Journal of Institutional Economics (2025), **21**, e39, 1–19
doi:10.1017/S174413742510026X

CAMBRIDGE UNIVERSITY PRESS

RESEARCH ARTICLE

Norms are relational: cognitive institutions, practices, and the 'where' question

Enrico Petracca[1] and Shaun Gallagher[2,3]

[1]Konrad Lorenz Institute for Evolution and Cognition Research, Klosterneuburg, Austria, [2]Department of Philosophy, University of Memphis, Memphis, USA and [3]SOLA, University of Wollongong, Wollongong, Australia
Corresponding author: Enrico Petracca; Email: enrico.petracca@kli.ac.at

(Received 13 December 2024; revised 4 September 2025; accepted 4 September 2025)

Abstract
Post-Northian institutional economics has been predicated on the socially extended and enactivist concept of cognitive institution. It has recently been suggested that this framework should include North's definition of institutions as 'rules of the game'. In this paper, we agree with this normative turn but take issue with the mental-model framework in which it is proposed. Retaining both shared mental models and rules of the game remains too 'Northian', even if complemented with enactivist dynamic principles of mental-model change. We propose an alternative enactivist concept of norm that entirely avoids mental models. We base it on an alternative social ontology that considers norms as located *in the relation* between agents and institutions. The implications of this relational ontology for the norms (or principles) of rationality are also discussed. We argue that a truly relational framework requires abandoning the adaptationist norm of rationality in favour of coordinative rationality principles.

Keywords: cognitive institutions; extended mind; norms; post-Northian institutional economics; rationality

> Whether we are dealing with formal or informal rules,
> we need to consider the ways in which rules are enacted
> (Hodgson, 2006: 12)

Introduction

We don't know whether, in drawing attention to the enaction of rules, in the epigraph, Geoffrey Hodgson was referring to 'enactivism' in the philosophy of the cognitive sciences (Gallagher, 2017; Varela *et al.,* 1991). We take his words, however, as an indication that some convergence between institutional economics and enactivism was occurring below the radar before it recently surfaced to the light of day (Gallagher *et al.,* 2019; Petracca and Gallagher, 2020; Frolov, 2023a). Importantly, Hodgson's words concern a topic central to institutional economics: the concept of *rules* that, after Douglass North, has shored up an entire view of institutions – the idea of institutions 'as the rules of the game in society' (1990: 3).

In this paper, we explicate what remains implicit in Hodgson's epigraph and explore some consequences that follow from this. These tasks are crucial because the consequences of an explicit enactivist turn may be momentous for institutional economics. We argue that the enactivist view of

how rules are enacted is at odds with long-established traditions in institutional economics, notably with North's 'cognitive institutionalism' (North, 2005; Mantzavinos *et al.,* 2004; see also Mantzavinos, 2001). Although studying how rules are enacted looks like a natural follow-up to North, once enaction is understood in terms of philosophical enactivism, especially its 'radical' variants, not much room is left for North's internalist view of 'shared mental models' (Denzau and North, 1994). Institutions are more than just rules/norms, but if norms operate as cornerstones of institutions, enactivism provides an alternative understanding of their cognitive, behavioural, and moral dimensions – here comprehensively called 'normativity'.

Although the 'post-Northian' label embraces all attempts to overcome North's philosophical and cognitivist foundations, there has been some convergence in tying post-Northianism to the enactivist concept of a 'cognitive institution' (Dekker, 2022; Dekker and Remic, 2024; Frolov, 2023a; 2024; Gallagher *et al.,* 2019; Gallagher and Petracca, 2024; Kuchař, 2025; Petit and Ballet, 2025; Petracca and Gallagher, 2020; 2025; Remic, 2021). Within this enactivist turn, Daniil Frolov (2023a) has linked North's idea of rules of the game with an enactivist conception of norms. While we consider Frolov's work an important contribution to the study of cognitive institutions, we suggest that there is an alternative way to integrate norms within the enactivist framework without relying on the Northian concept of mental models. To this effect, we offer a *strictly* externalist, relational, and radical (i.e., non-mental-model-based) framework for understanding norms in cognitive institutions. A useful way to contrast our proposal to Frolov's is in terms of where we think norms are 'located' in the ontology of cognitive institutions. Beck (2024) calls this sort of ontological issue the 'where' question (see also Hindriks, 2013, concerning the 'location problem' in social ontology). Addressing the *where* question also entails addressing the broader question of 'what' norms are. Our view is that a norm is located *in the relation* between individuals and institutions, and not, as currently argued, in *both* individuals *and* institutions. Our more accentuated distance from North is ultimately justified by what we think is a more manageable social ontology that dispenses with (possibly redundant) mental models, the distinction between a norm's content and its form, and the necessity to identify different locations for content and form.

Our argument unfolds in three parts. First, before proposing our enactivist view of norms, a few clarifications concerning the post-Northian debate seem necessary. Understanding where norms are ontologically located in post-Northian cognitive institutions requires understanding the concept of extended mind as it has been recently applied in institutional economics. Section '*Waves' of the extended mind and institutional theory* surveys the use of the extended mind in this context and warns about the possible informal use/overuse or dilution of the concept in seeking 'alignments' with extant approaches to institutional economics. Adopting an extended mind framework means raising genuinely new ontological questions and providing new answers. In Section *Where are the norms in cognitive institutions?*, we discuss the enactivist claim that *norms are in the relation* and offer a radical interpretation of this statement. We argue that to frame norms as radically relational, practice is the most suitable construct, which will lead us to argue that a norm lies ontologically in the practice. Finally, in Section *Normativity and rationality: another location problem*, we discuss the implications of this relational concept for a related location problem concerning the norms of rationality. We suggest that the ontology of cognitive institutions makes the *adaptive* norms of bounded/ecological rationality problematic to use. In cognitive institutions, it is hard to disentangle (locationally) institutions as part of the agent-institution coupling from institutions as part of the task environment in which rationality is to be assessed. For this reason, we advocate the introduction of new concepts, terminology, and, in particular, new normative coordinative principles and criteria of institutional rationality.

'Waves' of the extended mind and institutional theory

We start with a few clarificatory remarks about what has been identified as three 'waves' of extended mind theory. Since the enactivist view of cognitive institutions builds on the extended mind concept, it's important to clarify these different waves and their relevance for the broader post-Northian debate.

https://doi.org/10.1017/S174413742510026X Published online by Cambridge University Press

The extended mind hypothesis proposes that environmental factors contribute to cognitive processes and can thereby form part of an individual agent's mind. On this externalist view, mental processes are not confined to the head but can extend into the environment.

- The first wave of extended mind theory, associated primarily with Andy Clark and David Chalmers (1998), contends that mind extension depends on a functional equivalence existing between internal and external processes (the so-called 'parity principle').
- The second wave takes issue with the parity principle and argues for a complementary or integration relation between internal processes and external resources (Sutton, 2010; Menary, 2010).
- The third wave focuses on the more dynamical aspects of these relations and emphasises the enactive intersubjective/social dimensions of the 'socially extended mind' (Cash, 2013; Gallagher, 2013; 2018; 2023; Kirchhoff, 2012).

