

Is There a Neurological Foundation for Institutions? Mirroring, Sympathy, and Institutional Design

L. Lynne Kiesling*

Department of Economics
Northwestern University
2001 Sheridan Road
Evanston, Illinois 60208
847.491.8250
lkiesling@northwestern.edu

June 2010

Abstract: Is there a neurological foundation for institutions? Recent neuroscience research has expanded our understanding of the cognitive foundations of economic decision-making at the individual level, but the implications of this research for institutions and institutional design have not yet been fully explored using the tools and concepts of new institutional economics. This paper explores these implications generally, focusing in part on the implications of mirror neuron systems research and the possible connection of the mirror system with the concepts of sympathy and mutual sympathy that Smith (1759) characterized as being at the foundation of evolved human institutions that enable civil society. Smith's connection between mutual sympathy and institutions also generates some hypotheses regarding the role of social distance in the form that institutions can take (an analysis that is similar in many ways to that of Hayek (1973) on laws and extended orders), and why those institutions change, for example, in the move from personal to impersonal exchange. The connection of mirror neurons with these concepts is consistent with their hypotheses, as well as the experimental research on the effects of social distance and social cues in dictator games.

*Prepared for the annual meetings of the International Society for New Institutional Economics. I am grateful to Matthew Coffey, Steven Horwitz, Yannay Spitzer, and Bart Wilson for conversations that helped me crystallize my thoughts.

I. Introduction

One of the persistent questions of political economy over the past 400 years has been the extent to which man is a social animal by nature, and the social (economic, legal) institutions we devise to promote the benefits of that sociability. Do we devise institutions to curb our innate, atavistic self-interest-maximizing tendencies (narrowly defined) and escape a Hobbesian state of nature, or do we devise institutions to enhance the benefits of our innate sociability? Not surprisingly, research in political economy (and evolutionary biology) suggests that both are the case, and that “society” is the evolving, emergent outcome of the interplay of our genes and our institutions. As Matt Ridley observed,

Society works not because we have consciously invented it, but because it is an ancient product of our evolved predispositions. It is literally in our nature. (1996, p. 5)

Our neurological composition also plays a role in the emergent order known as society. What do the neurological foundations of “man as a social animal” contribute to our analysis and understanding of social institutions? Consider the following stylized narrative of economic history, the formation and evolution of institutions, and the move from personal exchange to impersonal anonymous exchange.¹ In early human history, people lived together in families and small communities, in which they interacted personally and repeatedly. In such societies, the rules governing their behavior arose by trial and error, through repeated personal interaction. Rules that generally led to better outcomes were selected and survived, and economic growth was evidence of the benefits of such rules. These rules became custom.

Over time, success in these communities led both to population growth and to the extension of interaction with other communities. Interaction and exchange became less personal and more anonymous, and the share of exchange that was impersonal increased. In these more impersonal exchanges, each party brought a concept of customary law to the transaction. Resolving differences in custom and tradition while retaining customary rules that were mutually

¹ This narrative is my high-level interpretation of extensive work in economic history, including (but not limited to) Mokyr (1990), North (1990), North (2006), Greif (1989), and Milgrom, North and Weingast (1990). Hayek (1973) also discusses the evolution of legal institutions from more of a cultural anthropological perspective.

beneficial was (and remains) part of the institutional evolution that accompanies, and further enables, increasing impersonal exchange.

Hence the surviving customary legal institutions were codified into formal law over time, with the most obvious example of this process being English common law. The codified formal institutions generally retain features of the customary traditions, but have also evolved to incorporate features that are better suited to impersonal exchange, and to facilitating impersonal exchange. Thus, for example, consumer credit transactions no longer require personal discussion and a handshake, but require only a credit card and a signature. This institutional evolution incorporates a level of abstraction into legal institutions that is not necessary in face-to-face exchange.

What is the human neurological basis for this pattern of economic history, and does that basis help us analyze and understand social institutions more deeply? This neurological basis is, in part, found in the mirror neuron system. Recent neuroscience research has explored the existence and function of the mirror neuron system in humans. The mirror neuron system is a distributed neural network that seems to be the foundation of our ability to understand the actions and intentions of others, form expectations for the behavior of others, and develop sympathy and shared emotional connections with others. Here I argue that the mirror neuron system therefore provides a neurological foundation for the origin and evolution of customary institutions in face-to-face, personal exchange. That observation begs the question, though, of whether the mirror neuron system plays a role in the evolution of institutions for impersonal exchange that do not involve face-to-face-interaction. In this paper I argue that the mirror neuron system's hypothesized role in language acquisition means that it plays a role in the human capability for abstraction and conceptual knowledge. That capability is essential to the kind of generalization from individual experience that is important for developing institutions that support and facilitate impersonal exchange.

This exploratory paper sets the stage for further research (Kiesling & Spitzer 2010, Kiesling & Horwitz 2010) by focusing on three topics:

- An overview and literature survey of the interdisciplinary mirror neuron literature in cognitive neuroscience, cognitive psychology, and philosophy;

- An application of this recent research to Adam Smith's (1759) argument for sympathy as the foundation of social institutions; and
- An application of this recent research to selected topics in new institutional economics to assess the relevance of mirror neuron research to testing further hypotheses in new institutional economics.

