

Trust as the glue of cognitive institutions

Shaun Gallagher & Enrico Petracca

To cite this article: Shaun Gallagher & Enrico Petracca (16 Oct 2022): Trust as the glue of cognitive institutions, *Philosophical Psychology*, DOI: [10.1080/09515089.2022.2134767](https://doi.org/10.1080/09515089.2022.2134767)

To link to this article: <https://doi.org/10.1080/09515089.2022.2134767>



© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



Published online: 16 Oct 2022.



Submit your article to this journal [↗](#)



Article views: 1691



View related articles [↗](#)





View Crossmark data [↗](#)



Citing articles: 1 View citing articles [↗](#)

Trust as the glue of cognitive institutions

Shaun Gallagher ^{a,b} and Enrico Petracca ^c

^aDepartment of Philosophy, University of Memphis, Memphis, TN, USA; ^bSchool of Liberal Arts, University of Wollongong (AU), Wollongong, Australia; ^cSchool of Economics, Management and Statistics, University of Bologna, Bologna, Italy

ABSTRACT

In this paper we consider the importance of trust, in the context of economic institutions, and specifically with respect to questions about market mechanisms and the role of social interactions. We review recent advances in institutional economics closely tied to developments in philosophy of mind and cognitive science, involving extended and enactive cognition. We argue that the analysis of different conceptions of institutional mind extension, in Denzau and North's shared mental models, Clark's extended mind, and a more enactive approach that emphasizes the importance of social interaction and personal relationships, can benefit from Kathrine Hawley's distinction between reliability and trust. Institutional arrangements based solely on the reliability of impersonal mechanisms can lead to a variety of social pathologies and, at the extreme, a form of cognitive atrophy, all of which can undermine the sustainability of institutions. Even if trust comes with risks and some degree of unpredictability, it turns out to be a necessary glue-like ingredient in the constitution of social institutions.

ARTICLE HISTORY

Received 27 July 2022
Accepted 5 October 2022

KEYWORDS

Extended mind; socially extended mind; cognitive institutions; markets; trust; reliability

1. Introduction

Katherine Hawley (2014) noted that philosophers often distinguish between trusting others and trusting (or simply relying on) an instrument. An instrument is trusty if it functions properly – does what it is designed to do. This is a kind of shallow trust which can be equated with “perceived reliability” (David et al., 2021, p. 55). There is no sense that the instrument may be dishonest or may try to hide something, or try to mislead me in some circumstances, or that it could betray me (Baier, 1986). In contrast, in the case of other persons we need to consider just these possibilities. Again, as Hawley points out, trust is also something that may (or may not) define our relations with institutions or collectives. In this paper we explore the notion of trust in personal and institutional relations that include relations with economic institutions, such as markets. One question is whether we

CONTACT Shaun Gallagher  s.gallagher@memphis.edu  Department of Philosophy, University of Memphis, Memphis 38152, TN, USA

© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

should think of trust in (or distrust of) such institutions in terms of a shallow instrumental reliability or in terms closer to the deeper trust we might have in other persons. After all, institutions are not just tools that we can use; they include people and are in some limited way expressions of ourselves. To be clear, the question is not, “Should we trust institutions?” – rather it is, “What is the nature of the kind of trust that is operative in regard to institutions?” This question was raised by Hawley (2014): “it is no accident that within this philosophical paradigm we rarely discuss trust or distrust as directed toward collective entities, since even the most inflationary accounts of collective agency hold back from treating such entities as full-fledged persons in every respect, on a par with individual human persons” (pp. 2–3). She also demonstrates the complexity of finding an answer given the different ways that philosophers and sociologists define groups, organizations and institutions.

We want to explore this question in relation to economic institutions. To develop an answer, we take our bearing from the concept of cognitive or “mental institution” (Gallagher & Crisafi, 2009). We can trace the development of this notion in institutional economics starting from Arthur Denzau and Douglass North’s (1994) idea of institution understood in terms of “shared mental models”. Motivated by Denzau and North’s views, Andy Clark (1997) took this one step further, introducing the notion of “scaffolding institution”, based on the concept of the extended mind. The notion of cognitive institution, or the socially extended mind (Gallagher, 2013), goes even further, drawing on 4E (embodied, embedded, extended, enactive) cognition, and attempts to capture more fundamental and salient aspects of economic institutions, such as markets (Gallagher et al., 2019; Petracca & Gallagher, 2020). After summarizing these developments, and locating ourselves on the landscape of extended and socially extended cognition, we easily find a model of trust that would treat the scaffolding institution according to the shallow instrumental version of reliability. We argue that this model is insufficient for the dynamical workings of the socially extended cognitive institution.

2. Three models of institutions

As Markey-Towler (2017) points out, “adherence, knowing or unknowing, to a particular philosophy of mind can have profound effects on the ontology of the economy” (p. 203). We think it can also have important consequences for how we understand a central concept in economics, namely the concept of trust and the related notion of trustworthiness. In philosophy of mind, externalist 4E approaches that do not restrict the ontology of mind to the internal workings within an individual’s head, but rather emphasize the constitutive role for embodied and environmental

processes (see, e.g., Newen, De Bruin, & Gallagher 2018), provide alternatives to the more or less standard internalist positions found in institutional theory (e.g., Searle, 2005). Consider three such alternatives.

2.1 *Shared mental models*

Denzau and North (1994) attempt to create something like a bridge between the internal realm of representations and the external institution. Institutions stem from individuals' "shared mental models", and institutional change takes place when these mental models change (North, 2005). "The mental models are the internal representations that individual cognitive systems create to interpret the environment; the institutions are the external (to the mind) mechanisms individuals create to structure and order the environment" (North, 2005, p. 4). More specifically, they define institutions as "the rules of the game of a society [consisting] of formal and informal constraints constructed to order interpersonal relationships." The sharing of mental models and the formation of institutions are motivated by an attempt to reduce uncertainty in human, and specifically economic relations.

Under conditions of uncertainty, individuals' interpretation of their environment will reflect their learning. Individuals with common cultural backgrounds and experiences will share reasonably convergent mental models, ideologies, and institutions; and individuals with different learning experiences (both cultural and environmental) will have different theories (models, ideologies) to interpret their environment.

(North, 2005, pp. 3–4).

2.2 *Extended mind*

Clark (1997) focuses more on the external mechanisms, and borrows significantly from Denzau and North in order to conceive of institutions as structures that provide epistemic scaffolding. As such, they involve strong constraints and incentives that predictably direct agents' behaviors. Institutions are able to produce a "cognitive economy" guiding individuals' decisions and actions. Accordingly, they significantly reduce the cognitive effort for information processing, by externalizing a number of processes. Clark (1997) suggests that the fact that economic decision-making takes place in such highly scaffolded environments explains why neoclassical economics – the economics of perfectly rational actors – works "(insofar as it works at all)" (p. 271).

This view of economic institutions is consistent with the "extended mind" hypothesis (EM) (Clark & Chalmers, 1998; Clark, 2008) – the hypothesis that cognitive processes are not limited to what happens in the head but may occur

by allowing the body and the external world to do some of the information-processing work, for example, using pencil and paper to do math, or a notebook to remember addresses. In the EM model, the emphasis falls on the use of factors external to brain and body (instruments and tools) that may become part of the overall cognitive system. This extension of the mechanisms (or vehicles) of cognition includes specific kinds of actions that manipulate an external tool or instrument in a way that is functionally equivalent to processes that occur in the head (Clark & Chalmers, 1998). Clark and Chalmers emphasized functional equivalence in what became known as the “parity principle”.

