidence ("winner takes all" mental procedure).

Based on these studies, the process of addiction in absence of macroscopic or microscopic brain lesions represents a dynamic alteration of the reward system in that both these situations bias the dopaminergic circuit which, after losing the physiological feedback between the subcortical and cortical areas, is in a condition of disequilibrium between midbrain and orbito-frontal cortex. These consequences are also described as correlated with iatrogenic causes: for instance, some pharmacologic treatments of Parkinson's disease may induce gambling, which is related to the altered distribution of dopamine in the above-mentioned areas (Heiden et al., 2017). This alteration determines a strong hedonistic feeling correlated to the habit of gambling which is an integral part of maintaining such a bad habit. Moreover, it is remarkable that drugs which depress the hedonistic components of a bad habit also decrease the "craving" related to the appetitive gambling behavior in these subjects (Muroni et al., 2011; Bortolato et al., 2012), thus suggesting a plastic modulation by the pleiotropic assembly of neurotransmitters and neuromodulators, though only few of them have hitherto been studied (Devoto et al., 2012). Furthermore, neuropathologists also consider "bad" habits a number of involuntary complex dystonic attitudes represented by the altered postures observed in several neurodegenerative conditions (i.e., Parkinson disease, Huntington's chorea, Tourette's syndrome, among the most common) (Graybiel and Smith, 2014). It is peculiar that, in association with the more complex bad habits such as gambling, in these conditions apparently involving only muscular districts, a peculiar hedonistic counterpart correlated to the moments preceding the motoric expression of the Tourettic tics has also been described. Indeed, the accounts given by these subjects underlie the fact that the patient experiences the moment preceding the tic explosion as "a paramount hedonistic moment" (personal communication to FM by a Tourettic patient) despite their somewhat impressive and very particular violence; for instance, many subjects affected by Tourette syndrome reported cervical herniations induced by the violent displacement of the neck propelled forward during tics. Interestingly, all the pharmacologic treatment for these altered patterns of movement are based on reducing the quantum of dopamine available in neural synaptic clefts of the striatal-cortical circuitry: for instance they can be reduced by the same drugs used to contrast the parkinsonian gambler (Bortolato et al., 2007). Finally, while describing pathologic habits in relation to breaking norms, a peculiar mention should be reserved for the cases of frontotemporal dementia. In this globally-spreading disease, there is a deep alteration of the neuronal circuitries dedicated to keeping the most basic impulses under control; as mentioned above, these areas are located in frontal regions and, partially, in certain deep areas of the temporal lobe (amygdala and accumbens). The neurodegenerative processes, which frontotemporal dementia belongs to, are widespread neurological diseases characterized by neuroanatomical alterations still scarcely known despite massive economic investments and considerable scientific efforts.

## NORMS AND HABITS: COMMON NEURAL NETWORKS BUT DIFFERENT FUNCTIONAL TRAJECTORIES

From the issues mentioned above, it is hard at first sight to dismiss the fact that there is a striking similarity between many cerebral areas involved in the acquisition of a habit (either considered as simple sequential motor patterns or more complex behaviors) and in the brain regions involved in neural operations that represent how social norms are processed (Xiang et al., 2013). However, in contrast with the sequential pattern of brain networks activated by habits (either good, neutral or bad), we don't yet know whether the representation of a norm is activated in those neural circuitries in the form of a norm-oriented complex of behaviors or in an abstract tendency of a social tropism toward the observance of a norm. The first hypothesis echoes William James' theories that norms are sequentially expressed by behaviors which are an adequate reflection of rules that are expected to act according to a dominant position: in simple words, the norm looks like a kind of super habit. The latter position, however, deals with a tautological difficulty as it needs to operate following the speculation that certain brain areas can process abstract reasons whose final outcome assigns to these areas abstract procedures aimed at driving the best way to handle a given habit. In order to align norms with habits, the role of a high social cohesion which depends in part on the grade of interindividual social accountability must be acknowledged. Accordingly, the last part of the procedure represents the main difference between a norm-oriented complex of behaviors and the abstract tendency toward the observance of a norm.