The mirror system neuroscience research is in its early stages, which prevents substantial definitive conclusions because neuroscientists continue to learn more about what the mirror system does and does not do with respect to human social cognition. The clearest connections between the mirror system and social institutions at this point are, in order of decreasing generality:

- The mirror system in humans enables interdependence among individuals, which shapes the social interactions between (and among) people.
- The mirror system's apparent role in social cognition contributes to the formation and evolution of customary institutions among people who engage in face-to-face, personal exchange.
- The mirror system's apparent role in language acquisition, the social role of language, and the degree of abstraction and conceptual knowledge that such a capability enables are important contributors to the ability to formulate and understand abstract concepts. Such an understanding is essential for the evolution of institutions that support and facilitate impersonal exchange out of customary rules based on personal interaction.

Section II of this analysis provides a literature survey in the neuroscience research in mirror neuron systems. In Section III I discuss the general connection between this research and Adam Smith's contention that sympathy is the foundation of the institutions of civil society, and in Section IV I connect the neuroscience research to some existing and potential research areas in new institutional economics (NIE).

II. Recent neuroscience research on mirror systems

A. What is the mirror system?

Recent neuroscience research has revealed that existence of mirror neuron systems in the brains of monkeys and humans. The mirror neuron system is active both when we engage in a motor action (such as grasping an object) and when we observe someone engaging in a motor action. The pioneering research on monkeys identified mirror neurons firing when another monkey or a human performed an action (Rizzolatti et. al. (1996), Gallese et. al. (1996), Umiltà et. al. (2001)), and even upon hearing, but not seeing, an action (Kohler et. al. (2002), Keysers et. al. (2003)). Other research provides evidence that mirror neuron systems also code, or act as a neurological determiner of, facial expressions (Ferrari et. al. (2003), Carr et. al. (2003)). Corroborating evidence on the presence of mirror systems in humans has resulted from functional magnetic resonance imaging (fMRI) studies (e.g., Decety et. al. (1997), Iacoboni et. al. (1999)). A recent study using neurological data from implanted electrodes in epilepsy patients provides further corroboration of mirror system presence and activity in humans, concluding that “these findings suggest the existence of multiple systems in the human brain endowed with neural mirroring mechanisms for flexible integration and differentiation of the perceptual and motor aspects of actions performed by self and others” (Mukamel et. al. (2010), p. 755).

The mirror system is a distributed, complex neural network, located in several regions of the brain and differentially active depending on the nature of the action undertaken or observed.² Although networked across different areas, *mirror neurons* are concentrated in the premotor cortex, an area in the frontal lobe that uses sensory information to plan, choose, and implement motor action. Most of its activity is not conscious, and occurs without our having a sense of developing abilities to perform actions effortlessly. As part of the larger sensory-motor system, mirror neurons communicate information gathered either visually or aurally:

² For a thorough survey of the neuroscience literature on the mirror system, see Rizzolatti and Craighero (2004). For a general overview as well as a critical analysis of mirror neuron research and hypotheses addressing action understanding specifically, see Hickok (2009), who explores alternate hypotheses for mirroring. One alternate hypothesis, for example, is that these patterns arise from learned associations at the neural level, rather than from a distinct, distributed mirror neuron network.

Each time an individual sees an action done by another individual, neurons that represent that action are activated in the observer's premotor cortex. This automatically induced, motor representation of the observed action corresponds to that which is spontaneously generated during active action and whose outcome is known to the acting individual. Thus, the mirror system transforms visual information into knowledge ... (Rizzolatti & Craighero 2004, p. 172)

Neuroscience research thus suggests that mirror neurons provide the foundation for understanding the motor actions of others, and are therefore important factors in social interaction and social cognition.³

The premotor cortex also has large concentrations of *canonical neurons* that complement mirror neurons by firing upon observation of objects used in actions; canonical neurons fire upon seeing a bicycle, for example, even if the observer is seated and only imagining riding a bicycle. To illustrate the complementarity of mirror neurons and canonical neurons, consider the general example of the simple motor task of grasping a cup. Mirror neurons fire when an individual grasps a cup or observes someone else grasping a cup, while canonical neurons fire when an individual grasps a cup or simply observes a cup. For this reason, canonical neurons are more associated with the ability to interpret context and potential, and recent research suggests that they work with mirror neurons to create understanding, although the mirror neurons appear to be responsible for the ability to infer the intentions of others (Iacoboni et. al. 2005). For the remainder of this analysis I will refer to this combined system of mirror neurons and canonical neurons as the *extended mirror neuron system*.

Gallese and Lakoff (2005) characterize this extended mirror system as part of the "shared neural substrate" that facilitates both performing an action and imagining an action. This sharing, they argue, connects our sensory-motor action understanding to more abstract conceptual knowledge:

³ Mirror neurons can be further categorized as *strictly congruent* and *broadly congruent* based on differences in their activities: "For instance, although the term "mirror" implies a strong similarity between the executed and the observed actions, only one third of mirror neurons—the so-called strictly congruent mirror neurons—fire for the same executed and observed action. The remaining two-thirds of mirror neurons—the so-called broadly congruent mirror neurons—fire for executed and observed actions that are not the same but either achieve the same goal or are logically related ..., thus forming some sort of sequence of acts, as for instance observed placing food on the table and executed grasping food and bringing it to the mouth." (Iacoboni 2009, p. 660) In other words, strictly congruent mirror neurons fire upon execution and observation of the same action, while broadly congruent mirror neurons fire upon execution and observation of actions with the same goal, even if the actual actions are not precisely the same.