If, as we confront some task, a part of the world functions as a process which, were it done in the head, we would have no hesitation in recognizing as part of the cognitive process, then that part of the world is (so we claim) part of the cognitive process.

(1998, p. 8).

The parity principle states that information processing done via the use of external tools and instruments can be functionally equivalent to information processing done by neural mechanisms in the brain.

In what is sometimes called the “second wave” of the extended mind theory (Sutton, 2010), the notion of functional integration came to hold more importance. Second wave theorists challenge the parity principle. Even if in some cases there is a functional isomorphism between inner and outer processes, as John Sutton (2010) makes clear, there may also be many significant differences.

Sutton defends the importance of complementarity, indexed by differences in individual cognizers, and differences in particular environments. Richard Menary (2007) puts this in terms of an integration that is generated in specific kinds of cognitive or epistemic practices. One gets an integrated cognitive system, not because there is an automatic fit between neural processes and pieces of the environment, but because there is an active coupling involving reciprocal causal connections activated by active bodily manipulations of the environment. These manipulations, according to Menary (2010) include epistemic actions, sensory-motor couplings, the use of language and props to guide completion of tasks, and cognitive practices. The latter involve “the manipulation of external representational and notational systems according to certain normative practices – as in mathematics” (p. 237). Numbers, diagrams, drawings, maps, charts, etc. are external representations that allow us to accomplish cognitive tasks. We manipulate such things using pencils and paper, computers, reorganizing spatial arrangements, and we do so following norms that are culturally established and learned.

2.3 Socially extended mind

The idea of the socially extended mental (SEM) or cognitive institution has been associated with a “third wave” of extended mind (Cash, 2013; Kirchhoff, 2012). Regardless of how one defines these theoretical waves, we can say that SEM institutions return to the idea of institutions in a way that incorporates a second wave perspective that emphasizes active integration. Clark and Chalmers (1998, pp. 17–18) had anticipated the role of the social environment in suggesting that “my mental states might be partly constituted by the mental states of other thinkers . . . one’s beliefs might be embodied in one’s secretary, one’s accountant, or one’s collaborator.” The SEM institution, however, takes the idea that cognition may be a matter of shared communication, shaped by normative practices, much further. It suggests that just such practices allow for the establishment of cognitive institutions (Gallagher, 2013). In other words, extended mind is not just about the use of hand-held notebooks, writing tablets, iPhones, diagrams, maps, or specific kinds of exograms, etc. It’s also about our use of, or our engagement with large-scale institutions – academic, scientific, cultural, economic, legal, etc. These are institutions that enable cognition, and, indeed, as we engage with them, they allow for specific types of cognitive accomplishments. This to the point of altogether constituting cognitive processes, thus leading to the definition of cognitive institutions as institutions “without which some of the agents’ cognitive processes would not exist or even be possible” (Petraçca & Gallagher, 2020, p. 747).

The legal system has been a favorite example in this respect. It consists of a set of structures and practices that include normatively defined cognitive practices. When we engage with the legal system (that is, when we interact with, or are coupled to it in the right way), it extends our cognitive processes and helps us to solve problems of a particular type.

Contracts, for example, embody conceptual schemas that contribute to and shape some of our cognitive practices. They are themselves products of specific cognitive exercises, but they are also used as tools to accomplish certain aims, to reinforce certain behaviors, and to solve certain problems. Institutions of property, contract, rights, and law, and the precise way that we use them, not only constrain our thinking about social arrangements, and about acceptable behavior, they allow us to think in ways that would not be possible without such institutions.

(Gallagher 2018, p. 434; see Gallagher & Crisafi, 2009; Gallagher, 2013/2018).

Contracts not only play a role in legal contexts, but are economic instruments. Gallagher et al. (2019) and Petraçca & Gallagher (2020) have also argued that markets are cognitive institutions. A market is a social institution, and it emerges as such because it involves intersubjective interactions embedded in social and cultural practices that constrain and enable economic agents’ behavior. The constraints imposed by market interactions, as

well as the possibilities enabled by such interactions, suggest that economic reasoning is not just an individual process characterized by autonomous decision making, as it is classically understood. Considering the perspective of the agents involved, the market as cognitive institution includes Clark's notion of scaffolded cognition, (i) extending/supporting the participant's cognitive processes of economic reasoning, and (ii) both constraining and enabling the actions and interactions of embodied and embedded agents in the economy. These processes could easily be modeled on functional equivalence (first wave) and integration (second wave).

Two further characteristics of SEM-style cognitive institutions, however, add some complexity to the notion of institution. First, that institutions emerge from intersubjective interactions. They not only constrain/enable social interaction, but they are built upon (constituted by) social interactions. If one takes away social interactions, then the legal system, markets, educational institutions, science, etc., and the cognitive processes they make possible no longer exist. If this seems obvious, its importance is tied to the way that we understand the nature of intersubjective interaction. This is not something we will explore here (see Gallagher, 2020), except as it relates to the issue of trust (in the next section). Second, as clarified by Marc Slors (2019), SEM/cognitive institutions can be differentiated from an extended-mind conception of institution as a causal-functional unit. In contrast to a conception of functional equivalence or functional integration, where the cognitive process may be internal or external or some combination of internal and external information processing, the concept of a SEM institution incorporates an enactive perspective, according to which cognition is constituted in a specific form of dynamical engagement with the world (Di Paolo & Thompson, 2014; Gallagher, 2017; Kirchhoff, 2012; Stewart et al., 2010; Varela et al., 1991). Slors thus contrasts the functionalist view intrinsic to Clark's concept of institution, with what he calls "symbiotic" processes involved with SEM institutions. Slors (2019) defines the notion of symbiotic cognition in terms of "task dependency".

"Task dependency" is the extent to which the intelligibility of a task depends on a larger whole of coordinated tasks. Task dependency is a notion that is connected with coordination and planning. It is a normative notion in the sense that high task dependency means that tasks play specific roles in the overall organization of a cognitive system or a cultural cognitive ecosystem; roles that can be played properly or improperly.

(p. 1190)

Slors cites the legal system as a good example of high task dependency: judge, prosecutor, defense attorney, clerk, and other court officials are responsible for tasks organized (and inter-defined) in a holistic way. In other words, for example, what a defense attorney does is understandable

only by referring to what judges and prosecutors do. This indicates a division of labor in a symbiotic system.

Division of labor involves a specific type of offloading, one which is typical for symbiotic cognition but not for extended [mind]. Every participant in a symbiotic system profits from whatever the system as a whole offers (education, justice, social coordination) while contributing only a small part. The tasks, jobs and roles of others in the system co-define and enable one's own task.

(p. 1198).

Both the role of intersubjective interaction and task dependence (and the implied division of labor) hold some implications for how we think about trust or reliance in cognitive institutions in ways that differ from the extended-mind conception of scaffolding institution.

3. Trust vs. reliability in institutional theory

Questions about trust in the case of extended-mind institutions are tied to the idea that the primary function of institutions can be framed in terms of information processing (e.g., Hayek, 1945), whether that happens in a functionally equivalent manner or in terms of a more complex functional integration. In this respect, an institution is treated as a resource that provides or processes information as a precondition for pursuing whatever other institutional function may be at stake (Hindriks & Guala, 2021). We might then ask, how much trust should we put in such resources. There is already a good answer to this question in the EM literature. Clark and Chalmers (1998) had already defined a set of criteria, as a supplement to the parity principle, that became known as the “trust and glue” criteria. These criteria are meant to head off the “cognitive bloat” objection, which is a worry about how far the mind can be extended – does it include all of the resources of the internet, for example? (Rupert, 2004). The answer is that we should include only those resources as part of extended cognition that meet the parity principle and these three criteria. They are clearly stated in Clark (2008).