The first wave: North and Clark

The economic literature on cognitive institutions has duly traced the concept to the philosophical debate over the extended mind hypothesis and its 'waves'. The short version of the story is that in the mid-1990s, North entertained a fruitful exchange with Andy Clark, philosopher and co-progenitor with David Chalmers of the 'extended mind' concept (Clark and Chalmers, 1998), and that this collaboration led to a new framework for economic institutions aligning with the 'first wave' of extended-mind theory. Although overall correct, this story risks overlooking subtle differences in North and Clark's contributions, including the fact that they introduced two alternative concepts of institutions: institutions as 'shared mental models' (Denzau and North, 1994) and 'scaffolding institutions' (Clark, 1997; 1998). The two views are closely related but differ on important points, which ultimately reflect different degrees of externalism. For Denzau and North, the starting point of analysis is, from the cognitive viewpoint, shared mental models. It is usual to see North's contribution as composed of two sides: the cognitive side, consisting of shared mental models, and the institutional side, consisting of rules; but this does not mean that the two sides are equal for North: shared mental models run deeper than institutions.

> 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. Some types of mental models are shared intersubjectively. If different individuals have similar models they are able to better communicate and share their learning. Ideologies and institutions can then be viewed as classes of shared mental models. (Denzau and North, 1994: 4)

Institutions can be classes of shared mental models external to the mind only if the former are *emanations* or consequences of the latter. As Denzau and North put it, 'institutions, clearly are a reflection of the evolving mental models' (1994: 13). As systems of rules, institutions are the outcome, in terms of constructed environment, of shared mental models, and shape the latter in the process of learning or belief transmission, after institutions are set up. Moreover, institutions change only when shared mental models change (North, 2005), confirming the initial prominence of the latter.

Clark's view of institutions as scaffolds is symmetrical to North's, but with much less emphasis on mental models. Clark's concept of extended mind sees external resources as the means for 'off-loading' mental models. Once offloaded, the use of external resources or mechanisms can substitute for internal processes, just as the use of a notebook can substitute for the workings of long-term memory, or the use of pencil and paper can ease the computation of a long multiplication (Clark and Chalmers, 1998). Pen,

https://doi.org/10.1017/S174413742510026X Published online by Cambridge University Press

paper, and the external manipulation of symbols can scaffold and even replace some internal operations of mathematical reasoning. The same externalist logic applies to institutions.

> We can now see, in barest outline, how institutions, firms, and organizations seem to share many of the key properties of pen, paper, and arithmetical practice in this example. Pen and paper provide an external medium in which we behave (using basic on-line resources) in ways dictated by the general policy or practice of long multiplication. [. . .] Similarly, firms and organizations provide an external resource in which individuals behave in ways dictated by norms, policies, and practices. (Clark, 1998: 185)

Now, if North's shared mental models 'emanate' institutions, such that the relevant nexus runs from the inside (mental models) to the outside (institutions), on Clark's view, the nexus seems to be reversed: institutions provide a scaffold or relief to the mind, substituting or easing the necessity for mental models. Institutions operate as external scaffolds, resources or constraints for cognition and behaviour in the forms of 'norms, policies, and practices that may even become internalized as mental models' (Clark, 1997: 279).

Why does this difference between North and Clark matter?[1] For one, it bears on the extent to which their respective institutional models align with the first-wave extended mind predicated on the *parity principle* (Clark and Chalmers, 1998). The core of this principle is that to qualify as an external resource able to extend the mind, the resource – pen and paper or an entire institution – must perform the same function as an internal resource. Just as pen and paper substitute for the work of our neurons in complex multiplications, for Clark institutions do likewise as we engage in solving societal problems. This, however, doesn't seem to be exactly North's position. As indicated, for North institutions are emanations, reflections of beliefs, such that they don't entirely substitute for beliefs. Considered to be both 'shared mental models' and 'humanly devised constraints' (North, 1990: 3), institutions and internal mental states often redundantly bear the same content. In contrast to the parity principle, understood as a *substitution* principle requiring that an external resource do the work in place of an internal resource, North seems to express an *emanation* principle. Even if North endorses a form of external offloading, this externalism continues to depend on internal processes.

Other (mis)alignments?

It has recently been hypothesised that other influential approaches in institutional economics are aligned with the first-wave extended mind (Frolov, 2024). One of the 'aligned' approaches is Kahneman's behavioural economics, which is said to resonate, in his 'social mind' version of nudge, boost, and other policy interventions, with the first-wave extended mind. As seen above, to have genuine first-wave alignment, the parity principle must hold, but it is not obvious that what happens in a nudge or a boost is functionally equivalent and substitutes for what happens in individuals' minds. While a nudge may in some meaningful sense substitute for agents' internal resources, this does not seem to be the case for boost. In boost policies, the environment provides agents with cognitive resources (in terms of education, competences, etc.) that are supposed to persist after the coupling is broken (Hertwig and Grüne-Yanoff, 2017), which would suffice to make a first-wave alignment improbable.

In the philosophical literature, more specific criteria have been introduced to decide whether an external resource qualifies as part of an extended cognitive system. David Kaplan (2012), for instance, appeals to the criterion of 'mutual manipulability' (MM), i.e., if an intervention on a component of the system changes the system, and if an intervention at the systemic level changes that component, then that component is a constitutive part of the system (Craver, 2007). On this interpretation, something

[1]The difference between North and Clark lies in what they identify as priorities. When they consider the dynamic interplay between beliefs and institutions, the two positions converge and become indistinguishable.

https://doi.org/10.1017/S174413742510026X Published online by Cambridge University Press

would count as a constitutive part of an extended cognitive system if it passes the MM test. Kaplan (2012), referencing an example given by Clark and Chalmers (1998), of a person, Otto, with poor memory using a notebook to remember, suggests that Otto's notebook is part of his cognitive system since if we take away the notebook, his cognitive capacity declines, and if we change the memory task, he will rewrite his notebook. While nudges and boosts are designed to change an individual's behaviour, it is not obvious that manipulating the individual's behaviour would change, in turn, the nudge or the boost. They would not pass the MM test. This suggests that superimposing a first-wave alignment on nudge-boost behavioural economics may be hasty. To be clear, redrafting nudge and boost theory to make them more dynamic and compatible with the extended mind is possible, and indeed welcome (e.g., Kirchhoff, 2017). But since this is not the current understanding of nudge and boost, a retrospective extended-mind interpretation seems not entirely justified.

The second influential view, deemed to align with the first-wave extended mind, is Gigerenzer's 'fast-and-frugal' heuristics (FFHs). FFHs offers a criterion of environmental *adaptation* for discriminating whether heuristics – easy-to-use judgement and decision procedures – are effective in specific environments. The formula of ecological/adaptive rationality reads: '[a] heuristic is ecologically rational to the degree that it is adapted to the structure of the environment' (Gigerenzer and Todd, 1999: 13). Although the environment is a crucial component in explaining the effectiveness of FFHs, ecological rationality and extended mind deal with environments differently, making 'the two perspectives hard to integrate' (Petracca, 2021: 8). For one, the use of FFHs doesn't seem to satisfy the parity principle. On this point, we concur with Arnau et al. (2014), who have explicitly denied that environments can be considered pieces of the extended mind in FFHs (see also Petracca and Grayot, 2023). They contrast an 'ontological' environment, where the environment is a resource extending individuals' minds, with a 'methodological' environment, where the environment offers only conditions for discriminating the success of heuristics. If environments play only a methodological role, as in the current FFHs perspective, an alignment with first-wave extended mind may be another unnecessary superimposition on the extant theory.