Moreover, in both cases, the interplay between norm and habit represents a complex brain process of internal evaluation which ought to consider the implicit consequences of following or breaking a norm because of a particular habit. However,

neuroscientists' answer to these problems are largely incomplete as we have just begun to grasp the involvement of these brain networks in terms of neurophysiological operations involving sets of cortico-cortical and cortical-subcortical processes. It is worth mentioning that these mostly neurochemically regulated processes are aimed at appropriately signaling the outcome (the value) correlated to habits and norms, as well as the different activities of these areas while following a norm and/or breaking it (O'Doherty et al., 2003). Together, experimental and clinical data point to the same areas as major players of both habits and norms: however, while we are more informed on the functions of these neural circuits when they operate in habit mode, we can only understand that the areas involved in building the bricks of the norm are the same as those used in building habit (Xiang et al., 2013), but to date it is difficult to explain how neurochemical/neuroelectric recordings can help in differentiating the subtle differences between norms and habits in the same areas.

In addition, a speculative aspect of this relation between norm and habit can be illustrated by the role of particular brain areas in situations when there is an initial fuzzy possibility of a rough evaluation of a situation (as proposed by Kahneman and Tversky, 1979). In this case it has been found that the insula is involved in processing unexpected results issued by particular situations and unrelated to the fact that they are better or worse than expected (Preuschoff et al., 2008). In this situation, one can speculate that the norm can drive the habit according to a neuroeconomic distinction between risk and ambiguity that can be handled more successfully by a norm which might orient the habit in a situation of risk or ambiguity. Unfortunately, though many models of evaluation in conditions of risk/ambiguity are available, none of them has yet obtained powerful statistical support, despite some elegant approach (Hsu et al., 2005).

However, the estimation of the evaluation systems operated by habits increases with experience, so that the norms can help a habit, which learns slowly, and that "might forecast the value of actions incorrectly immediately after a change in the action-reward contingencies" (Rangel et al., 2008) by a frame of vectorial approximation of different valuations. In addition, norms are also gauged by multifactorial variables that together can bias a "correct" habit and that suggest the stronger role of the cortical areas in modulating norms. A semi-serious example of this possibility can be found in what is now a classic experiment lead with a team of fMRI scientists which showed that a solid habit such as the appreciation of the quality of wine as rated by professional sommeliers, can be erroneously rated by introducing a false price in the bottle; following the "consequent" high rating of less fine wines, the fMRI showed a surprising activation of brain areas involved in the positive appreciation of good taste, supporting this (Plassmann et al., 2008).

As for the neuroscientist perspective, it is possible that our most sophisticated neuroimages (i.e., PET, fMRI) are inadequate for achieving a contemporary detection of brain activity (BOLD effect) and for giving an account on how the priority of certain areas common to norm and habit are depicted under a detailed time-locked sequence of the networks involved. As example, one critical aspect involving this problem is the activation of electrical signals (synaptic activity which eventually releases quantums of neurotransmitters) that according to variations in their frequency and firing rate can change in time without variations in their number and metabolic costs: the net result of this operation (a very common economic operation of the brain) is that though the energetic cost is the same (thus PET and fMRI detect the same metabolic rate), the brain operation in terms of engagement of neural networks can be very different in spite of zero changes in metabolic rates (Buzsaki, 2006).

Setting aside the pathologically deranged habits, in a physiological situation it is likely that a norm involving a social value helps an individual habit to rely on common rules in making decisions. This particular point has been raised in other terms by modern neurophysiology. Montague and Lohrenz (2007) show that using a computational depiction of an economic use of norms (i.e., the "Ultimate Game") is possible to demonstrate that social emotions are part of the strict observation of a norm and that the sense of guilt is correlated with the violation of norms. The "Ultimate Game" is based on how fair the division of a sum of money (real money) is between an owner granted the sum and allowed to share it freely with another person. As the money is merely granted, without no counterpart, the only rule for a definitive possession of that money for both players is determined by the acceptance of the owner's proposal for how the sum will eventually be divided: any refusal of the proposal by the owner means that both players will end up without money. Despite the fact that theoretically any sum is good (after all it is a gift!), when it is divided unfairly (i.e., 50 dollars divided in 40 to the owner and 10 for the other person), the proposal is usually rejected. The social group may also react to other emotional responses (usually variations of anger), when they are in the presence of individuals who violate a tacit norm, since one expects a proposal of no less than 40% (60% for the owner) (Sanfey et al., 2014).

## CONCLUSIONS

The purpose of this study is to provide an outline of the role that neurosciences play in understanding the genesis of habits and the feedback which results from the interplay between social norms and habits. In this far from exhaustive essay, we show that several brain areas come into common play to define both habits and behaviors aimed at social and legal norms. A crucial part of this study are the new considerations spurred by the notion that an evaluation system operates inside the practice of a habit and, at least in physiological conditions, that this system operates a continuous trade-off with norms. Indeed, though this part seems the least well-grasped by neuroscientists, nonetheless our current knowledge represents a promising field of investigation for the future. Following this line of research, another relevant issue related to neurophysiological investigation is whether the incorporation of negative and positive components of a particular habit is correlated with the areas involved in processing uncertain outcomes (Daw et al., 2005). The next years will play an important role in understanding several open questions. Among them the following three seem to be more urgent:

(i) considering the economic aspect of making decisions (either according to short or/and long temporal intervals), how can the habit be modulated by norms and/or with what type of gradation can a norm act to modulate habits?