- (1) That the external resource be reliably available and typically invoked.
[...]
- (2) That any information thus retrieved be more-or-less automatically endorsed. It should not usually be subject to critical scrutiny (unlike the opinions of other people, for example). It should be deemed about as trustworthy as something retrieved clearly from biological memory.
- (3) That information contained in the resource should be easily accessible as and when required. (p. 79)

Although second-wave EM questions the parity principle, it generally accepts these criteria of reliability, trustworthiness, and accessibility (see Aizawa, 2018). The third wave, SEM model, however, tends to reject these criteria as overly narrow.¹ This becomes evident as soon as we try to apply these criteria to the institutional domain. When Clark says “[i]nstitutions, firms, and organizations seem to me to share many of the key properties of pen, paper, and arithmetical practice (Clark, 1997, p. 279), he seems not to consider that institutions would not pass his own three criteria of mind extension. For example, the legal system may not be easily accessible for all parties, but it can still work when one does gain access; accordingly, easy accessibility should not be a litmus test of cognition. Likewise, it’s not clear why automatic endorsement is required to qualify a process as part of extended cognition since critical scrutiny (which may actually involve the opinions of others) itself qualifies as cognitive, and indeed may very well add a trustworthiness or reliability to my thought processes. Indeed, postulating automatic endorsement Clark seems to rule out interaction as an institutional process of trust-making. These considerations suggest that Clark’s “trust and glue” package is unable to account for trust as a criterion of mind extension in circumstances where other agents are involved.

Hawley’s (2014) proposal to distinguish between trust and reliability works well in this context. Roughly, on her view, we can think that reliability applies to impersonal resources – mechanisms or sources of information, tools, and instruments, whereas trustworthiness applies to interpersonal relations where some form of assurance, responsibility or obligation is involved. The “trust and glue” criteria say nothing about assurance, responsibility or obligation. We suggest that these criteria are all about reliability – they aim to test the dependability and availability of resources and the information they convey. Indeed, to the extent that this information could include “the opinions of other people” we have reason to discount it since such opinions would require critical scrutiny. Perhaps that would indicate some degree of distrust of that source of information, although whether this is distrust or simply a judgment of unreliability will depend on the nature of the relation one has with that other person (see below). These criteria, then, suggest that resources such as institutions, including economic institutions, operating as information sources, may be treated as one might treat what Hawley (2017) views as group epistemic testimony – group testimony can have epistemic significance, without entailing group assurance, responsibility or trustworthiness. Trustworthiness, however, becomes an issue in the case of individual epistemic testimony.

The case is different for cognitive institutions understood on the SEM model, specifically because intersubjective interaction and task dependency play essential roles both in the formation and the operation of such institutions (Slors, 2019; see above). They operate as sites of social and cultural

practices with normative expectations and constraints that involve responsibility for carrying out tasks. The model of engagement or coupling here is action and interaction rather than just testimony. Operative members of an institution conduct the affairs of the institution at least in part by means of interpersonal interaction or communication. Operative members are those members who are responsible for decision making, determined by the structure (formal or informal rules/norms) of the institution (Tollefsen, 2004; Tuomela, 1995). As individuals or as a collective, operative members engage with or engage in the operations of the institution by engaging, directly or indirectly with others. The institution functions only through social interactions constrained by normative, task-related factors.

In the example of the legal system, even mulling over contracts, or attempting to come to a judgment based on evidence, involves the work of others who are responsible, and who attest to following the proper legal procedures. Even more clearly, in legal proceedings that involve judges, jurors, attorneys, etc., the interactions come with specifications of tasks and responsibilities for carrying out those tasks. Understood as cognitive institutions, science (Slaby & Gallagher, 2015), educational institutions, corporations, markets, etc. work in a broadly similar fashion.

It's important to recognize that the operation of the cognitive institution is not simply a case of the action of an individual writ large. Hawley (2017) suggests that trustworthiness is tied to the idea that the individual has a certain autonomy. "Regarding an individual as trustworthy is typically caught up with respect for her as autonomous in some way: someone who reliably follows a benevolent despot's orders under duress is not displaying her trustworthiness. Trustworthy behavior often, though not always, reflects a determination to fulfil obligations or commitments which were voluntarily acquired" (p. 246). She then asks:

Do we need to regard groups and organizations in the same way? It may depend upon the group, its constitution and supposed purpose, including questions about whether an organization is an element of the state or a private entity, and whether membership of a particular group is chosen or imposed. To what extent do certain groups have obligations to individuals, or to other groups? As these considerations indicate, issues of trustworthiness in group action quickly involve larger questions about collective responsibility and commitment; a fruitful line of inquiry would be to investigate whether or not these questions can be sidestepped by thinking in terms of reliability rather than trustworthiness.

(p. 246)

Individuals within the context of cognitive institutions may indeed have a kind of autonomy (perhaps best thought of as "relational autonomy," that is, a form of autonomy that is either enabled or constrained by social and normative factors – see Cash, 2013; Mackenzie & Stoljar, 2000). With respect to the functioning of the cognitive institution, it is their responsible

action or work within a task-related framework that counts. The judge must do her work according to the law; the scientist works according to scientific procedure; in market relations, individuals or collectives follow the shared rules that define them as buyers, sellers, producers, consumers, etc. These laws, rules and procedures specify interdependent task responsibilities. In such contexts, trustworthiness is not simply directed to the individual, but also to the role they are playing. We trust that the judge is capable, not only of autonomous behavior, but of making a legal ruling following proper procedure. Indeed, this trust is in addition to any concern about reliability, which may itself qualify or undermine the trustworthiness of the judge who is herself subject to the limitations of her bodily engagement. We may find both the judge and the system unreliable if the judge has not eaten a proper breakfast, and that's different from finding them untrustworthy. A recent study reinforces the idea that hunger can shape, and perhaps even distort, cognitive processes. Danziger et al. (2011) show that whether the judge is hungry or satiated may play an important role in the rational application of legal reasoning (the percentage of favorable rulings drops significantly between breakfast and lunch and returns abruptly after a [food] break). In such cases, the judge may be entirely trustworthy in terms of her intentions to fulfill her legal responsibilities; but she may be entirely unreliable simply because of her metabolic state.

Consider another example more directly relevant to markets. The Securities and Exchange Commission (SEC) imposed a large fine on the Boeing company and its CEO for "putting profits over people" and misleading investors with the aim of improving its image. We know that in the case of the 737 Max aircraft, the Boeing manufacturing process was seriously unreliable, leading to two air disasters and the loss of hundreds of lives. According to Gary Gensler, chairman of SEC, "[i]n times of crisis and tragedy, it is especially important that public companies and executives provide full, fair, and truthful disclosures to the markets."² Boeing failed in this regard too. Its reputation was severely damaged by its attempt to cover up a major design problem, and it lost public and investor trust. Boeing turned out to be both unreliable and untrustworthy, and this case makes it clear that these are two different things.

Reliability and trustworthiness, however, although they can be distinguished, are not unrelated. Contingencies connected with reliability can undermine the degree of trust we may have for the operations of both individuals and the institutions. A similar dynamic is at work in markets where a lack of reliability or trust can introduce inefficiencies and costs, or put the brakes on collaborative enterprise.