One can reason about grasping without grasping; yet one may still use the same neural substrate in the sensory-motor system. Indeed, that is just what we shall argue. In doing so, we will extend what we know about doing and imagining sharing a common substrate via the following hypothesis: *The same neural substrate used in imagining is used in understanding*. ... Accordingly, we will argue that a key aspect of human cognition is *neural exploitation*—the adaptation of sensory-motor brain mechanisms to serve new roles in reason and language, while retaining their original functions as well. (Gallese & Lakoff 2005, p. 456)

Thus they argue for a “neuro-scientific account of conceptual content as nested in the activity of the brain ...“ (p. 457), based on the idea that doing, perceiving, and imagining are all activities that are embodied in the interaction between the brain and the physical world. Under this argument, not just an action itself, but the concept of an action, is embodied in the brain’s sensory-motor circuitry. This hypothesis implies that action, action understanding, and abstract conceptual knowledge are neurologically connected through the extended mirror system. Gallese & Lakoff also characterize mirror systems as multimodal, meaning that they respond to multiple triggers (such as seeing and doing), which contribute to the sophistication and nuance of the functionality of the mirror system and related neurological systems. When combined with the social interaction aspects of the mirror system, this hypothesis is consistent with some of the mirror system’s capabilities most closely associated with social cognition and learning, and are therefore likely to have implications for human social institutions.⁴

B. Functionality of the extended mirror system

A pervasive and straightforward function of the mirror system in humans is action understanding, or the ability to understand or interpret the motor actions of others. The capability for action understanding is what creates meaning, and is a neurological foundation for abstraction and the formation of concepts that generalize beyond a specific, personal situation (Iacoboni et. al. (2005), Rizzolatti & Craighero (2004), Rizzolatti & Singalia (2010), Hickok (2009), Gallese (2003)).

Another function of the mirror system is imitation of a motor action, such as grasping, upon physical observation of another party performing that action. This sensory-motor capability also

⁴ Gallese extends this argument and model in Gallese (2008).

thus enables not just direct imitation from observation, but also emulation and learning (Iacoboni (2009)).⁵

This capability also extends to enabling the understanding of intentions and formation of expectations, both for our own behavior and for the behavior of others. Because the extended mirror system builds in an inherent external focus, we can recognize, understand, interpret, and form expectations and inferences about the actions of others in addition to ourselves. The external capability combines with our internal focus on our own behavior to enable the coordination of behavior and expectations across individuals. Gallese notes that

Social interaction, interpersonal coordination and joint actions require the capacity to understand and anticipate partners' behavior. The MNS [mirror neuron system] seems to provide the right neurofunctional substrate to all these capacities. ... The evidence collected so far suggests that the MNS is sophisticated enough to enable its exploitation for different social purposes. ... (Gallese 2008, p. 320)

This ability to recognize and understand others' actions thus contributes to the predictability of behavior in social contexts.

Mirror systems research has also focused on the role that such neurological capacities can play in the formation of sympathy and empathy – shared understanding of feelings and emotions (Decety et. al. (1997), Iacoboni et. al. (1999), Iacoboni (2009)). Thus the research suggests that the extended mirror system provides neural mechanisms for social relationships with other people. Furthermore, an extensive literature in the philosophy of mind explores the implications of mirroring for “mind reading” (e.g., Hurley (2005), Borg (2007)), in part building on this potential for empathy between the self and other.

Within the extended mirror system, canonical neurons in particular contribute to an individual's imaginative capacity. Canonical neurons fire in ways that show processing and understanding of intentions, not just actions, based on seeing objects. Intention understanding employs a capacity for imagining potential actions one could take with an object.

⁵ Among primates, only humans and apes are capable of motor imitation, suggesting that imitation cannot be the *primary* function of mirror neurons generally (Rizzolatti & Craighero 2004, p. 172). However, in humans the mirror system does seem to support imitation at both a motor and a cognitive level.

The combination of the functionality of sensory motor imitation, learning, understanding of intention and potential, expectations formation, sympathy, and embodiment of abstract conceptual knowledge has led to extensive research over the past decade on the implications of mirroring for language acquisition (Gallese (2008), Gallese & Lakoff (2005), Corballis (2010)).

With the advent of language, and even more with the “discovery” of written language, meaning is amplified as it frees itself from being dependent upon specific instantiations of actual experience. Language connects all possible actions within a network expanding the meaning of individual situated experiences. ... When we speak, by means of the shared neural networks activated by embodied simulation, we experience the presence of others in ourselves and of ourselves in others. (Gallese 2008, p. 322)

Thus the mirror system contributes to the development of abstract meaning, with language as an abstract system for the representation of meaning.

C. Summary of current mirror system hypotheses

While new, and very much under development, the neuroscience research into the extended mirror system over the past 15 years has generated support for some hypotheses that are relevant to the formation and evolution of human social institutions. The hypotheses of material relevance in this analysis are:

- (H1) The (extended) mirror system is a multimodal, shared neurological system/substrate comprising mirror neurons (strictly congruent and broadly congruent) and canonical neurons.
- (H2) The (extended) mirror system contributes to our individual capabilities to imitate, learn, recognize, understand, and infer in observation and interaction with others.
- (H3) Canonical neurons specifically contribute to inference and understanding of intent and the capacity for imagination
- (H4) The (extended) mirror system provides a neurological foundation for language acquisition, and for the social nature of language as an embodied sensory-motor capability that enables abstraction from a particular physical circumstance or situation to more general concepts and principles.

These neurological foundations of our social interactions may also have implications for the development and evolution of economic and social institutions, and for the form that they take in different contexts.