The second wave and cognitive norms in new institutional economics

Second-wave extended mind (Menary, 2010; Sutton, 2010) rejects the parity principle. What enhances individuals and their cognitive success, on this view, is the heterogeneity between internal and external resources: their (functional) complementarity, not their (functional) parity. As John Sutton (2010: 194) puts it, 'different components of the overall (enduring or temporary) system can play quite different roles and have different properties while coupling in collective and complementary contributions to flexible thinking and acting'. The process of mind extension thus becomes a process of *integration* between heterogeneous internal and external resources (Menary, 2007). It is in this context that Richard Menary introduces a discussion of norms and rules, and more generally highlights the role of normativity for mind extension. Integration requires some form of active agential manipulation of external resources, a manipulation that Menary considers to be inherently normative. '[A] crucial role of integration-style arguments is to show that we cannot make good on the manipulation thesis without understanding the normativity of the bodily manipulations of external vehicles of cognition' (Menary, 2010: 228).

Following up on Clark's example of pen-and-paper multiplication, Menary's example also concerns mathematical practice but is more general. It concerns the entire gamut of ways in which individuals couple with the mind-extending resources offered by mathematics. To perform pen-and-paper multiplication, agents must follow an algorithm that gives a procedural *command*, which is inherently normative because it must be executed exactly to return a correct result. The kind of normativity at work in this example is ubiquitous to the point that Menary sees mathematics as a whole as a normative system regulating the manipulation of symbols and operations. To explain individuals' ability to manipulate external resources like mathematics, Menary introduces the key concept of a 'cognitive norm'.

https://doi.org/10.1017/S174413742510026X Published online by Cambridge University Press

> We are able to manipulate external vehicles [e.g., mathematical symbols] because we gain manipulative abilities that are governed by cognitive norms. These are norms that govern manipulations of external representations, which aim at completing cognitive tasks. (Menary, 2010: 238)

Cognitive norms offer a bridge with the institutional economics literature in the possible affinity, noticed by Frolov (2023a), between Menary's concept and what Greif and Mokyr (2017) call 'cognitive rules'. This affinity has led Frolov to see an interesting alignment between second-wave extended mind and post-Northian institutional economics, although he, more interested in the third-wave, doesn't discuss the details of this possible alignment. In this section, we'll examine the terminological and conceptual affinity between Menary and Greif & Mokyr to see if, or how, the latter's view can qualify as an instance of second-wave mind extension.

Greif and Mokyr define a cognitive rule as a social construct that conveys information, 'which distills and summarises society's beliefs and experience' (2017: 2). Although Greif and Mokyr aim explicitly to go 'beyond' North, it is difficult to tell whether they succeed. On the one hand, they criticise North's reduction of institutions to rules and suggest a transition toward an integrative definition consisting of both rules and equilibria (see Hindriks and Guala, 2015). On the other hand, their cognitive assumptions seem to entirely reflect North's. Similarly to shared mental models, cognitive rules 'summarize and aggregate society's beliefs and attitudes' (2017: 3) and do other things that North ascribes to shared mental models, such as imposing constraints and providing cognitive order. Nonetheless, Greif and Mokyr do point in a genuinely new and promising direction beyond North when they emphasise *social interactions* (2017: 2, fn1).

At this point, we can return to the question of the alignment of Greif and Mokyr with Menary. A way to reconcile the two views is to see cognitive rules (in Greif and Mokyr's sense) as the rules that individuals use in a specific kind of environmental manipulation (in Menary's sense), namely, in social interactions. In other words, individuals would use cognitive rules as a means to extend their minds through social interactions. However, both approaches eventually fall short of delivering mind extension through social interactions. On Menary's side, although his view notionally includes institutions, these don't seem to be of the interpersonal and socially interactive kind. If mathematics counts as an institution, it is not immediately recognisable as one where social interactions play any central role. But we are also short on Greif and Mokyr's side, and for the simple reason that they don't seem to have an even remote interest in the extended mind (which cannot be said of North, despite his differences with Clark). To fill these gaps, we now discuss third-wave extended mind and cognitive institutions.

The third wave and cognitive institutions: some initial considerations

As noted, it is hard to think of mathematics as an institution in the usual sense of institutional economics. Although mathematics certainly involves normativity, it seems to fall short of resembling an institution in that Menary's cognitive norms are not 'social rules' (e.g., Dequech, 2009). Although related to cultural practices, cognitive norms for Menary involve enculturation and 'public systems of representation' (2018: 194) rather than primarily sociality and social interactions, which makes the alignment between Greif & Mokyr and Menary appealing but possibly incomplete. This is a gap that the concept of cognitive institution aims to fill, hinged on an alternative form of mind extension called 'socially extended mind' (SEM) (Gallagher, 2013; Gallagher and Crisafi, 2009). As part of the extended mind's third wave, SEM holds that the kind of mind extension reached through the enactive conception of social interactions and cognitive institutions is ubiquitous.

Although this paper is more about enactive normativity than cognitive institutions per se, in line with this section's clarificatory intent, some general remarks may be useful. Given the direction of the debate, the main point that deserves clarity is what it means for an agent's mind to be extended by an institution. To address this point, it is useful to remember that institutional mind extension was already

https://doi.org/10.1017/S174413742510026X Published online by Cambridge University Press

a possibility in the first wave, in Clark and to a lesser extent in North. Why would we need a SEM, then? Part of the answer, as noticed by second-wave scholars, is that the first wave requires parity, while what institutions do is not necessarily functionally equivalent to what individuals do in their heads. Another answer, more relevant for the point we want to make here, involves the 'cognitive bloat' (Rupert, 2004), namely, the objection that the extended mind hypothesis indiscriminately inflates (or includes too much in) the concept of mind. As a response, Clark (2018) shored up the parity principle with more restrictive criteria of mind extension: in addition to being functionally equivalent to internal resources, the external resource should also be reliably available and easily accessible when needed, and its content more or less automatically endorsed and not usually subject to critical scrutiny. If applied strictly, however, these additional criteria would rule out the very possibility of institutional mind extension: institutions are not easily accessible like Otto's pocket notebook, nor immune to critical scrutiny.

Third-wave theorists offer a different response to the cognitive bloat objection, which is also relevant for the post-Northian debate. According to third-wave scholars, the process of mind extension takes place only when individuals cognitively engage with tools, technologies, or institutions to make decisions or solve problems. This means, symmetrically, that the mind is *not* extended when individuals don't actively engage with the external resource. It is the active engagement and coupling with the cognitive institution that makes it a constituent part of an agent's cognitive processes (Gallagher, 2013).