Just as institutions do not regulate all possible contingencies, so knowledge of another person can never be so perfect that her decisions can be anticipated with certainty. Also for the calculation of the intentions of the exchange partner it holds true that, even with all attempts at calculation, a final doubt cannot be removed. Despite social norms, a harmful violation of the norms by the trust-giver always remains on the horizon of possibility. Because of the increase of decisions as well as the number of actors involved, time restrictions, scarcity of information, and the need to delegate decisions become more urgent problems. Especially market relations in modern societies demand that risky advance concessions must be made without precise knowledge of the exchange partner, without relying on long-term relations, and despite only incomplete observation of his action with others, as in situations that are not extensively guaranteed by power or norms. The core of the problem of trust consists precisely in this “middle condition between knowing and not-knowing” (Simmel), even when institutions, social norms, power and calculations of interest contribute to forming expectations.

(Beckert, 2005, p. 14)

Judgments of reliability may just be a matter of having sufficient information about and signals from the source. This may be the kind of shallow “trust” one finds in game theory.³ In markets, perceived reliability can be externalized “to interfaces and related protocols among them in their niche to reduce costs of communication” or to signs or signals (David et al., 2021, p. 51). We could think of these as a kind of automatic stand-in for trust. “For example, traffic signals allow drivers to externalize their trust to signals which inform the projection of other drivers’s behavior, as opposed to being left to develop trust with other drivers in order to share the road” (David et al., 2021, p. 51). Likewise in institutions, trust can be externalized to organizational arrangements that are reliable indicators of behavior (a contract, for example, may serve as a barrier to withdrawal [Beckert, 2005]).

Trust may operate in some deeper sense, with respect to the individual, correlative to the psychological complexity of the resource; it has to penetrate to the realm of motives, and to considerations that despite well-meaning advice, the other may in fact not know that a particular action or decision is the wrong one. This deeper, interpersonal trust, has its roots in what Matthew Ratcliffe (2017, p. 150) calls a “primitive, affective, nonconceptual form of trust,” which “develops within, and is then sustained by, the interpersonal environment.”⁴ With respect to the institution, interpersonal trust may be correlative to the organizational complexity of the resource and an understanding of how the system works. One would need an endorsement of not only the reliability of the resource, but the honesty of the resource, based on prior experience with it, or perhaps the testimony of others whom one trusts, where such trust involves the same kind of dialectic with reliability.

4. Trust and reliability in markets

Adam Smith's famous passage on self-interest as market mover has something to do with the trust/reliability distinction.

It is not from the benevolence of the butcher, the brewer, or the baker, that we expect our dinner, but from their regard to their own interest. We address ourselves, not to their humanity but to their self-love, and never talk to them of our own necessities but of their advantages.

(Smith, 1976a, I.ii.2)

Smith can here be understood as saying that to provide for our needs we do not need to trust in whatever sentiment the butcher, the brewer, or the baker may have toward us; rather, we should rely on their self-interest. The mere expectation that those market agents would do what their roles dictate them to do is not enough for considering them recipients of our trust – that is, it would be inaccurate, even paradoxical, to say that we trust in their self-interest. It takes Smith's strictly complementary work as a moral philosopher focusing on people's mutual "sympathy" (Smith, 1976b) to acknowledge that also the proper, relational notion of trust, as economists have recently come to recognize, "is an economic primitive" (Berg et al., 1995, p. 123). It is important to note that Smith's discussion took place at the end of a historical period where markets were such a central part of the economy that the economic organization itself was called a "market economy," but it was not yet, as Fernand Braudel (1977) pointed out, a capitalist system. The advent of capitalism brought about structural changes in markets able to transform the notions of trust and reliance themselves. If Smith was still able to see trust and reliance (or lack thereof) as primarily relational constructs involving strong symbiotic relationships between agents recognized as the butcher, the brewer, and the baker, this would soon not be entirely the case.⁵

Growth in market size, the concentration of market actors, the rise of all sorts of intermediaries, and dematerialization, are just some of the phenomena that make Hawley's distinction between trust and reliance increasingly crucial for understanding contemporary markets. When we today say that we trust or don't trust the market to take care of one or another part of our lives, we almost invariably mean, not that we trust or distrust market actors, but that we find the market mechanism reliable or not reliable. This goes well beyond the issue, by itself not trivial, of distinguishing between individual and collective actors as recipients of trust⁶: it concerns the distinction between actors and instruments itself. Discussing the nature of trust-takers, Beckert (2005), much in agreement with what Hawley would later say, indicates that

Trust is a social expectation which means that trust always refers to another person or a corporate actor. I can trust my friends, my business partners, my bank, but not my bike or the safety of a railing.

(p. 7).

The point is that a market mechanism has much more to do with the safety of a railing than with a friend. Mechanism and railing are both designed to accomplish something. They are, as such, quintessential recipients of reliance, not trust. In this scenario, trust enters into play at best “at access points where the nexus of system and person takes place” (Beckert, 2005, p. 19). It would be hard to overemphasize the centrality of market design (or mechanism design) today in both economics and the economy (Roth, 2008). The more a certain market is considered to be critical, the more it is painstakingly engineered to reach specific goals, among which efficiency usually comes first.⁷ In brief, market mechanisms are designed and employed as reliable instruments of efficiency.⁸ Such a view of markets is part of a more general approach to institutions called “institutional design” (e.g., Goodin, 1998), which sees institutions as entities given “from above” in a top-down fashion (Smith, 2007).

This discussion of market design has much to do with the extended mind foundations of institutions and the idea of trust they support. Not only does Clark (1997, 1998) present markets as pivotal examples of scaffolding institutions, but his way to institutional scaffolding seems to require markets to be the kind of entities given from above as objects of market design. He mentions in this regard the work of Gode and Sunder (1993) showing that certain market rules can lead to efficiency even in the case of unsophisticated cognitive agents called “zero-intelligence” traders. This kind of research seems to dovetail perfectly with Clark’s idea that mind extension occurs when individuals off-load or “externalize” (a term with similar economic and philosophy of mind implications) their cognition onto external resources, in this case the market mechanism. Clark shifts the focus of market scaffolding from individuals to market mechanisms in a way that, as it is usually said in the economic literature, “[s]tructural constraints, not individuals, do much of the explanatory work” (Hodgson, 2004, p. 438). When discussing markets, Clark (1998) vividly conveys the idea that the intelligence of individuals is traded for the intelligence of the designed system:

we [humans] excel in one crucial respect: we are masters at structuring our physical and social worlds so as to press complex coherent behaviors from these unruly resources. We use intelligence to structure our environment so that we can succeed with less intelligence. Our brains make the world smart so that we can be dumb in peace!

(p. 180)

As this approach to institutional mind extension does not directly concern other agents, zero-intelligence seems to have “zero-trust” as a corollary; since these agents’ have only a fiduciary relationship with a mechanism, they are rather “maximum-reliability” agents.