III. Mirror system neuroscience and Adam Smith's concepts of sympathy⁶

A. Smith's sympathy, mutual sympathy, and institutions

One avenue for connecting mirror systems neuroscience research to new institutional economics is Adam Smith's development of the concepts of sympathy and mutual sympathy in *The Theory of Moral Sentiments* (1759).⁷ Smith's work of moral philosophy is in part an exploration of the principles and institutions that enable individuals to live together in civil society, and is thus focused directly on the factors that affect the formation, evolution, and form of economic and social institutions.

Sympathy is the ability to imagine the situation in which another person stands, to evaluate that situation and that person's actions and feelings. Smith defines sympathy as "... our fellow-feeling with any passion whatever." (1759, I.I.5) Passions with which we sympathize include joy, happiness, gratitude, grief, sorrow, and misery. In general, throughout *The Theory of Moral Sentiments* Smith calls sympathy a "fellow-feeling".⁸ The modern use of the term has narrowed, and is now more focused specifically on compassion or commiseration.

Smith characterized our desire for sympathy as an innate part of human nature; in fact, he opens *The Theory of Moral Sentiments* with this assertion: "How selfish soever man may be supposed, there are evidently some principles in his nature, which interest him in the fortune of

⁶ This discussion is condensed from Kiesling and Spitzer (2010).

⁷ All paragraph references herein refer to the online version of the text, available at <http://www.econlib.org/library/Smith/smMS.html>.

⁸ This definition is consistent with usage at the time, and with one of the many nuanced meanings of the word, according to the *Oxford English Dictionary*: "The quality or state of being affected by the condition of another with a feeling similar or corresponding to that of the other; the fact or capacity of entering into or sharing the feelings of another or others; fellow-feeling. Also, a feeling or frame of mind evoked by and responsive to some external influence."

others, and render their happiness necessary to him, though he derives nothing from it except the pleasure of seeing it.” (1759, I.I.1). He also conceptualized sympathy as a complex synthesis of our passions and our reason that contributed to the formation of our social relationships and institutions; as such, for Adam Smith sympathy is an instrument of ecological rationality (V. Smith (2008), Marciano (2009)). In contrast to the Cartesian deductive rationality that is assumed in many economic models, ecological rationality reflects the empiricism, inductive learning, and knowledge gathering that characterizes the context of individual decision-making more realistically. Thus by its nature ecological rationality is a bounded rationality concept, as is reflected in how Smith argues that sympathy functions for individuals and societies.

Sympathy bridges the gap among one’s internally-experienced sentiments, one’s externally-expressed passions, and one’s judgment of those passions and associated actions, grounded in reason. The form that sympathy takes relies crucially on analogy, and on the inescapable fact that we cannot fully experience the passions, perceptions, and realities of others:

Whatever is the passion which arises from any object in the person principally concerned, an analogous emotion springs up, at the thought of his situation, in the breast of every attentive spectator. ... In every passion of which the mind of man is susceptible, the emotions of the by-stander always correspond to what, by bringing the case home to himself, he imagines should be the sentiments of the sufferer. (1759, I.I.4)

We are incapable of experiencing fully and accurately the passions of others – there will always be a gap between the self and others – but we can analogize based on how we imagine we would feel, and should react, if in the same situation.⁹ Note also that this analogizing is a heuristic that enables sociability and cooperation in the presence of bounded ecological rationality.

Due to an individual’s inability to know fully another person’s true internal sentiments, sympathy is by necessity based on imagination, and Smith relies extensively on the role of imagination in our ability to form sympathy and mutual sympathy. This use of imagination is morbidly obvious in Smith’s discussion of how we even sympathize with the dead:

⁹ A detailed discussion of the “impartial spectator” device in Smith is beyond the scope of this analysis, but the capacity for understanding, inference, and abstraction that the mirror system enables is consistent with the functions he assigns to the impartial spectator.

The idea of that dreary and endless melancholy, which the fancy naturally ascribes to their condition, arises altogether from our joining to the change which has been produced upon them, our own consciousness of that change, from our putting ourselves in their situation, and from our lodging, if I may be allowed to say so, our own living souls in their inanimated bodies, and thence conceiving what would be our emotions in this case. It is from this very illusion of the imagination, that the foresight of our own dissolution is so terrible to us, and that the idea of those circumstances, which undoubtedly can give us no pain when we are dead, makes us miserable while we are alive. (1759, I.I.13)

Smith's concept of sympathy is grounded in both passions and reason, and it is from reason that individuals judge the passions of others and of themselves, based on an individual evaluation of the passion's or action's propriety. From that position of propriety, embodied in the abstract concept of the impartial spectator, Smith argues that different passions and actions merit different degrees of sympathy from others (1759, I.I.6-I.I.9). Observers cannot access the bodily passions or internal sentiments of individuals, but can imagine related reactions and feelings. Pain is an example of this asymmetry; I cannot experience your pain, but I can imagine how it feels, and I can experience an associated fear of pain (Hermann-Pillath (2009)). An "unsocial" passion such as anger deserves little sympathy, unless the passion is justified because of it resulting from some unjust situation. Smith observes that "social" passions, such as compassion, merit full sympathy from an observer. The passions he characterizes as "selfish", such as joy or grief, fall in between the unsocial and the social passions, and generally merit sympathy as long as the passion is in balance and accord with the action or situation that brought it about (as judged by the observer's standard of propriety): "Sympathy, therefore, does not arise so much from the view of the passion, as from that of the situation which excites it." (1759, I.I.10) Note that this evaluation of the "situation which excites it" requires a degree of abstract, conceptual reasoning that extends beyond the individual's personal, physical experience in the moment.