This argument counts as a reply to the cognitive bloat objection in that it introduces an inherent limit to mind extension. In doing so, it also denies, as a corollary, that external resources – notebooks, smartphones, the entire realm of mathematics, social relations, institutions, and so on – can be considered 'cognitive' per se. External resources become cognitive only when individuals engage and create the right coupling with the resource. If the external resources were considered inherently cognitive, without anyone engaging with them, the cognitive bloat objection would present a difficulty. In contrast, institutional resources count as cognitive institutions only if and when they are enactively in use. This is a philosophical point with relevant implications for the study of cognitive institutions in economics.[2]

To solve the conundrum of how institutions can be studied if they are not an independent object of analysis but are constituted in social interactions, it is important to recognise that institutions and individual processes of mind extension are two sides of the same coin; we can't have one without the other. The domain of SEM includes all the mind-extending processes that individuals enact in the social domain; in other words, it is about agentive engagement with institutions and complex aggregates of institutions ('institutional entanglements', see Slaby and Gallagher, 2015), and about understanding social interactions and social practices as mind-extending 'affordances', namely, properties of the individual-institution coupling that afford specific patterns of social action (and make other patterns less affordable). Accordingly, cognitive institutions include those institutions whose properties emerge from the agents' engaged actions and interactions, and then loop back, recursively, to shape their ongoing actions and interactions. This formulation reflects the reciprocal links between agentive processes, social interactions, and institutional processes, and makes one set of processes dependent on the other. This clarification is an attempt to immunise against the temptation, often found in institutional theory, to detach collective entities like institutions from those processes that create their properties. In other words, the study of cognitive institutions is never the study of independent or abstract institutional systems.

[2]Some recent discussions of cognitive institutions in institutional economics may incur the risk of considering institutional resources as cognitive per se, i.e., independently of agents who enact them. Frolov (2023a: 177), for example, emphasises a systems perspective: '[In third-wave extended mind] [p]rimary attention is paid not to individuals (whose cognition is extended with the help of external minds, tools, and institutions, as was held in the first and second waves) but to socially and culturally distributed cognitive systems'. Dengsø and Kirchhoff (2023) also advocate a systems view. We disagree with this shift in focus, since there are always agents operating within institutional systems. If the agents were removed from such systems, the systems would collapse. What is the legal system, for example, without judges, clerks, lawyers, defendants, etc. Institutional systems are always made by and for agents who are socially interacting. Accordingly, we advocate a simultaneous focus on the *relational* aspects of individual agents and systems.

https://doi.org/10.1017/S174413742510026X Published online by Cambridge University Press

Where are the norms in cognitive institutions?

The previous section can be read in two ways: as a set of clarifications about the use (and possible overuse) of the extended mind concept in institutional economics, or as a summary of the (mostly implicit) answers to the 'where' question about norms in the institutional literature. Norms have been alternatively seen as concerted deliberations in individual heads (Searle, 2005), models shared by many heads (North), or environmental resources working either as substitutes (first wave) or complements (second wave) to in-the-head processes. In what follows, we provide our 'where' answer in the framework of cognitive institutions. To this end, we structure the discussion dialectically with Frolov (2023a; 2024), the only other approach in the literature to (implicitly) address the 'where' question for norms in a third-wave fashion.

Frolov: norms as combinations of content and form

In his articles, Frolov makes what we think is still a 'Northian' methodological choice, namely, he suggests retaining mental models as necessary constructs for explaining norms. 'I think it is better for institutional economists to change their understanding of mental models than (for now) to abandon them entirely' (2023a: 179). This choice of retaining mental models is based on a premise we contest; one which leads, if accepted, to what we see as unwanted consequences. The premise is that a 'shared mental *process*', advanced by Petracca and Gallagher (2020) as an alternative to North's 'shared mental model', is not able to address the fact that a norm is a contentful construct. To save norms' contentful status, Frolov's move is to supplement the process view with content understood as mental models. The possibly unwanted consequence of adopting this content *and* process view is that it may lead to a hard-to-manage and redundant ontology. Our goal in the next sections is to provide a more parsimonious analytical and ontological alternative.

We consider points of agreement first. Frolov (2023a: 180) provides a compelling definition of cognitive institutions as 'interactively and polycentrically co-produced cognitive norms (or rules)'. For the sake of synthesis, we take the liberty of translating this definition into a version that directly answers the *where* question. If a holder of this view were explicitly asked, 'where is the norm?' the answer would probably be: 'the norm is in the relation between individuals and institutions'. We take this to be the case given Frolov's third-wave commitment and his insistence on interactionism, with which we entirely agree. Disagreement starts to emerge when we try to pin down what exactly it means for norms to be ontologically located *in the relation*.

Frolov's approach 'combines rule[content]-based and process-based views' on the postulation that 'content [of the norm] refers to what is co-produced in shared mental processes; form [of the process] is how it is co-produced' (2023a: 176, 180). Mere content of the norm can't explain, on this view, how the norm has been formed; likewise, the mere formation process can't explain the norm's content, that's why we would need both. As we'll soon see, this content-form distinction may be too rough-edged; it doesn't consider that a process can be a self-sufficient explanatory concept.

The establishment of the content-form distinction leads Frolov, and even requires him to salvage the Northian concept of a mental model as a cornerstone of institutions:

> The model-based view is valuable in that it allows us to see not only the amorphous and dynamic side of socially extended cognitive processes, but it also enables us to take into account that these processes give rise to more or less distinct and stable collective [content-based] mental models. (Frolov, 2023a: 179–180)[3]

In this formulation, the existence of mental models is offered as the only explanation able to account for the existence of normative content in the world. If mental models didn't exist, the existence of content

[3]We note that there seems to be a conceptual leap from the assumption that a norm is a meaningful construct, with which we agree, to the Northian position that content can only come from a mental model. More on this below.

https://doi.org/10.1017/S174413742510026X Published online by Cambridge University Press

would be unexplainable.[4] On this assumption, we ask, where is a norm *really* located? Is it *really* in the relation? If so, is it *only* in the relation? Our sense is that, on Frolov's account, it would be more descriptively exact to say that the norm is *in many places*.

Let's consider a simple case: the rules of the road enforced by a traffic warden at the crossroads. In the act of directing the traffic, a traffic warden moves her arms and hands in conventional ways to stop or let cars proceed. Where is the rule according to the mental-model view? First of all, in the heads of car drivers who know they need to follow the (contentful) rule: 'abide by the traffic warden's conventional signs'. But the rule is also in the mind of the traffic warden, who not only learned the rule's content, along with the conventional signs and how to apply them, but was also trained to enforce the rules, knowing that road-users must follow her command. On this view, being 'in the mind' means that, as mental models, they are representations physically located in the head/brain. And this is not the only place where rules are located. As linguistic representational content, rules are also located in environmental artefacts like law books. There was probably a first time that a traffic code was written down, which included this specific content ('the traffic warden is in charge of regulating the traffic'), and since then, the same content has been variously formulated, translated into different languages, and stored in copies of law books. The number of places where this rule-as-content is located does not stop here, since it is to be found also in driving license manuals, news stories about rules being disregarded, causing an accident, as well as in the heads of the legislators who wrote the norm, and so on. For all these reasons, we tend to doubt that '[t]his view', as claimed, 'is one of minimal representationalism' (Frolov, 2023a: 179). On the contrary, there seems to be an *inflated overabundance* of mental models and representations in this account (see Petracca, 2022).

What about the process part of this approach? Frolov's content-form approach treats the norm-production process as form. A process enters the institutional picture to account for the dynamical, partial, transitory status of content. Mental models are not 'static mental representations', they 'are a dynamic part of dynamic interactions with the environment' (Frolov, 2023a: 181). Process, on this view, refers to the production or co-production of content by individuals within institutions. The metaphor Frolov offers for the production of cognitive institutions is that of a digital cloud.