(ii) are norm-oriented habits represented as a kind of brain "default-mode" (a genetic blueprint in neural networks) for the expression of good habits or is the observance of a norm mirrored in the beginning by imitation, and then autonomously developed?

(iii) given that the number of individual habits exceeds the number of social norms, how could a habit without a specific norm tend toward good?

Indeed, the answers to these questions are relevant for solving new correlated problems. For neuroscientists, investigating on how norms are correlated with habits in several social contexts is of great interest. Confrontations and collaborations of multiple disciplines, including the philosophy of law, should be considered in the perspective that work in these areas can give new outlooks, not only in the reality of relations between norms and habits, but also in the process of opening up new strategies and fresh possibilities to neuroscientific investigation on the biological mechanisms subtended by these issues.

These examples suggest the need of an interdisciplinary collaboration in investigating norms vs. habits, a task that can indeed be exploited according the peculiarities of each discipline with their own information, notions and debates. Old and new disciplines (social sciences, neuroeconomics, computer sciences, neurosciences) will be draw the most beneficial advantages from this program.

## AUTHOR CONTRIBUTIONS

This article is the result of joint research undertaken by the two authors.

## FUNDING

Supplementary Research Fund 2016, Dipartimento di Scienze Mediche e Sanità Pubblica, University of Cagliari.

## REFERENCES

Aristotle (2007). *Metaphysics.* Mineola, NY: Dover.

Augustine (Saint). (1997). *The Works of Saint Augustine: A Translation for the 21st Century*. New York, NY: New City press.

Balleine, B. W. (2005). Neural bases of food-seeking: affect, arousal and reward in corticostriatolimbic circuits. *Physiol. Behav.* 86, 717–730. doi: 10.1016/j.physbeh.2005.08.061

Barandiaran, X. E., and Di Paolo, E. A. (2014). A genealogical map of the concept of habit. *Front. Hum. Neurosci.* 8:522. doi: 10.3389/fnhum.2014.00522

Barrett, N. F. (2014). A dynamic systems view of habits. *Front. Hum. Neurosci.* 8:682. doi: 10.3389/fnhum.2014.00682

Bortolato, M., Cannas, A., Solla, P., Bini, V., Puligheddu, M., and Marrosu, F. (2012). Finasteride attenuates pathological gambling in patients with Parkinson disease. *J. Clin. Psychopharmacol.* 32, 424–425. doi: 10.1097/JCP.0b013e3182549c2a

Bortolato, M., Muroni, A., and Marrosu, F. (2007). Treatment of Tourette's syndrome with finasteride. *Am. J. Psychiatry* 164, 1914–1915. doi: 10.1176/appi.ajp.2007.07060978

Buzsaki, G. (2006). *Brain Rhythms*. New York, NY: Oxford University Press.

Conte, A. G. (2000). Nomotropismo: agire in funzione di regole. *Sociol. Diritto* 27, 1–27.

Coutureau, E., and Killcross, S. (2003). Inactivation of the infralimbic prefrontal cortex reinstates goal-directed responding in overtrained rats. *Behav. Brain Res.* 146, 167–174. doi: 10.1016/j.bbr.2003.09.025

D'Ardenne, K., McClure, S. M., Nystrom, L. E., and Cohen, J. D. (2008). BOLD responses reflecting dopaminergic signals in the human ventral tegmental area. *Science* 319, 1264–1267. doi: 10.1126/science.1150605

Damasio, H., Grabowski, T., Frank, R., Galaburda, A. M., and Damasio, A. R. (1994). The return of Phineas Gage: clues about the brain from the skull of a famous patient. *Science* 264, 1102–1105.