Not only do individuals have an innate desire for sympathy and an imaginative capacity that enables them to sympathize with others – we also seek out and benefit from mutual sympathy, according to Smith. Mutual sympathy arises when emotion, passions, or sentiments are coordinated across people, when their passions are in agreement in substance and in magnitude. Each person derives pleasure from seeing his/her sentiments echoed, whether positive or negative, weak or strong. Indeed, Smith contends that the desire for this pleasure is so strong that people will modulate their passions to achieve this concordance, and that

modulation process is informed by observation, imagination, judgment in the form of the impartial spectator, and a concept of propriety in that situation. Thus striving for the great pleasure of mutual sympathy is a discovery and coordination process that, even more than sympathy, relies on an ability to bridge between internal and external, and between passions and reason. This drive for mutual sympathy is the foundation of social cohesion and community, implemented by repeated reciprocal adjustments of behavior and sentiments over time.¹⁰

Sympathy and mutual sympathy are also partly a function of social context, and changes according to social distance (1759, Part VI). Smith contended that sympathy and mutual sympathy were strongest and most likely among family members, and then among friends, and then acquaintances and neighbors, with potential sympathy and mutual sympathy decreasing as social distance increases and relationships become more impersonal. If Smith is correct and sympathy is the core human foundation of social institutions, the relationship between sympathy and social distance has implications for the form of social institutions in different contexts, and for the heterogeneity and diversity of institutions across those contexts.

For Smith, social institutions such as justice and beneficence are grounded in these notions of sympathy and mutual sympathy (1759, Part II). An individual's adoption of and adherence to rules, whether formal or informal, is driven by the innate desire for mutual sympathy. The process of iterated, reciprocal adjustment of passions and actions to enable that mutual sympathy reinforces rules, typically by trial and error. These rules become custom, and over time can be codified into formal rules of (commutative) justice protecting individual rights to life, liberty, and property.

Smith's model of sympathy and mutual sympathy as the foundation for social institutions relies on an assertion that sympathy and mutual sympathy are innate human desires. Thus, to the extent that a connection exists between Smithian sympathy and the mirroring function of the extended mirror system, this system provides a deep neurological foundation for institutions, and for institutional differences, by providing support for Smith's assertion.

¹⁰ I am indebted to Otteson (2002) for contributing to this interpretation and analysis.

B. Mirroring and Smith's concepts of sympathy and mutual sympathy

The mirror systems hypotheses explored in neuroscience are strongly consistent with Smith's concepts of sympathy and mutual sympathy. The mirror system relies on observation and enables action understanding, imitation, emulation, and learning, even in situations where the observer is not engaging in the action; Smithian sympathy relies on the observation of the passions and actions of others, even in situations where the observer herself is not experiencing that passion or engaging in that action. In both instances the observation need not be visual, but the effects are influenced by distance, either physical or social depending on the context; thus both mirroring and Smithian sympathy are partly informed by social context. The mirror system bridges our internal physical and emotional experiences and the external world, and that bridging between internal and external is necessary for individuals to develop and experience Smithian sympathy and mutual sympathy.

The extended mirror system is also likely to play a crucial role in our ability to understand and evaluate situations that we observe. To the extent that it contributes to our individual capabilities to imitate, learn, recognize, understand, and infer in observation and interaction with others, the mirror system is consistent with the component of Smithian sympathy that evaluates and judges the passions and actions of others. This function of sympathy synthesizes the passions and reason and relies on our evaluative abilities that have foundations in the mirror system – an individual cannot form sympathy without the ability to recognize, understand, and infer the actions and reactions of others.

One of the most striking connections between the extended mirror system and Smithian sympathy is the crucial role of imaginative capacity in the formation of sympathy and mutual sympathy. The contribution of canonical neurons to our inference and understanding of intent and capacity for imagination provides a neurological foundation for the imaginative capacity that is essential for enabling an individual to bridge from her personal experiences, sentiments, and passions to any sort of “fellow-feeling” with the observed passions and actions of others.

IV. Institutional implications and hypotheses of the extended mirror system

The connection between the mirror system and institutions seen in the analysis of Smithian sympathy generalizes to other situations in which institutions evolve as interactions become more impersonal, or where outcomes vary depending on the nature of the interaction. In this brief analysis I focus on how an understanding of the mirror system is relevant to two examples of NIE-related research:

- Human nature, sociability, the capacity for abstraction, and institutions; and
- Social distance, mirroring, and institutional form.

This (selective) overview suggests that the mirror system provides a plausible hypothesis for explaining many of the institutional patterns and phenomena we have seen in human history.

A. Human nature, sociability, the capacity for abstraction, and institutions

In writing about the social evolution theories of Darwin and Hayek, Marciano (2009) discusses both as working within the framework of ecological rationality developed by the Scottish Enlightenment scholars, including Smith and Hume. As seen in the previous section, the mirror system contributes some understanding and support to the ecological rationality framework and its institutional implications. We can trace those implications from Smith to current research in emergent legal orders. In brief, this area of analysis supports the stylized narrative in the introduction about the co-evolution of legal institutions and impersonal exchange.