> As in the case of shared files hosted on the cloud and edited collaboratively by many users, there is little point in internalizing ('downloading') cognitive institutions, since these institutions (like shared files) are different at each moment. [. . .] People always internalize partial, incomplete, and inaccurate 'copies' of cognitive institutions. (Ibid.)

Although the process of internalisation is described as itself a 'metaphor' (ibid.; see also Frolov, 2023b), it is consistent with the idea that institutions exist primarily as content. And if they exist as content, they need to be located somewhere (in the cloud, in single devices, or both);[5] or to put it differently, in philosophical jargon, if they are pure content, they need a 'vehicle'.[6] But vehicles seem to be everywhere: in the heads of all the actors involved, in traffic lights and other signs, in legal codes, etc. In what

[4]In a previous version of this paper, we wrote that Frolov views processes as inherently 'meaningless'. We agree with one reviewer that this wording may miss the clear integrative intent of Frolov's approach. We now use the more specific word 'contentless' (a process is defined as a *how*, not as a *what*).

[5]It is not straightforward to assess the ontological status of a view in which internalisation and storage are described as metaphors (along with the processes of 'uploading' and 'downloading'), but that at the same time rejects the very ideas of internalisation and storage.

[6]The extended mind hypothesis is typically discussed as a form of 'vehicle externalism' (Hurley, 1998). How does Frolov's content-form distinction relate to the content-vehicle distinction in extended mind discussions? One reviewer suggested that Frolov doesn't read the content-form distinction through the usual categories of the extended-mind debate, namely, the distinction between content and vehicles of content, but through the lens of dialectical philosophy, where a norm can be seen interchangeably as content and form. We have no clear evidence to tell whether or to what extent this interpretation is correct; we just notice that a reconciliation with the terms of the extended mind debate and the formulation of a clear ontology would be needed to avoid possible misunderstandings. We have nothing in principle against dialectical philosophy, and in fact appreciate it as a fruitful philosophy of *praxis* compatible with enactivism (e.g., Di Paolo and Potapov, 2024).

https://doi.org/10.1017/S174413742510026X Published online by Cambridge University Press

follows, we will argue that this institutional ontology is 'inflationary'. We propose a different ontology for cognitive institutions that dispenses with the multiple representations involved in mental models and identifies practice itself as intrinsically meaningful and normative.

Norms in the process or practice

Although in theory mental models and processes are claimed to be equally important (Frolov 2023a: 179), mental models have a sort of priority in the actual workings of Frolov's account. In the example of the traffic crossroads (the example is ours), all the causal power seems to be in mental models: they lead the traffic warden to raise her arm and hand in one way or another, and once they get reactivated in the car driver's head, they lead the car driver to stop or proceed. Even a driver's decision not to stop at the traffic warden's command would reference a mental model (about the costs and benefits of breaking the rule). On the content view, mental models have a *causal* priority over processes. There seems to be no moment in which a process can change a mental model without being initiated by a previous mental model. Another way to say this is that norms don't seem to be *actually* located in the relation (rather, they seem to be located in minds as causal 'effectors' and in the process as causally 'effected'). In this section, we want to provide a view more consistent with the claim that norms are *in the relation*.

SEM emphasises processes and practices. Agents engage in cultural practices, habits, customs, skills based on know-how, and the like. Frolov objects, as indicated, that processes and practices are different from content, and for this reason, are unable to constitute a norm that has (or even *is*) content.

> Cognitive institutions (rules or norms) result in a wide range of practices, but *practices themselves are not rules or norms*. Practices are specific patterns of norm-following behaviors; a cognitive norm can drive many behaviors and manifest in various practices. (2023a: 180, emphasis added)

Priority is here again explicitly assigned to mental models, in the form of rules or norms, as they are said to 'result in' practices and 'drive' behaviours. Practices follow norms, are not themselves norms, nor are they constitutive of a norm's content. This is the main difference from our SEM approach. We argue that to be genuinely in the relation, a norm has to be enacted in the practice, and does not exist outside of a set of practices that instantiate it. The practice is in a constitutive relationship with the norm, so that the norm itself is in-the-practice – a *rule-in-action*.[7]

A practice, we argue, can be meaningful ('contentful') per se. Let's return to the crossroads. At the crossroads, there is no cloud (or cloud-like institution) where people dynamically upload and download copies of mental models, but there is the enaction of coordinated practices. From a very young age, people are exposed to the kind of normativity that Slors (2020) calls 'symbiosis'. What one does in the social system strictly depends on what other people in the system do, that is, it depends on their reciprocal roles and how they interact. What a driver does is dependent on what a traffic warden does, and this form of normativity doesn't need to be stored in any code; rather, in the specific situation, it is hermeneutically meaningful to all the actors involved. The process of socialisation we are all exposed to, which is also a process of enculturation, enacts normativity such that individuals typically don't need to represent to themselves a rule that at a crossroads they must follow the traffic warden's indications. Nor do drivers need any special learning to stop at a crossroad when the traffic warden shows them a raised arm and a culturally embedded stop gesture.[8]

[7]Philosophically, our argument is in line with Lambros Malafouris's (2013) criticism of Searle's concept of intentionality. For Searle, an intention (a rule, in this case) being 'in action' means that it originates in a previous mental model (a prior intention), such that the action is a vehicle of that mental model. This is a 'rule in action' view that we think aligns with Frolov's and other representational approaches. We suggest changing it into a 'rule-in-action', where the hyphenation means that rules do not originate in previous mental models but are truly immanent in actions.

[8]This is a Wittgensteinian point as clearly explained by Charles Taylor: 'A rule that exists only in the practices it animates, and does not require and may not have any express formulation. How can this be? Only through our embodied understanding'. This also relates to Pierre Bourdieu's concept of 'habitus', which, Taylor explains, 'is "a system of durable and

https://doi.org/10.1017/S174413742510026X Published online by Cambridge University Press

In this respect, it is important to emphasise that the rule is not in the traffic warden's gesture but in the coordinated activity of the system comprising the traffic warden, car drivers, bikers, skaters, pedestrians and so on approaching that crossroads. Such practices become sedimented in the history of the norm's enaction by all these subjects. In this sense, the norm is *relational*. At the crossroads, a concerted (re)*enaction* of rules and norms takes place. This leads through repeated interactions and sedimentations to relatively stable formations like habits and routines that don't need to be in addition represented in mental models (Gallagher, 2025). This processual and practice-based enaction is more essential, we argue, than some ideal enaction of a pure contentful rule. Consider, for example, a situation at the crossroads where there is no traffic warden, but, having come upon some children in the road, a passing pedestrian holds up her hand to stop traffic. There is no rule prescribing to stop at that command; yet most of us would stop. We could, of course, formulate some ad hoc rules to justify everyone's actions. Generally, however, there is no need for such rules (which would be difficult to formulate on the spot in any case). Things happen in a way that follows a set of practised dispositions to interact – to see the others, to stretch out a hand, to stop and ask what's going on, etc. – practices that end up having a normative effect. These normative practices may then sediment, becoming institutions.