Daw, N. D., Niv, Y., and Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat. Neurosci.* 8, 1704–1711. doi: 10.1038/nn1560

Devoto, P., Frau, R., Bini, V., Pillolla, G., Saba, P., Flore, G., et al. (2012). Inhibition of 5a-reductase in the nucleus accumbens counters sensorimotor gating deficits induced by dopaminergic activation. *Psychoneuroendocrinology* 37, 1630–1645. doi: 10.1016/j.psyneuen.2011.09.018

Di Chiara, G., and Imperato, A. (1988). Drugs abused by humans preferentially increase synaptic dopamine concentrations in the mesolimbic system of freely moving rats. *Proc. Natl. Acad. Sci. U.S.A.* 85, 5274–5278. doi: 10.1073/pnas.85.14.5274

Di Francesco, M., Motterini, M., and Colombo, M. (2007). In search of the neurobiological basis of decision making: explanation, reduction and emergence. *Funct. Neurol.* 22, 197–204.

Gallagher, S., and Zahavi, D. (2008). *The Phenomenological Mind: An Introduction to Philosophy of Mind and Cognitive Science*. New York, NY: Routledge.

Graybiel, A. M. (2008). Habits, rituals, and the evaluative brain. *Ann. Rev. Neurosci.* 31, 359–387. doi: 10.1146/annurev.neuro.29.051605.112851

Graybiel, A. M., and Smith, K. S. (2014). Good habits, bad habits. *Sci. Am.* 310, 38–43. doi: 10.1038/scientificamerican0614-38

Güell, F. (2014). Pre-dispositional constitution and plastic disposition: toward a more adequate descriptive framework for the notions of habits, learning and plasticity. *Front. Hum. Neurosci.* 8:341. doi: 10.3389/fnhum.2014.00341

Hayek, F. A. (1982). *Law, Legislation and Liberty: A New Statement of the Liberal Principles of Justice and Political Economy*. London: Routledge & Kegan Paul.

Heiden, P., Heinz, A., and Romanczuk-Seiferth, N. (2017). Pathological gambling in Parkinson's disease: what are the risk factors and what is the role of impulsivity? *Eur. J. Neurosci.* 45, 67–72. doi: 10.1111/ejn.13396

Hsu, M., Bhatt, M., Adolphs, R., Tranel, D., and Camerer, C. F. (2005). Neural systems responding to degrees of uncertainty in human decision-making. *Science* 310, 1680–1683. doi: 10.1126/science.1115327

Hymanand, S. E., Malenka, R. C., and Nestler, E. J. (2016). Neural mechanisms of addiction: the role of reward-related learning and memory. *Annu. Rev. Neurosci.* 29, 565–598. doi: 10.1146/annurev.neuro.29.051605.113009

James, W. (1890). *The Principles of Psychology*. New York, NY: Holt.

Kahneman, D., and Tversky, A. (1979). Prospect theory: an analysis of decision under risk. *Econometrica* 47, 263–292.

Kauer, J. A., and Malenka, R. C. (2007). Synaptic plasticity and addiction. *Nat. Rev. Neurosci.* 8, 844–858. doi: 10.1038/nrn2234

Killcross, S., and Coutureau, E. (2003). Coordination of actions and habits in the medial prefrontal cortex of rats. *Cereb. Cortex* 13, 400–408. doi: 10.1093/cercor/13.4.400

LeDoux, J. (2003). The emotional brain, fear, and the amygdala. *Cell. Mol. Neurobiol.* 23, 727–738. doi: 10.1023/A:1025048802629

Locke, J. (1975). *An Essay Concerning Human Understanding,* ed P. H. Nidditch. Oxford: Oxford University Press.

Mead, M. (2001). *Coming of Age in Samoa: A Psychological Study of Primitive Youth for Western Civilization*. New York, NY: Harper-Collins, Spring.

Merleau-Ponty, M. (2012). *The Phenomenology of Perception*. Transl. by D. A. Landes. London; New York, NY: Routledge.

Mirenowicz, J., and Schultz, W. (1994). Importance of unpredictability for reward responses in primate dopamine neurons. *J. Neurophysiol.* 72, 1024–1027. doi: 10.1152/jn.1994.72.2.1024

Montague, P. R., Dayan, P., and Sejnowski, T. J. (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J. Neurosci.* 16, 1936–1947. doi: 10.1523/JNEUROSCI.16-05-01936.1996

Montague, P. R., and Lohrenz, T. (2007). To detect and correct norm violation and their enforcement. *Neuron* 56, 185–196. doi: 10.1016/j.neuron.2007.09.020

Muroni, A., Paba, S., Puligheddu, M., Marrosu, F., and Bortolato, M. (2011). A preliminary study of finasteride in Tourette syndrome. *Mov. Disord.* 26, 2146–2147. doi: 10.1002/mds.23810

Nozik, R. (2001). *Invariances. The Structure of the Objective World*. Cambridge, MA: Harvard University Press.