Smith (and Hume, Darwin, and Hayek) starts from the presumption that humans are social animals who have always lived in a social state, in contrast to social contract theorists like Hobbes. As Marciano observes:

... the assumption of a non-constructivist rationalism, necessarily implies that civil societies were not created by a social contract, or to put it differently, the idea that human beings could have given birth or created, instrumentally and purposefully, their societies is incompatible with the theoretical framework Hume and Smith adopt. In effect, individuals are not able to envisage and build institutions that they have not experienced, so institutions actually exist because they have always existed. In effect, in contrast to what individualist theorists (in particular social contract theorists like Hobbes) have argued, human beings never lived isolated and separated from others; they did not rationally choose to live with others in social groups, but have always lived in social states. If a state

of nature ever existed, it never resembled to what Hobbes describes. (2009, p. 55)

The ecological rationality framework implies that this human sociability is innate. The enabling role of the mirror system is one of providing the neurological bases for that sociability – the understanding of actions, goals, and intentions; the imaginative capacity; and the capacity for abstraction that, in Smith’s model, grounds institutions in mutual sympathy. Darwin (DATE) and Hayek (1977, 1988) characterize innate human sociability as “instinctive”: “Instincts thus form the “cement” of primary and elementary social groups, first of all because of their biological nature. They exist before any other faculty and, accordingly, make the development of mental faculties possible.” (Marciano 2009, p. 55) Whether articulated as human nature or instinct, the neurological capabilities of the mirror system are consistent with an innate human capacity for sociability, and thus support the ecological rationality framework.

Furthermore, this first step in the analysis extends our understanding of the evolution of legal institutions. Hayek (1973) argues that law and formal legislation are not the same, and that law is an emergent outcome of the process of (innately sociable) individuals in a community coordinating their behavior to create a peaceful and prosperous social existence.

In communities living in close personal proximity with repeated face-to-face interaction, the combination of action understanding from the mirror system, reputation incentives from repeated interactions, and Smith’s argument for the desire for mutual sympathy combine to form rules of conduct that emerge as common codes of behavior for the community.¹¹ More generally, Hayek notes that

... individuals had learned to observe (and enforce) rules of conduct long before such rules could be expressed in words; and second, that these rules had evolved because they led to the formation of an order of the activities of the group as a whole ... the efficiency of the resulting order of actions will determine whether groups whose members observe certain rules of conduct will prevail. (1973, p. 74)

Thus, again, understanding the mirror system provides some insights that complement a body of existing work on emergent legal institutions.

¹¹ Note that this argument also supports those in Greif (1989) on the self-enforcing contracts employed by the Maghribi traders in the medieval Mediterranean, as well as Greif (1998) more generally on the methodology of historical and comparative institutional analysis.

A next step in the analysis is examining the extent to which the mirror system increases understanding of the evolution of institutions supporting impersonal exchange out of these customary, personal institutions. At its most basic level, the mirror system is a sensory-motor network that primarily works on visual and auditory information. How can such a physical system enhance the understanding of something as abstract as formal legal institutions that enable exchange and civil society among strangers?

Take, for example, the accounts of the Law Merchant in Benson (DATE) and Milgrom, North, and Weingast (1990). The Law Merchant was a system of commercial law, developed and adjudicated by merchants at trade fairs, to settle disputes quickly between merchants who were typically strangers from a wide range of countries and cultures. It evolved from their separate practices and traditions, emerging from the practice of having an independent merchant adjudicate such disputes.

In contributing to our understanding of examples like the Law Merchant, the mirror system plays two plausible roles. First, as discussed above and in the previous section, the mirror system's functions provide neurological mechanisms for coordination on a set of customary rules within a personal community. Second, the role of the mirror system in language acquisition enables the capacity for abstraction that can carry those customs into more impersonal situations (in which some bargaining and negotiation would certainly be necessary, though). The iterative interactions that fed into the Law Merchant codified a set of rules at the intersection of many different customary systems, and our neurological capacity for abstraction contributes to making that evolution possible, because that capacity allows individuals to generalize from personal experience.¹²

B. Social distance, mirroring, and institutional form

The neuroscience research and the discussion of Smith's grounding of social institutions in sympathy suggest that social distance is an important variable in the formation of institutions and in the functions that institutions serve within the broader context of providing a context for

¹² Hayek (1973, Chapter 4) has an extensive discussion of law before language, and the role of language in coordinating the actions, codes of conduct, and expectations of community members, ultimately leading to codification into formal law.

cooperation and civil society. Perhaps more importantly, these ideas indicate that as social distance changes, institutional forms do, and should, change accordingly. I apply this observation briefly to the experimental literature on social distance and social cues in dictator game experiments. For many of the functions tested in the mirror system research, physical proximity or the lack thereof changes outcomes and affects whether or not the mirror system engages. Generally, visual or auditory cues are important for engagement. This finding provides neurological support for the results of such research.

Dictator game experiments are used to test hypotheses about voluntary giving to charity or public goods, the effect of trust on giving, and the effect of social distance and social cues on giving. In the dictator game, two players (Dictator and Recipient) split an endowment given to D; D determines the amount to give to R, and R cannot respond (i.e., R has no veto power). Game theory models predict that D would keep the entire endowment, a result that is generally not borne out in laboratory experiments. The social distance between D and R affects D's choice – more anonymity increases the amount D keeps, although in most experiments D does give some of the money to R. Face-to-face interaction generally increases payments to R – the lower the social distance and stronger the social cues, the larger the payments to R.

An extensive literature exists on the differences in outcomes in dictator games depending on social cues and social distance; see, for example, Camerer (2003), List (2007), and Rigdon (2003) for thorough overviews and surveys of the dictator game literature. Recently, Rigdon et al. (2009) performed a dictator game experiment in which even a weak social cue of the dictator being observed led the dictator to offer higher average payments to the recipients.