There are multiple explanations for why we stop at the traffic warden's command that do not appeal to contentful rules. One is that, as proponents of institutions-as-equilibria would say (e.g., Aoki, 2001), stopping for the traffic warden is not strictly a rule but a self-enforcing equilibrium that could be costly to break, as the risk of an accident would rise dramatically. Another explanation is that once the users of the road have become practised in their behaviour as road users, the rules of the road are no longer explicitly invoked; they become embodied. Hubert Dreyfus (2006) draws a vivid analogy between rules and training wheels that get removed at some point in the process of learning to ride a bicycle. Likewise, experienced drivers don't need to activate a mental model when they approach a crossroads.

The metaphor of training wheels leaves open the possibility that norms are established as content before dissolving into practice. We claim, instead, that rules get primarily established (enacted) in practice; practice comes first. Once established, rules may in some cases become formalised, which can be useful in a learning stage (although in developmental contexts learning is typically practical and experiential), before being removed (like training wheels) in everyday practice. The need for crossroads regulation probably emerged as a very practical problem as vehicles became faster and crashes piled up; the solution arguably came first in the form of self-organised normative arrangements before transitioning into law books and then disappearing again in everyday car drivers' practice. In the context of the legal system, this is a complex issue involving codified legal institutions as manifestations of specific cultures and their practices (e.g., Berman, 1983).

The crossroad example has its limitations. Even people convinced by the example may argue that what happens at a crossroads cannot be extended to entire institutions, and even less to rule-intensive institutional contexts like a financial market or a parliament. We disagree. It's important to be clear that here we are not denying that some institutions are more intensely regulated than others. Rather, our claim is that practice is a thick enough concept to be explanatorily self-sufficient across a variety of contexts; practice is not an empty box that needs to be animated or filled by mental content. Practices are pieces of continually enacted meaning that form non-representational cornerstones of cognitive institutions.

Non-representationalism in the study of social phenomena is an increasingly available option for situations that prima facie appear to be 'representation-hungry' (e.g., Hutto and Myin, 2017; Gallagher, 2020a; Petracca, 2023). One of its main epistemic merits, parsimony, avoids the *inflation problem* in third-wave representationalism, namely, the multiplication of loci for contents and processes, and the necessary activation of content simultaneously in multiple locations. In the SEM approach to what a rule or a norm is, and how it works, the focus on the *relational aspects of practice*, its being continually enacted, permits a more parsimonious (deflationary) dynamical explanation (Zednik, 2011).

transposable dispositions", dispositions to bodily comportment, to act, to hold oneself, or to gesture in a certain way. A bodily disposition is a habitus when it encodes a certain cultural understanding' (Taylor, 1995: 178; see Gallagher, 2020b for discussion).

https://doi.org/10.1017/S174413742510026X Published online by Cambridge University Press

Our proposal also enables us to address the problem of norm stability. Frolov introduces mental models as an attempt to capture the fact that some norms and institutions are fairly stable over time, under the assumption that the process-view would be unable to account for such stability. We have already motivated our reservations about this assumption, and we add that, at least since Thorstein Veblen (1922[1899]; see Hodgson, 2004a; 2004b), the process-based concept of habit has been invoked to explain stability and regularity in institutional economics. Moreover, and in contrast, the mental-model view depicts content as being itself quite unstable, thus undermining the main virtue for which it was introduced in the first place. This instability is rendered clear through the cloud metaphor.

> Co-production most resembles collaborative document editing: the document is stored in the cloud, and a lot of people have real-time access to it and can simultaneously make edits, additions, and comments. Likewise, cognitive institutions are (to varying degrees) open to 'co-editing'; they are social constructs that people co-produce on a conversation-by-conversation basis by participating in personal and digital communications. (Frolov, 2024: 4–5).

From such continuous editing, uploading and downloading, content does not remain the same, and the issue is which version is relevant. '[Cognitive institutions] are constantly changing, and by connecting to them, people "update" their personal mental models to new versions of cognitive institutions' (Frolov, 2023a: 181). Indeed, this continuous running after the most updated content may suggest that the distinction between content and process itself becomes blurred. As Frolov says,

> The choice of whether to describe a cognitive institution as a 'model' or a 'process' depends on the scholar's subjective viewpoint; even a stone can be considered as an extremely slow process of formation and deterioration. (Frolov, 2023a: 179)

We think that focusing on norms as meaningful processes (practices) may be both ontologically and explanatorily frugal. It removes the uneasy feeling of having to choose each time between a content and a process description.

Market rules as practices

At this point, it will be helpful to discuss an economic example. The market, a staple of post-Northian analyses, is a good one (Gallagher *et al.,* 2019; Petracca and Gallagher, 2020; Dekker, 2022; Kuchař, 2025). This section briefly discusses how market rules can be considered meaningful practices without invoking the idea of mental models. The view of market rules as practices is not new. It can be traced back to the discussion of routines (Nelson and Winter, 1982) and habits (see Hodgson, 2004a). More specifically, market rules have been conceived as practices in the work of Callon (1988) and Bourdieu (2005). This section aims to connect this literature with that of cognitive institutions.

In a famous computer simulation, Gode and Sunder (1993) show that the rules of a double auction can lead to high levels of market efficiency even when the market is populated with cognitively unsophisticated agents, called 'zero-intelligence' traders. Since markets continue to work efficiently in the absence of agents' mental models, this result has been taken to mean that external market rules can almost entirely replace mental models. North and Clark are the main champions of this view, since they see market rules respectively as externalisations of and substitutes for mental models. In the context of Gode and Sunder's analysis, though, it is not clear what benefit ensues from talking about mental models at all: externalisation and replacement of mental models need to be invoked only if mental models are assumed in the first place. North and Clark's view that the rules of double auction are pure content is clearly redundant, since zero-intelligence traders experience market rules as behavioural constraints to their action, not as pieces of content (qua unsophisticated automata, their behaviour consists more of reactions than information processing).

https://doi.org/10.1017/S174413742510026X Published online by Cambridge University Press

It would be unwarranted, though, to establish our view of market rules as meaningful practices based exclusively on a post-Northian interpretation of Gode and Sunder's example: market agents are never completely cognitively unsophisticated. On the other hand, the lesson from the various strands of bounded rationality is that agents are not as sophisticated as economic theory would assume or prescribe. Market participants demonstrate, it seems fair to say, a broad spectrum of abilities and success in the use of markets. The crucial question remains whether mental models in market participants' heads or as offloaded content onto market rules are necessary for demonstrating ability and success in markets.

In this context, we can understand rules not as content but as practices stemming from practical needs. The role of an auctioneer is analogous to that of a traffic warden, and the role of a trader is analogous to that of a car driver: like traffic regulations, auction rules emerged out of the very practical need of regulating transactions, not out of fiat deliberation. Moreover, it would be interesting to know how well stock market traders know the details of a double auction. It's not far-fetched to say that most people engage with the stock market without knowing in detail its rules. They experience the double auction mostly as a set of consequences to their actions, either as askers or bidders, which is another way to say that they experience the market as a set of reciprocal practices. Agents may read textbooks on auction theory, hoping thus to increase their understanding or improve their trading performance, but it is doubtful whether this is essential for success; either way, not much pure content remains after traders become expert practitioners. Most of what goes on in markets is describable in terms of intersecting practices, namely, agents' practices continually colliding with those of other agents populating the market.