Although this is the most common way markets are understood today, an alternative approach called the “relational paradigm” emphasizes interpersonal relationships in markets (Bowles, 2016; Bruni & Zamagni, 2017; see also Sugden, 2018). Although not usually framed in terms of extended cognition, this approach can lead us toward a third-wave understanding of economic institutions and trust. To do this, consider first the limits that the relational paradigm identifies in the mechanism-focused understanding of markets. The idea of a “relational good” introduced by Carole Uhlaner (1989) can be useful in this regard. The traditional approach to economic transactions would be unable to address, Uhlaner says, an entire class of goods which are called relational exactly in that they are neither created nor consumed by a single person. One example of a relational good is a mattress purchased in the nearby mattress shop with the help of a shop assistant, in contrast to one ordered online. Another example is going to a service to help prepare your taxes; using tax preparation software is not. As these examples show, what is called in economic jargon a relational “good” includes in fact an interpersonal experience, the sort of value-related phenomena that market mechanisms have difficulties accounting for. There are clear third-wave (SEM) elements in the definition of a relational good, for instance their co-creational nature, as “the adjective *relational* is intended to specify what these goods are made of, not just one of their qualities” (Gui, 2013, p. 296, emphasis in original). There is also a way to extend the concept from a collective but still limited dimension, usually two or a few people, to entire institutions through concepts such as “social capital” which point to an entire society’s disposition to engage in relational experiences. There are, however, other aspects which make the analysis of relational goods too limited from the point of view of the SEM model, or third-wave extended mind. Although of course important, the aim of catching “what makes personalized different from non-personalized interaction” (Gui, 2013, p. 299) is somewhat limited. To see what makes this goal limited, consider Julie Nelson’s objection to the idea of a relational good from a radical viewpoint. She points out that we are “in very large degree constituted by our relationships. [. . .] [W]e are continually *created and shaped by* the encounters in which we participate” (Nelson, quoted in Gui, 2013, p. 300, emphasis in original). The idea of a relational good would be problematic if it sees interpersonal relations as an important but dispensable category. Advocates of relational goods may say that consuming a meal at a restaurant with a friend, rather than eating a pre-prepared meal from the fridge alone, increases happiness and well-being, but radical theorists like Nelson would object that such practices do much more than that: they enact

a social practice or institution, and they contribute, in a constitutive way, to who we are, or, to say it differently, they are part of the physiology of a SEM institution. Still another way to put it is to say that lack of interpersonal relations would put in jeopardy the very existence of trustworthy sociality. Given its importance, the next section will be dedicated to a discussion of this point.

5. Institutional physiology: Trust as a criterion of sustainable institutions

Nelson's critique of the insufficient radicalism of the relational paradigm raises a crucial question: are interpersonal relationships and relational trust required for SEM/cognitive institutions in the sense that they are constitutive of such institutions? The emphasis Nelson puts on the "constitutive" role of interpersonal relationships suggests that they are. This is a view consistent with the enactivist emphasis on social interaction as a constitutive process (De Jaegher et al., 2010). As we've seen, however, not all cases of institutional mind extension are considered socially extended. Defining a cognitive institution as an institution "without which some of the agents' cognitive processes would not exist or even be possible" (Petracca & Gallagher, 2020, p. 747), Clark could object that designed markets with no real interpersonal relationships involved can also create new cognitive processes. Dekker (2022) discusses cognitive processes that markets make possible, and some of them seem *prima facie* compatible with either interpersonal relational or non-relational market structures (see also Callon & Muniesa, 2005). We all continuously rely on utterly impersonal mechanisms that involve no relational trust but that nonetheless do extend our minds. Even if most markets are hybrids of impersonal and interpersonal structures, involving both functional integrative and symbiotic processes, both reliability and trust, there seem to be clear examples of purely impersonal market mechanisms running on automatic pilot, and requiring zero-intelligence while paradoxically extending individual cognition.

All this is to say that in some common examples of institutional mind extension, neither interpersonal relationships nor trust seem necessary; the mind-extending coupling between individuals and institutions can sometimes do without them, and be satisfied with reliability. Indeed, some may take this possibility to be an advantage, not just in terms of efficiency, but in terms of overall moral benefit. After all, any arrangement that entails trust, may also result in justified mistrust. The social extension of cognition does not always take desirable forms (Gallagher, 2020). Intrusive institutions and distorted communicative practices can "colonize" the mind, extending it in pathological directions (Castoriadis, 1997). These considerations notwithstanding, we still think that relational trust is a requirement for institutional

mind extension, even as an important part of the impersonal scaffolding type. Clark may have a point in saying that designed scaffolding institutions can extend the mind – sometimes more systematically and efficiently than interpersonal relational SEM institutions – but he says nothing about the sustainability of this form of mind extension. In what follows, we argue that the lack of both interpersonal relationships and relational trust makes institutional mind extension fundamentally unsustainable. Therefore, we introduce trust as a dynamic criterion of institutional mind extension more generally.

The discussion can start with what makes Clark's approach fundamentally static, his idea that mind extension requires the automatic endorsement of external resources (Clark, 2008; see above). Assuming automatic endorsement as a criterion, understanding it normatively, as Clark does, means ruling out most cases of institutional mind extension that take place via interpersonal relationships. The fact that one person rarely (virtually never) endorses "automatically" what another person does or says, makes it difficult for interpersonal relationships to meet Clark's criterion. A consequence of this is, therefore, a significant restriction in the cases of institutional mind extension. But we do not limit our analysis to the constraining limitation of this criterion. Rather, we can follow Clark's logic to the end, accepting the requirement of automatic endorsement of institutional resources as a criterion of mind extension, if only to see that this is dynamically unsustainable. If Clark's criterion implies or leads to the dispensability of interpersonal relationships, our unsustainability argument demonstrates that such relationships are (dynamically) necessary.

To make our case, that is, to show why Clark's criterion of automatic endorsement leads to unsustainability, consider some institutional pathologies related to the idea of automatically endorsed mechanisms. One of them is the reification of relational potential. Reification "means a forgetting of the primal recognition that two humans accord each other in a fundamental process of intersubjective interactions" (Jay, 2008, p. 8). In modern markets, reification is not merely a byproduct of market mechanisms but their direct implication, probably even an assumption, i.e., the assumption that if they did not produce reification – a price for everything – markets would not work as efficiently as they do. But reification cannot be considered a neutral feature of markets, insofar as it has important effects, such as crowding out intrinsic motivation (Frey, 1994). Another pathological effect is counterperformativity (Bamford & MacKenzie, 2018). If reification and its effects are generally expected in modern markets, counterperformativity includes all the unintended ways in which mechanisms may fail or may work differently than expected in changing circumstances. Reliability tends to be short term, not only because mechanisms sometimes break down, but oftentimes because circumstances (including

human inclinations) change. In market counterperformativity, the market mechanism not only fails but creates the conditions for its own failure: it shapes an inflexible reality in which it is bound to fail.

Another pathology is institutional lockout. Even institutions that are perfectly designed internally, sometimes just do not work, because they do not get integrated into the larger institutional system. A conspicuous example is the persistent failure to implement both market-based and state-based development policies (Gunn, 1978). Add to this that an institution or a policy imposed from above may sometimes generate “reactance” in people who feel deprived of their autonomy (Sunstein, 2017).

As the list of pathologies could easily go on, we’ll mention just one more that is wryly presented by Clark as an advantage: the lack of necessity for individual intelligence (a phenomenon not unconnected with the advances of AI). It is easy to suspect that a consistent practice of automatic endorsement may cause the atrophy of cognitive capacities rather than the extension of those capacities; the high reliability of our technical instruments may preempt anything like imaginative cognitive engagement and simply lead to cognitive atrophy. Moreover, this would not be restricted to cognitive capacities since relational capacities would become atrophic as well. Combined this with counterperformativity, and the path clearly leads to unsustainable arrangements.

The pathologies listed above stem from a static and impersonal understanding of institutional mind extension. They all, individually or together, pose a serious threat to the survival of institutions. These issues could be addressed, supporters of the relational paradigm advocate, by restoring the physiology of interpersonal relationships and relational trust (Bowles, 2016; Bruni & Zamagni, 2017; Sugden, 2018). Adam Smith’s message that “[w]ithout trust there can be no progress” (Evensky, 2011, p. 253) can be read in this vein, understanding trust as a condition of institutional continuity.