Mirror systems research on empathy is consistent with this pattern of results. Iacoboni (2009) surveys the cognitive neuroscience and cognitive psychology literature on imitation and empathy; he finds that humans tend to echo the behavior of others in personal interactions, feel more empathy toward those with whom they share echoed behavior, and tend to have less mirror system response, less imitation, and less empathy as visual or aural connections are decreased. Iacoboni identifies the mirror system as the network that connects imitation, broadly defined, and empathy:

Mirror neurons, in contrast, provide a prereflective, automatic mechanism of mirroring what is going on in the brain of other people that seems more compatible with our ability

to understand others effortlessly and with our tendency to imitate others automatically, as we have discussed in this review. A further implication of the recent work on the relationships between mirror neurons, imitation, and empathy is the consideration that the evolutionary process made us wired for empathy. This is a major revision of widely held beliefs. Traditionally, our biology is considered the basis of self-serving individualism, whereas our ideas and our social codes enable us to rise above our neurobiological makeup. The research on mirror neurons, imitation, and empathy, in contrast, tells us that our ability to empathize, a building block of our sociality (Adolphs 2009) and morality (de Waal 2008, Tangney et al. 2007), has been built “bottom up” from relatively simple mechanisms of action production and perception (Iacoboni 2008). (2009, p. 666-667)

He concludes that the mirror system, as a neural mechanism for interdependence and connecting people, plays an important role in our ability to empathize and cooperate with others. Not only is this conclusion consistent with the dictator game literature, but it is also consistent with the earlier analysis of the role of the mirror system in coordinating expectations and behavior.

V. Conclusion

By surveying the results of mirror neuron research, connecting mirror systems and Smithian sympathy, and indicating the mirror system’s role in institutional diversity, this analysis suggests that the human mirror neuron system may have an explanatory role in the formation and evolution of institutions, and the different forms they take in different contexts. The mirroring function contributes to our desire for sympathy and mutual sympathy, enabling our coordination on customary rules in personal communities. However, its ability to facilitate abstract understanding and language acquisition also suggests that the mirror system enables the evolution of more abstract institutions during the move from personal to impersonal interaction, thereby facilitating even further such interaction and the benefits from it. The mirroring and social distance hypothesis also suggests the existence of a personal, neurological foundation for the benefits of institutional diversity and heterogeneity, and for nested institutions more generally.

The connections from the mirror system to institutions through social distance and learning raise some important institutional change questions, two of which I will address and both of which provide avenues for further research. First, if the processes of learning and language acquisition rely on visual and aural cues and are grounded in sensory motor systems, then how do they

connect to something as abstract as social institutions? Language as an abstract representational framework is a plausible hypothesis that requires further examination.

The connection among mirroring, social distance, and institutional design is broader than the dictator game suggests. The mirror system research reinforces and provides a neurological foundation for the general results from Ostrom (1990, 2005) that our institutional structures in complex social systems are typically nested, and that communities with more face-to-face interaction are generally better able to coordinate and develop decentralized governance institutions.

Those results, though, probably conflate three mutually reinforcing dimensions of interaction – the face-to-face, or social distance, aspect of the interaction, the size of the group, and the repeated nature of the interaction. Closer social distance, smaller group size, and repetition of interaction are three separate hypotheses for the ability to coordinate actions, and should be examined separately. Further experimental research in this area could disentangle these hypotheses.

Second, note that the neuroscience research suggests that the mirror system plays a role not only in learning and language acquisition, but also in adaptation. However, the neuroscience research emphasizes the role of the mirror system in enabling us to coordinate our actions and decisions, and consequently in the formation of expectations. This focal-point coordination suggests a certain degree of stasis, while adaptation, and by extension institutional change, are dynamic processes of change. Does the mirror system contribute to maintaining a status quo, to a degree of inertia in processes of change, because of this focal point coordination? Both consequences are probably true, but a systematic evaluation of the role of mirroring in institutional change in different contexts may enhance our understanding. One challenge for experimental design is coming up with an appropriate alternative treatment – what is the counterfactual to the mirror system hypothesis?

Put another way, what is the balance between coordination around a stable outcome and coordination around a dynamic evolutionary process? What are the processes by which adaptation, including institutional adaptation, occurs? This general question has been important for all studies of institutional evolution and change, including (but not limited to) Hayek (1979,

1988), Ostrom (1990, 2005), and North (1990, 2006). Mirror system research may yield some insights into institutional inertia in the presence of dynamic change, and into the processes of institutional change themselves.

References

Benson, B. (1989) *The Enterprise of Law*.

Borg (2007)

Camerer C. (2003) *Behavioral Game Theory*. Princeton: Princeton University Press.

Carr L, Iacoboni M, Dubeau MC, Mazziotta JC, Lenzi GL. (2003) "Neural mechanisms of empathy in humans: a relay from neural systems for imitation to limbic areas." *Proc. Natl. Acad. Sci. USA* 100: 5497–502.

Darwin, C. (1871) *The Descent of Man*.

Decety, J., Grezes, J., Costes, N., Perani, D., Jeannerod, M., Procyk, E., Grassi, F. & Fazio, F. (1997) *Brain* 120: 1763–1777

Ferrari PF, Gallese V, Rizzolatti G, Fogassi L. (2003) "Mirror neurons responding to the observation of ingestive and communicative mouth actions in the monkey ventral premotor cortex." *Eur. J. Neurosci.* 17:1703–14.

Gallese, V. (2008) "Mirror neurons and the social nature of language: The neural exploitation hypothesis." *Social Neuroscience* 3-4: 317-333.