O'Doherty, J. P., Dayan, P., Friston, K., Critchley, H., and Dolan, R. J. (2003). Temporal difference models and reward-related learning in the human brain. *Neuron* 38, 329–337. doi: 10.1016/S0896-6273(03)00169-7

Ohman, A. (2005). The role of the amygdala in human fear: automatic detection of threat. *Psychoneuroendocrinology* 30, 953–958. doi: 10.1016/j.psyneuen.2005.03.019

Padoa-Schioppa, C., and Assad, J. A. (2006). Neurons in the orbitofrontal cortex encode economic value. *Nature* 441, 223–226. doi: 10.1038/nature04676

Plassmann, H., O'Doherty, J., Shiv, B., and Rangel, A. (2008). Marketing actions can modulate neural representations of experienced pleasantness. *Proc. Natl. Acad. Sci. U.S.A.* 105, 1050–1054. doi: 10.1073/pnas.0706929105

Preuschoff, K., Bossaerts, P., and Quartz, S. R. (2006). Neural differentiation of expected reward and risk in human subcortical structures. *Neuron* 51, 381–390. doi: 10.1016/j.neuron.2006.06.024

Preuschoff, K., Quartz, S. R., and Bossaerts, P. (2008). Human insula activation reflects risk prediction errors as well as risk. *J. Neurosci.* 28, 2745–2752. doi: 10.1523/JNEUROSCI.4286-07.2008

Prigogine, I., and Stengers, I. (1984). *Order Out of Chaos*. New York, NY: Bentam.

Rangel, A., Camerer, C., and Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nat. Rev. Neurosci.* 9, 545–556. doi: 10.1038/nrn2357

Redish, A. D. (2004). Addiction as a computational process gone awry. *Science* 306, 1944–1947. doi: 10.1126/science.1102384

Sanfey, A. G., Stallen, M., and Chang, L. J. (2014). Norms and expectations in social decision-making. *Trends Cogn. Sci.* 18, 172–174. doi: 10.1016/j.tics.2014.01.011

Schultz, W., Dayan, P., and Montague, P. R. (1997). A neural substrate of prediction and reward. *Science* 275, 1593–1599. doi: 10.1126/science.275.5306.1593

Scofield, M. D., Heinsbroek, J. A., Gipson, C. D., Kupchik, Y. M., Spencer, S., Smith, A. C., Kupchik, Y. M., et al. (2016). The nucleus accumbens: mechanisms of addiction across drug classes reflect the importance of glutamate homeostasis. *Pharmacol. Rev.* 68, 816–871. doi: 10.1124/pr.116.012484

Searle, J. R. (2003). "Social ontology and political power," in *Socializing Metaphysics: The Nature of Social Reality,* ed F. F. Schmitt (Oxford: Rowman & Littlefield), 195–210.

Sellars, W. (1949). "Language, rules and behavior," in *John Dewey: Philosopher of Science and Freedom*, ed S. Hook (New York, NY: Dial Press), 289–315. (Reprinted in *Pure Pragmatics and Possible Worlds. The Early Essays of Wilfrid Sellars*, edited by Jeffrey F. Sicha, 129-55. Atascadero, CA: Ridgeview, 1980).

Smith, K. S., and Graybiel, A. M. (2013). A dual operator view of habitual behavior reflecting cortical and striatal dynamics. *Neuron* 79, 361–374. doi: 10.1016/j.neuron.2013.05.038

Smith, K. S., and Graybiel, A. M. (2014). Investigating habits: strategies, technologies and models. *Front. Behav. Neurosci.* 8:39. doi: 10.3389/fnbeh.2014.00039

Smith, K. S., and Graybiel, A. M. (2016). Habit formation. *Dialogues Clin. Neurosci.* 18, 33–43.

Steckler, C. M., Hamlin, J. K., Miller, M. B., King, D., and Kingstone, A. (2017). Moral judgement by the disconnected left and right cerebral hemispheres: a split-brain investigation. *R. Soc. Open Sci.* 4:170172. doi: 10.1098/rsos.170172

Varela, F. J., Thompson, E., and Rosch, E. (1991). *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge, MA: MIT Press.

Xiang, T., Lohrenz, T., and Read Montague, P. R., (2013). Computational substrates of norms and their violations during social exchange. *J. Neurosci.* 33, 1099–108. doi: 10.1523/JNEUROSCI.1642-12.2013

Yin, H. H., and Knowlton, B. J. (2006). The role of the basal ganglia in habit formation. *Nat. Rev. Neurosci.* 7, 464–476. doi: 10.1038/nrn1919

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.