Bourdieu has described the market as animated by 'the logic of the spontaneous orchestration of practices' (2005: 73). Spontaneous orchestration is increasingly visible as we distance ourselves from highly regulated settings (like auctions) and consider cases more similar to local markets or bazars. It would be an error, however, to think that these more informal settings lack norms: market students are only more prone to accept that the norms operating in those settings are called practices. The practice of handshaking is, in certain contexts, no different than a binding rule, and is commonly referred to as a practice. Our point, in brief, is that no ontological difference intervenes between bidding in a double auction and handshaking a business partner.

Normativity and rationality: another location problem

So far, we have dealt with the nature of norms and their location in the context of cognitive institutions. In this section, we discuss a topic closely related to norms, in fact, an entire branch of normativity: rationality. Rationality is guided by and assessed through norms, which we'll call here *criteria* or *principles* to distinguish them from social norms discussed in the previous sections. Based on the criteria and principles of rationality currently available in the literature, we can ask what it means for an agent to behave rationally in a cognitive institution. But we can also ask, perhaps more urgently, whether the available criteria and principles of rationality are adequate to guide and assess rationality in cognitive institutions.

For this paper, this is a relevant point because it addresses a related 'where' or location question, this time involving criteria and principles of rationality. Since the advent of bounded/ecological rationality, the criteria for guiding and assessing rationality have included the environment, such that in a (methodological, not ontological) sense, the environment has been considered *part of* rationality (see Petracca, 2017). As seen in Section *Other (mis)alignments?*, the criterion of rationality that ensues from including the environment in the rationality picture is a principle of *adaptation*: '[a] heuristic is ecologically rational to the degree that it is *adapted* to the structure of the environment' (Gigerenzer and Todd, 1999: 13, emphasis added). How would this criterion work in an institutional setting? And would it be suitable for cognitive institutions?

Frolov's work also plays an important role in this discussion since he deals with rationality within the framework of cognitive institutions. For this reason, our argument will shape up once again

https://doi.org/10.1017/S174413742510026X Published online by Cambridge University Press

dialectically with his. Frolov (2023a) presents his contribution in continuity with ecological rationality, which, as recognised by Dekker and Remic (2019), consists of two research strands: Gigerenzer's FFHs research program (Gigerenzer and Todd, 1999) and Vernon Smith's (2007) market institutionalism. Frolov aims to introduce a third, alternative version of ecological rationality, 'ER3', by integrating elements of polycentrism and co-production. Frolov (2024) calls his proposal 'enactive rationality', although he seems to distance this idea from the enactivist version of 'radical embodied rationality' that dispenses with mental models (Gallagher, 2018; Petracca, 2021; Petracca and Grayot, 2023; 2025; Rolla, 2021).

Following Rolla (2021), Frolov accepts Gigerenzer's adaptationist principle of rationality, namely, that rationality has to be framed in terms of adaptation to the environment. What differs in his proposal is the attempt to overcome Gigerenzer's exclusive focus on heuristics as sources of adaptive rationality. Institutions, Frolov claims, can also be considered sources of adaptive rationality (Frolov, 2023a: 183). ER3 or enactive rationality is, therefore, a version of ecological rationality that includes cognitive institutions as resources for adaptive rationality. As such, it leads to a revised adaptive rationality principle: '[i]ndividual decisions are ecologically rational if they are good enough or better than alternatives through the use of *cognitive institutions adapted to a given environment*' (2023a: 185, emphasis added). There is an annotation to this reformulation that we agree with: that '[o]ur cognitive activities are aimed not so much at passive adaptation to an environment with predetermined and fixed signals but rather at active exploration, redefinition, rethinking, probing, and testing the environment in search of cues and clues provided by cognitive institutions of varying degrees of complexity' (ibid). Adaptation is not only static or passive but also active. Yet, when it comes to summarising the main 'cognitive problem' which the research on cognitive institutions is supposed to address, the rationality principle is said to concern a property that is in fact static and passive, that is, '[t]he *fitness* of cognitive institutions to their environments' (Frolov, 2024: 13, emphasis added). Analogous to how heuristics need to match the structure of the environment, cognitive institutions, to lead to rational judgements and decisions, also need to match the structure of the environment.

Before offering some comments on this rationality proposal and briefly advancing our alternative, we note that extending the ecological framework from heuristics to institutions is not a trivial project. In Gigerenzer's framework, heuristics are individuals' internal (or sometimes embodied) resources whose rationality comes from their *direct* fitness with the environment; there is no mediation between the agent's cognitive resources and the environment. This is, however, not what happens in the case of institutions: individuals in Frolov's model need first to couple with external resources as complex as institutions before being able to make rational judgements and decisions in a given environment. In this case, the entity whose adaptation or fitness is assessed is not an internal resource, as in the case of heuristics, but a *new composite entity* constituted by the coupling between individuals and institutions. In other words, it is the coupling of *individual-plus-institution* that is assessed in terms of ecological rationality in any particular environment. This is more than a detail in the economy of this view, to the point that Frolov's proposal seems not purely ecological but more about the means-end/functionalist rationality of couplings, namely, an 'extended' (more so than enactive) form of rationality (Clark, 2001; see Petracca, 2021; Petracca and Grayot, 2023).

With this clarification in mind, we can now consider the specific location problem concerning the rationality of institutional mind extension. If we understand institutions as environmental resources with which we couple to make good judgments and decisions, a location problem arises in the following terms: *should institutions be considered part of the coupling or alternatively part of the environment against which the coupling is assessed?* (Another option is that institutions are part of both, which makes the location problem even harder to solve.) In the literature on the extended mind, when agents couple with tools and artefacts in the environment, like Otto with his notebook, the external resource (the notebook) can be neatly separated from the environment, such that it is conceptually straightforward to assess the rationality of the coupling (Otto plus the notebook) in that environment (defined by the memory task). But when we assess the rationality of an agent's coupling with an institution, the possibility of distinguishing the institution from the environment seems far more problematic. This is a

https://doi.org/10.1017/S174413742510026X Published online by Cambridge University Press

crucial issue since the ecological/adaptive framework stands on the assumption that we can neatly mark off the resource from the environment against which the resource's rationality is assessed. If this is not possible, the entire ecological framework is in question. In other words, what Arnau et al. (2014) call a 'methodological environment' needs to be neatly distinguished from the coupling to play its normative role.

First, consider a case where the distinction between the institutions we couple with and the methodological environment seems reasonably clear-cut. Harper (2010) makes the case that number systems (Arabic, Roman, . . . , binary, decimal, etc.) can be seen as cognitive and social technologies. They could therefore count as cognitive institutions. Building on this, Whitman (2025) suggests that certain number systems are better suited than others to solve certain society-relevant cognitive problems, making the convincing case that different number systems are suitable for different tasks. This idea seems overall consistent with Masahiko Aoki's (2011) institutions as 'cognitive media' (*media* because they somewhat interpose themselves between agents and the 'outer' environment). In this case, the use of an ecological framework where the institutional resource we couple with (i.e., the number system) is distinguished from the environment (a measuring or counting task) seems well supported by an adaptive story.