Gallese, V, Lakoff, G. (2005) "The brain's concepts: The role of the sensory-motor system in conceptual knowledge." *Cognitive Neuropsychology* 22(3/4): 455-479.

Gallese V, Fadiga L, Fogassi L, Rizzolatti G. (1996) "Action recognition in the premotor cortex." *Brain* 119: 593–609.

Greif, A. (1998) "Historical and Comparative Institutional Analysis." *American Economic Review* 88(2): 8-84.

Greif, A. (1989) "Reputation and Coalitions in Medieval Trade: Evidence on the Maghribi Traders." *Journal of Economic History* 49(4): 857-882.

Hayek, F. (1988) *The Fatal Conceit: The Errors of Socialism*. Chicago: University of Chicago Press.

Hayek, F. (1977) *Law, Legislation, and Liberty, Volume 2: The Mirage of Social Justice*. Chicago: University of Chicago Press.

Hayek, F. (1973) *Law, Legislation, and Liberty, Volume 1: Rules and Order*. Chicago: University of Chicago Press.

Hermann-Pillath, C. (2009) "Adam Smith's *Theory of Moral Sentiments* in the light of modern neuroscience and neuroeconomics." Presentation at the workshop, *Limits to Rationality in Financial Markets*, The Institute for Advanced Studies, Glasgow.

- Hickok, G. (2009) "Eight Problems for the Mirror Neuron Theory of Action Understanding in Monkeys and Humans." *Journal of Cognitive Neuroscience* 21(7): 1229-1243.
- Hurley (2005) "Social heuristics that make us smarter." *Philosophical Psychology* 18(5): 585-612.
- Iacoboni M. (2009) "Imitation, empathy, and mirror neurons." *Annual Review of Psychology* 60: 653-670.
- Iacoboni M, Molnar-Szakacs I, Gallese V, Buccino G, Mazziotta JC, et al. (2005) "Grasping the intentions of others with one's own mirror neuron system." *PLoS Biology* 3(3): e79.
- Iacoboni M., Woods, R. P., Brass, M., Bekkering, H., Mazziotta, J. C. & Rizzolatti, G. (1999) *Science* 286: 2526–2528.
- Keysers C, Kohler E, Umiltà MA, Nanetti L, Fogassi L, et al. (2003) "Audiovisual mirror neurons and action recognition." *Exp Brain Res* 153: 628–636.
- Kiesling L and Horwitz S. (2010) "Mirroring, Classification, and Social Institutions: Implications of Recent Neuroscience Research for Interpreting Hayek's The Sensory Order." Working paper.
- Kiesling L and Spitzer Y. (2010) "The Implications of Mirror Neuron Systems for Interpreting Sympathy and Mutual Sympathy in Adam Smith and Dugald Stewart." Working paper.
- Kohler E, Keysers C, Umiltà MA, Fogassi L, Gallese V, et al. (2002) "Hearing sounds, understanding actions: Action representation in mirror neurons." *Science* 297: 846–848.
- List J. (2007) "On the interpretation of giving in dictator games." *Journal of Political Economy* 115(3): 482-493.
- Mokyr J. (1990) *The Lever of Riches*. Oxford: Oxford University Press.
- Milgrom P, North, D, & Weingast, B. (1990) "The Role of Institutions in the Revival of Trade: The Law Merchant, Private Judges, and the Champagne Fairs." *Economics and Politics*.
- Mukamel R, Ekstrom A, Kaplan J, Iacoboni M, & Fried I. (2010) "Single-Neuron Responses in Humans during Executions and Observations of Actions." *Current Biology* 20: 750-756.
- North D. (2006) *Understanding the Process of Economic Growth*. Princeton: Princeton University Press.
- North, D. (1990) *Institutions, Institutional Change, and Economic Performance*. Cambridge: Cambridge University Press.
- Ostrom, E. (2005) *Understanding Institutional Diversity*. Princeton: Princeton University Press.
- Ostrom, E. (1990) *Governing the Commons*. Cambridge: Cambridge University Press.
- Otteson, J. (2002) *Adam Smith's Marketplace of Life*. Cambridge: Cambridge University Press.

- Polanyi, M. (1974) *Personal Knowledge*. Chicago: University of Chicago Press.
- Ridley, M. (1996) *The Origins of Virtue*. London: Penguin Press.
- Rigdon M, Ishii K, Watabe M, Kitayama S. (2009) "Minimal social cues in the dictator game." *Journal of Economic Psychology* 30: 358-367.
- Rigdon M. (2003). "Dictator games." In L. Nadel (Ed.), *Encyclopedia of Cognitive Science* (pp. 209–210). London: MacMillan.
- Rizzolatti G, Craighero L. (2004) "The mirror-neuron system." *Annual Review of Neuroscience* 27: 169-192.
- Rizzolatti G, Fadiga L, Gallese V, Fogassi L. (1996) "Premotor cortex and the recognition of motor actions." *Brain Res Cogn Brain Res* 3: 131–141.
- Rizzolatti G, Sinigaglia C. (2010) "The functional role of the parieto-frontal mirror circuit: interpretations and misinterpretations." *Nature Reviews Neuroscience* 11: 264-274.
- Smith, Adam. (1759) *The Theory of Moral Sentiments*. Indianapolis: Liberty Fund.
- Smith, Vernon. (2008) *Rationality in Economics*. Chicago: University of Chicago Press.
- Umiltà MA, Kohler E, Gallese V, Fogassi L, Fadiga L, et al. (2001) "I know what you are doing. A neurophysiological study." *Neuron* 31: 155–165.