The 2008 financial crisis is a recent historical example where institutional continuity was seriously put at risk. Although the crisis was not the result of one particular market pathology but of a nefarious combination of pathologies,⁹ the problems that led to this crisis all had something to do with either misplaced reliance on the market mechanism or the erosion of relational trust. Discussing the crisis in some detail may help show how relational trust at sensitive junctures of the economic system could have contributed to mitigating the crisis’ spread and its severity. It is widely known that much of what happened in 2008 had to do with mortgages: they were pooled in securities called asset-based securities (ABSs) which in turn became the underlying assets of another type of securities, collateralized debt obligations (CDOs). Serial pooling was supposed to decrease the risk of the resulting instruments (called ABS CDOs) and obtain good ratings although the quality of underlying mortgages declined as the number of

subscriptions increased. In this process, mortgages were detached from their genuine economic function (i.e., providing borrowers with the means to buy a house) and became mere raw materials for securitization; in a sense, they became “reified”.¹⁰ Brokers’ aggressive practices to procure mortgages – known as “predatory lending” – eroded relational trust between lenders and borrowers in an important transaction in the lives of borrowers (Keys et al., 2010; Mian & Sufi, 2009). A second pathology related to trust in this context was market opacity (or lack of market transparency). Securities like ABSs, CDOs – and a fortiori ABS CDOs – are inherently complex financial instruments. Their correct evaluation requires investors to be able not only to appreciate the instruments’ structure but also the quality of the underlying assets, a process that can take an excruciatingly long time. It was not uncommon at the dawn of the crisis to hear that full transparency would reduce the liquidity (number of exchanges) of securities markets (e.g., Madhavan et al., 2005). Evaluation of ABS CDOs was thus institutionally delegated to rating agencies, which played a central trust-related role in the crisis as the ratings they issued de facto had normative value for investment institutions. As MacKenzie (2011) shows, high ratings for risky ABS CDOs were the result of organizational imbalances within rating agencies in the face of a new exotic financial instrument as well as of an inadequate evaluation model. The model, Gaussian copula pricing models, is today infamously known as “the formula that killed Wall Street” (Salmon, 2009). Misplaced reliance on such a pricing mechanism was at the origin of a snowballing counterperformative process that did not necessarily imply fraud (MacKenzie & Spears, 2014, 2014b). Again, relational trust could have played a mitigating role had better communicative and interactional processes been possible, allowing dissenting voices about the unreliability of the pricing mechanism to be heard within the financial community (e.g., Taleb, 2007). Eventually, misplaced reliance and lack of relational trust at so many levels of the financial crisis caused generalized distrust of markets and other economic institutions (Stevenson & Wolfers, 2011).

Although it has little to do with markets, the timely example of vaccination campaigns can be useful to illustrate the role of interpersonal relationships and relational trust in creating institutional cohesion as a sort of institutional glue. It is well known that part of the population has been reluctant to take vaccines, and various experiments have shown that traditional communication from government or regulatory authorities has no effect, or even negative effects in promoting vaccinations (e.g., Betsch & Sachse, 2013). It has been shown, however, that small-group discussions with experts can have a real impact.

When people encounter a message that aims at changing their minds, they typically generate counter-arguments [...]. If they do not have an interlocutor who can address these counter-arguments (e.g., if they read a leaflet), they are less likely to change their minds. This likely explains why small-group discussion, in which counter-arguments can be addressed in the back and forth of discussion, is vastly more effective at changing people's minds than the simple presentation of arguments. [...] In line with this, direct communication with trustworthy professionals appears to be an efficient lever to increase vaccination acceptance.

(Altay et al., 2021, p. 2)

Although there are understandable concerns about the scalability of small-group discussions, recent evidence by Altay et al. (2021) that a chatbot can change people's attitudes toward vaccines could motivate the sort of critical questions we raised about Clark's approach: is such an approach sustainable? Is it possible to replace trust in interpersonal relationships with reliance on impersonal mechanisms without paying a price in terms of institutional cohesion and sustainability?

6. Concluding remarks

In considering the importance of trust, we've focused on economic institutions, and specifically questions about market mechanisms and the role of social interactions. It's been acknowledged that some recent advances in institutional economics are closely tied to developments in philosophy of mind and cognitive science, roughly, developments that concern extended and enactive cognition (Markey-Towler, 2017; Petracca & Gallagher, 2020). We've argued that the analysis of different conceptions of institutional mind extension, based on Denzau and North's shared mental models, Clark's extended mind, and a more enactive approach that emphasizes the importance of social interaction and personal relationships in SEM institutions, can benefit from Kathrine Hawley's distinction between reliability and trust. Institutional arrangements based solely on the reliability of impersonal mechanisms, we argued, can lead to a variety of social pathologies and, at the extreme, a form of cognitive atrophy, all of which can undermine the sustainability of institutions. Even if trust comes with risks and some degree of unpredictability, it turns out to be a necessary glue-like ingredient in the constitution of social institutions.

Notes

1. For that reason the concept of SEM institution is sometimes thought to succumb to cognitive bloat. There are other ways to answer this worry, however. Specifically, the SEM model suggests that resources can be treated as processes of extended cognition only when there is a proper coupling or engagement with them. If I am not engaged with some part of the legal system, mulling over a contract, or arguing before the bench, for example, that system, or that part of the system, is not part of my extended

cognition; if I am not actually using a search engine, then the internet is not part of my cognitive system. This retort to cognitive bloat works well if one considers extended mind to depend on active engagement.

2. <https://www.sec.gov/news/press-release/2022-170>.
3. As Beckert (2005) puts it, “the modeling of trust games based on game theory is theoretically problematic for assuming the possibility of rational interpretation of signals. They do not do justice to the fundamental uncertainty of the trust-giver, which also composes the core of any meaningful talk of trust” (p. 18).
4. “Trust, in the relevant sense, is something that first arises in infancy. It is a bodily, affective set of interpersonal expectations, which develop through patterned interactions with caregivers and later come to regulate encounters with people more generally. The developmental process and the effects of deviation from it are described by Fonagy and Allison (2014). They maintain that secure attachments in early life foster a sense of trust in others that later generalizes, disposing one to accept credible communications, while also instilling confidence in one’s own judgments and abilities (something that relies on feedback from others) Fonagy and Allison (2014, p. 374) add that attachment insecurity in adults is associated with epistemic biases that include intolerance of ambiguity, inflexible and dogmatic thinking, and a tendency to make judgments based on insufficient information” (Ratcliffe, 2017, p. 161). As we will discuss, for some social theorists this interpersonal kind of trust just is what trust means.
5. Smith, however, does not restrict trust and reliance to the interpersonal sphere. When he refers to recipients of trust, he often refers to collective entities, such as the government, or to less directly agentive entities, such as the law. In Smith’s reasoning there seems to be nonetheless a principle of discrimination. Consider Smith’s example of Hamburg (Evensky, 2011; Smith, 1976a, V.ii.f.10;), “a small republick” where people are said to pay their taxes “with great fidelity” because they “have entire confidence in their magistrates” (note here the symbiotic reference to magistrates and not to the government). Trust in collective entities seems to boil down to strict relational trust when the involved number of relationships is small.
6. Trust toward collective agents is called “institutional trust”. In certain circumstances, but not always, it can be considered as a form of relational trust (Bachmann & Inkpen, 2011).
7. The worldly success of market design is especially the merit of auction theory, famously used since the 1990s by the Federal Communication Commission (FCC) to allocate the frequencies of the radio spectrum.
8. Among other goals of market design are equity and stability.
9. Both narrative and quantitative reconstructions agree that the crisis was the result of a combination of circumstances rather than of a single major cause (e.g., Rose & Spiegel, 2012).
10. It must be stressed that the problem is not securitization per se (as it is a useful practice for funding lending), but how the genuine economic function of the practice can be distorted.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Notes on contributors

Shaun Gallagher, PhD, Hon DPhil., is the Lillian and Morrie Moss Professor of Excellence in Philosophy at the University of Memphis, and Professorial Fellow at the School of Liberal Arts, University of Wollongong (AU). His areas of research include phenomenology, philosophy of mind, embodied cognition, social cognition, and concepts of self. His recent publications include *Action and Interaction* (Oxford 2020); *Performance/Art: The Venetian Lectures* (Milan 2021); *The Phenomenological Mind* (Routledge 2021 – 3rd ed); and *Enactivist Interventions: Rethinking the Mind* (Oxford 2017). He is editor-in-chief of the journal *Phenomenology and the Cognitive Sciences*.

Enrico Petracca is a Research Associate at the School of Economics, Management and Statistics of the University of Bologna (Italy). His research interests are at the intersection between economics, embodied cognitive science, and philosophy of mind.

ORCID

Shaun Gallagher  <http://orcid.org/0000-0002-3147-9929>

Enrico Petracca  <http://orcid.org/0000-0001-7572-8117>

References

- Aizawa, K. (2018). Extended cognition, trust and glue, and knowledge. In C. J. Adam, A. Clark, J. Kallestrup, S. O. Palermos, & D. Pritchard (Eds.), *Extended epistemology* (pp. 64–78). Oxford University Press.
- Altay, S., Hacquin, A. S., Chevallier, C., & Mercier, H. (2021). Information delivered by a chatbot has a positive impact on COVID-19 vaccines attitudes and intentions. *Journal of Experimental Psychology Applied*. Advance online publication. <https://doi.org/10.1037/xap0000400>
- Bachmann, R., & Inkpen, A. C. (2011). Understanding institutional-based trust building processes in inter-organizational relationships. *Organization Studies*, 32(2), 281–301. <https://doi.org/10.1177/017084061039747>
- Baier, A. (1986). Trust and antitrust. *Ethics*, 96(2), 231–260. <https://doi.org/10.1086/292745>
- Bamford, A., & MacKenzie, D. (2018, September/October). Counterperformativity. *New Left Review*, 113(Sept/Oct), 97–121.
- Beckert, J. (2005). Trust and the performative construction of markets (Discussion Paper No. 05/08). Max Planck Institute for the Study of Societies. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2465811
- Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, reciprocity, and social history. *Games and Economic Behavior*, 10(1), 122–142. <https://doi.org/10.1006/game.1995.1027>
- Betsch, C., & Sachse, K. (2013). Debunking vaccination myths: Strong risk negations can increase perceived vaccination risks. *Health Psychology*, 32(2), 146–155. <https://doi.org/10.1037/a0027387>
- Bowles, S. (2016). *The moral economy: Why good incentives are no substitute for good citizens*. Yale University Press.
- Braudel, F. (1977). *La Dynamique du capitalisme [The dynamics of capitalism]*. Arthaud.
- Bruni, L., & Zamagni, S. (2017). *Civil economy: Another idea of the market*. Agenda Publishing.

- Callon, M., & Muniesa, F. (2005). Peripheral vision: Economic markets as calculative collective devices. *Organization Studies*, 26(8), 1229–1250. <https://doi.org/10.1177/0170840605056393>
- Cash, M. (2013). Cognition without borders: ‘Third wave’ socially distributed cognition and relational autonomy. *Cognitive Systems Research*, 25(1), 61–71. <https://doi.org/10.1016/j.cogsys.2013.03.007>
- Castoriadis, C. (1997). *The Imaginary Institution of Society* K. Blamey (Ed.). MIT Press. Original work published in 1975.
- Clark, A. (1997). Economic reason: The interplay of individual learning and external structure. In J. Drobak & J. Nye (Eds.), *The frontiers of the new institutional economics* (pp. 269–290). Academic Press.
- Clark, A. (1998). *Being there: Putting brain, body, and world together again*. MIT Press.
- Clark, A. (2008). *Supersizing the mind: Embodiment, action, and cognitive extension*. Oxford University Press.
- Clark, A., & Chalmers, D. (1998). The extended mind. *Analysis*, 58(1), 7–19. <https://doi.org/10.1093/analysis/58.1.7>
- Danziger, S., Levav, J., & Avnaim-Pesso, L. (2011). Extraneous factors in judicial decisions. *Proceedings of the National Academy of Sciences*, 108(17), 6889–6892. <https://doi.org/10.1073/pnas.101803310>
- David, S., Cordes, R. J., & Friedman, D. A. (2021). Active inference in modeling conflict. *Zenodo*. <https://doi.org/10.5281/zenodo.5750935>
- De Jaegher, H., DiPaolo, E., & Gallagher, S. (2010). Can social interaction constitute social cognition? *Trends in Cognitive Sciences*, 14(10), 441–447. <https://doi.org/10.1016/j.tics.2010.06.009>
- Dekker, E. (2022). How cognitive institutions and interpretative rationality enable markets with infinite variety. In D. J. D’Amico & A. G. Martin (Eds.), *Contemporary methods and Austrian economics* (pp. 151–167). Emerald Publishing Limited.
- Denzau, A. T., & North, D. C. (1994). Shared Mental Models: Ideologies and Institutions. *Kyklos*, 47(1), 3–31. <https://doi.org/10.1111/j.1467-6435.1994.tb02246.x>
- DiPaolo, E., & Thompson, E. (2014). The enactive approach. In L. Shapiro (Ed.), *The Routledge handbook of embodied cognition* (pp. 68–78). Routledge.
- Evensky, J. (2011). Adam Smith’s essentials: On trust, faith, and free markets. *Journal of the History of Economic Thought*, 33(2), 249–267. <https://doi.org/10.1017/S1053837211000034>
- Fonagy, P., & Allison, E. (2014). The role of mentalizing and epistemic trust in the therapeutic relationship. *Psychotherapy*, 51(3), 372–380. <https://doi.org/10.1037/a0036505>
- Frey, B. S. (1994). How intrinsic motivation is crowded out and in. *Rationality and Society*, 6(3), 334–352. <https://doi.org/10.1177/1043463194006003004>
- Gallagher, S. (2013). The socially extended mind. *Cognitive Systems Research*, 25(1), 4–12. <https://doi.org/10.1016/j.cogsys.2013.03.008>
- Gallagher, S. (2017). *Enactivist interventions: Rethinking the mind*. Oxford University Press.
- Gallagher, S. (2018). The extended mind: State of the question. *The Southern Journal of Philosophy*, 56(4), 421–447. <https://doi.org/10.1111/sjp.12308>
- Gallagher, S. (2020). *Action and interaction*. Oxford University Press.
- Gallagher, S., & Crisafi, A. (2009). Mental institutions. *Topoi*, 28(1), 45–51. <https://doi.org/10.1007/s11245-008-9045-0>
- Gallagher, S., Mastrogiorgio, A., & Petracca, E. (2019). Economic reasoning and interaction in socially extended market institutions. *Frontiers in Psychology*, 10, 1856. <https://doi.org/10.3389/fpsyg.2019.01856>

- Gode, D. K., & Sunder, S. (1993). Allocative efficiency of markets with zero-intelligence traders: Market as a partial substitute for individual rationality. *The Journal of Political Economy*, 101(1), 119–137. <https://doi.org/10.1086/261868>
- Goodin, R. E. (Ed.). (1998). *The theory of institutional design*. Cambridge University Press.
- Gui, B. (2013). Relational goods. In L. Bruni & S. Zamagni (Eds.), *Handbook on the economics of reciprocity and social enterprise* (pp. 295–305). Edward Elgar Publishing.
- Gunn, L. A. (1978). Why is implementation so difficult? *Management Services in Government*, 33(4), 169–176.
- Hawley, K. (2014). Trust, distrust and commitment. *Noûs*, 48(1), 1–20. <https://doi.org/10.1111/nous.12000>
- Hawley, K. J. (2017). Trustworthy groups and organisations. In P. Faulkner & T. Simpson (Eds.), *The philosophy of trust* (pp. 230–250). Oxford University Press.
- Hayek, F. A. (1945). The use of knowledge in society. *The American Economic Review*, 35(4), 519–530.
- Hindriks, F., & Guala, F. (2021). The functions of institutions: Etiology and teleology. *Synthese*, 198(3), 2027–2043. <https://doi.org/10.1007/s11229-019-02188-8>
- Hodgson, G. M. (2004). *The evolution of institutional economics*. Routledge.
- Jay, M. (2008). *Introduction A. Honneth*, reification: A new look at an old idea. Oxford University Press. (pp. 3–16).
- Keys, B. J., Mukherjee, T., Seru, A., & Vig, V. (2010). Did securitization lead to lax screening? Evidence from subprime loans. *The Quarterly Journal of Economics*, 125(1), 307–362. <https://doi.org/10.1162/qjec.2010.125.1.307>
- Kirchhoff, M. D. (2012). Extended cognition and fixed properties: Steps to a third-wave version of extended cognition. *Phenomenology and the Cognitive Sciences*, 11(2), 287–308. <https://doi.org/10.1007/s11097-011-9237-8>
- MacKenzie, D. (2011). The credit crisis as a problem in the sociology of knowledge. *The American Journal of Sociology*, 116(6), 1778–1841. <https://doi.org/10.1086/659639>
- MacKenzie, D., & Spears, T. (2014). ‘The formula that killed Wall Street’: The Gaussian copula and modelling practices in investment banking. *Social Studies of Science*, 44(3), 393–417. <https://doi.org/10.1177/0306312713517157>
- MacKenzie, D., & Spears, T. (2014b). ‘A device for being able to book P&L’: The organizational embedding of the Gaussian copula. *Social Studies of Science*, 44(3), 418–440. <https://doi.org/10.1177/0306312713517158>
- Mackenzie, C., & Stoljar, N. (Eds.). (2000). *Relational autonomy: Feminist perspectives on autonomy, agency, and the social self*. Oxford University Press.
- Madhavan, A., Porter, D., & Weaver, D. (2005). Should securities markets be transparent? *Journal of Financial Markets*, 8(3), 265–287. <https://doi.org/10.1016/j.finmar.2005.05.001>
- Markey-Towler, B. (2017). I, ‘Roboticus Oeconomicus’: The philosophy of mind in economics, and why it matters. *Cambridge Journal of Economics*, 41(1), 203–237. <https://doi.org/10.1093/cje/bew014>
- Menary, R. (2007). *Cognitive integration: Mind and cognition unbounded*. Palgrave Macmillan.
- Menary, R. (2010). Cognitive integration and the extended mind. In R. Menary (Ed.), *The Extended Mind* (pp. 227–243). MIT Press.
- Mian, A., & Sufi, A. (2009). The consequences of mortgage credit expansion: Evidence from the US mortgage default crisis. *The Quarterly Journal of Economics*, 124(4), 1449–1496. <https://doi.org/10.1162/qjec.2009.124.4.1449>
- Newen, A., De Bruin, L., & Gallagher, S. (Eds.). (2018). *Oxford handbook of 4E cognition*. Oxford University Press.
- North, D. C. (2005). *Understanding the process of economic change*. Princeton University Press.

- Petracca, E., & Gallagher, S. (2020). Economic cognitive institutions. *Journal of Institutional Economics*, 16(6), 747–765. <https://doi.org/10.1017/S1744137420000144>
- Ratcliffe, M. (2017). Selfhood, schizophrenia, and the interpersonal regulation of experience. In C. Durt, T. Fuchs, & C. Tewes (Eds.), *Embodiment, enaction and culture* (pp. 149–171). MIT Press.
- Rose, A. K., & Spiegel, M. M. (2012). Cross-country causes and consequences of the 2008 crisis: Early warning. *Japan and the World Economy*, 24(1), 1–16. <https://doi.org/10.1016/j.japwor.2011.11.001>
- Roth, A. E. (2008). What have we learned from market design? *The Economic Journal*, 118(527), 285–310. <https://doi.org/10.1111/j.1468-0297.2007.02121.x>
- Rupert, R. D. (2004). Challenges to the hypothesis of extended cognition. *The Journal of Philosophy*, 101(8), 389–428. <https://doi.org/10.5840/jphil2004101826>
- Salmon, F. (2009, February 23) Recipe for disaster: The formula that killed Wall Street. *Wired*. <https://www.wired.com/2009/02/wp-quant/>
- Searle, J. R. (2005). What is an institution? *Journal of Institutional Economics*, 1(1), 1–22. <https://doi.org/10.1017/S1744137405000020>
- Slaby, J., & Gallagher, S. (2015). Critical neuroscience and socially extended minds. *Theory Culture & Society*, 32(1), 33–59. <https://doi.org/10.1177/0263276414551996>
- Slors, M. (2019). Symbiotic cognition as an alternative for socially extended cognition. *Philosophical Psychology*, 32(8), 1179–1203. <https://doi.org/10.1080/09515089.2019.1679591>
- Smith, A. (1976a). *An inquiry into the nature and causes of the wealth of nations*. R. H. Campbell & A. S. Skinner, Eds.; W. B. Todd Textual (Eds. Eds.). Clarendon Press. Original work published in 1776.
- Smith, A. (1976b). *The theory of moral sentiments*. D. D. Raphael & A. L. Macfie (Eds.). Clarendon Press. Original work published in 1759.
- Smith, V. L. (2007). *Rationality in economics: Constructivist and ecological forms*. Cambridge University Press.
- Stevenson, B., & Wolfers, J. (2011). Trust in public institutions over the business cycle. *The American Economic Review*, 101(3), 281–287. <https://doi.org/10.1257/aer.101.3.281>
- Stewart, J. R., Gapenne, O., & DiPaolo, E. A. (Eds.). (2010). *Enaction: Toward a new paradigm for cognitive science*. MIT Press.
- Sugden, R. (2018). *The community of advantage: A behavioural economist's defence of the market*. Oxford University Press.
- Sunstein, C. R. (2017). Nudges that fail. *Behavioural Public Policy*, 1(1), 4–25. <https://doi.org/10.1017/bpp.2016.3>
- Sutton, J. (2010). Exograms and interdisciplinarity: History, the extended mind, and the civilizing process. In R. Menary (Ed.), *The Extended Mind* (pp. 189–225). MIT Press.
- Taleb, N. N. (2007). *The black swan: The impact of the highly improbable*. Random House.
- Tollefsen, D. (2004). Collective intentionality. *Internet Encyclopedia of Philosophy*. <https://iep.utm.edu/collective-intentionality/>
- Tuomela, R. (1995). *The importance of us*. Stanford University Press.
- Uhlaner, C. J. (1989). 'Relational goods' and participation: Incorporating sociability into a theory of rational action. *Public Choice*, 62(3), 253–285. <https://doi.org/10.1007/BF02337745>
- Varela, F. J., Thompson, E., & Rosch, E. (1991). *The embodied mind: Cognitive science and human experience*. MIT Press.