But what about other cognitive institutions, such as the legal system? Is there a clear-cut boundary between the legal system and the environment? Cognitive institutions are sometimes *so spread into the world*, and so intertwined with it and other institutions, to be massively porous, permeable, and hard to disentangle from the environment. One can couple with the legal system to solve a problem that is economic (e.g., a troubled relationship with a business partner) but also *legal* (i.e., it may involve the resolution of a contract), and perhaps moral: the legal system as a cognitive institution is embedded in the social, material, and moral environment. There is no clear-cut way to assess our coupling with the legal system in terms of adaptation because the legal system seems to be, at once, constitutive of the coupling *and* of the environment against which the coupling is to be assessed. Part of the problem relates to how broadly an institution is defined: narrow definitions of institutions, like in the example of number systems, are easier to reconcile with adaptive stories. But this is not obtainable in general. Even if we defined the legal system so narrowly to reduce it to the norms in a law book, in the very same law book we would encounter norms that individuals use to solve problems (e.g., the resolution of a business contract) and others that represent the more complex normative environment (e.g., constitutional, supra-national rules) or otherwise limit our legal possibilities (the resolution of a business contract can never entail the violation, for instance, of human rights). In all these cases, the distinction between institutional resources and environment is not possible, conceptually and locationally.

In addition, cognitive institutions present another similar but not identical 'where' issue. In contrast to what happens when an individual couples with other, less complex tools or external resources (e.g., notebooks), the individual in the case of institutional mind-extension is not outside of the cognitive institution, she is (ontologically) *part of it*, insofar as she enacts the institution in her practices and social interactions and is in a reciprocal/dynamical co-constructive relation with it. This is a further indication that institutional mind extension requires an entirely new set of location considerations. Figure 1 represents the two locational issues concerning rationality in cognitive institutions.

The ecological approach takes it for granted that the rationality associated with either the use of heuristics or cognitive institutions lies in their *adaptation* to the environment. If a cognitive institution, however, can't always be neatly separated from the environment, the principle of adaptation seems not to be generally applicable, and we need different criteria for assessing rationality (see Petracca, 2021; Petracca and Grayot, 2023). To this aim, we suggest taking the idea of *co-production* seriously: if we understand the production of cognitive institutions as a genuinely joint activity, that is, the idea that cognitive institutions are enacted in social interactive practices, co-production can hardly be reduced to either a matter of adaptation or of mere active construction by each single producer (Petracca, 2021). Accordingly, the idea of rationality in a co-production setting needs to depart from the classical means-end framework of individualistic forms of rationality (Petracca and Grayot, 2023; 2025). In the context

https://doi.org/10.1017/S174413742510026X Published online by Cambridge University Press

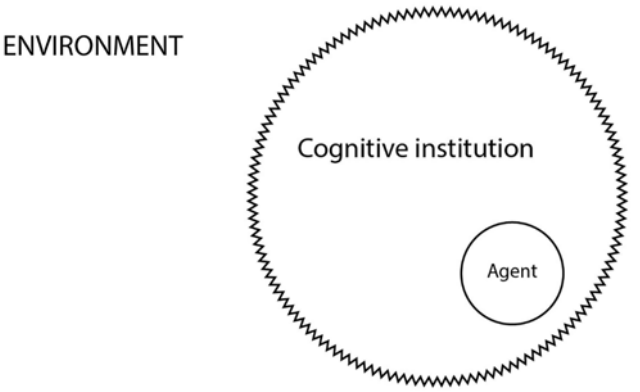

Figure 1. *Representation of two location issues concerning rationality in cognitive institutions*. (i) The boundary between cognitive institution and environment is porous and unstable (zig-zag line), making the adaptive rationality principle difficult to apply. (ii) Differently from when they couple with other tools, individuals are usually part of the cognitive institution they couple with. (The figure doesn't otherwise describe the agent-institution-environment coupling, which is inherently dynamic).

of the social practices that enact cognitive institutions, the equilibrium tradition in new institutional economics (e.g., Aoki, 2001) can provide useful insights. As co-producers, agents need to reach mutual coordination not for the sake of finding an 'equilibrium', but to achieve effective co-production and symbiotic mind-extension. Closing the circle with the arguments of the previous sections, it might be argued, in line with proponents of the rules-in-equilibria view (Hindriks and Guala, 2015), that coordination revolves around rules, and we would agree with that if it is also specified that coordination is of and about practices, not about contentful rules located in the head. Accordingly, instead of discussing coordination as a defining principle of institutions à la Aoki, we argue that it can represent an alternative to adaptation as a normative principle of rationality. It's not that adaptation is not at all usable in institutional settings; it's rather that any time the environment can't be told apart from the institution we couple with – which we take to happen often in cognitive institutions – the rationality game is more one of coordinating institutional resources with other co-producers than just adapting the institution to the environment. Of course, we are still far from resolving the conundrum of rationality in cognitive institutions; this section has just raised a conceptual issue and emphasised how it depends on an ontological, and not just methodological, location problem.

Concluding remarks

Along with the growing interest in post-Northian institutional economics, even outside the precinct of economics (e.g., Johnson *et al.,* 2024; Werner, 2024), the need becomes more pressing to figure out what we should retain of North's original framework, and what we shouldn't. North's contribution is typically said to lie in two parts, one social, the definition of institutions as norms and rules, and one cognitive, namely, that norms and rules are emanations of shared mental models. In this paper, we've argued that retaining the centrality of norms, and more generally the focus on normativity, is more fruitful than retaining the assumption that those norms or rules are mental models.

A post-Northian institutional economics focused on *both* norms and mental models seems too continuous with North to fit the 'post' category. Although an interactive and dynamical view of mental models is certainly welcome, it can hardly be said to be post-Northian since even Denzau and North (1994: 4) themselves claimed to be interested in 'the features and dynamics of mental models'. We've framed our alternative in terms of norm *location*, a question that we've liberally associated with the location problem in social ontology (Hindriks, 2013; Beck, 2024). Our tentative answer is that instead of being located in the head, in other vehicles, or in the interaction between content and vehicles, norms

https://doi.org/10.1017/S174413742510026X Published online by Cambridge University Press

are better seen as lying in the relation between agents and institutional resources, namely in agentive-institutional *practices*. This requires a twofold reconceptualisation: first, understanding processes as meaningful practices rather than empty or amorphous vessels, and second, understanding practices themselves not as consequences of mental models but as autonomous constructs provided with enacted meaning.

Among the benefits of our proposal is ontological and conceptual manageability. Since there seems to be general agreement that post-Northian institutional economics should be about interactions, processes and practices, an additional focus on mental models may be analytically and ontologically redundant, or, as we've called it, inflationary. Accordingly, we advocate a post-Northian institutional economics based on the more enactive third wave of SEM, involving norms and other mind-extending resources established on the self-sufficient concepts of interaction, shared meaningful processes, and practices.

Acknowledgements. We thank the seven anonymous reviewers for their stimulating and constructive comments. We are also grateful to the participants in the first '4E Economics' workshop (University of Memphis, November 2024) for their valuable feedback.

References

Aoki M. (2001). *Toward a Comparative Institutional Analysis*. Cambridge, MA: MIT Press.
Aoki M. (2011). Institutions as cognitive media between strategic interactions and individual beliefs. *Journal of Economic Behavior and Organization* **79**, 20–34.
Arnau E., Ayala S. and Sturm T. (2014). Cognitive externalism meets bounded rationality. *Philosophical Psychology* **27**, 50–64.
Beck L. (2024). Why we need to talk about preferences: economic experiments and the where-question. *Erkenntnis* **89**, 1435–1455.
Berman H.J. (1983). *Law and Revolution: The Formation of the Western Legal Tradition*. Cambridge, MA: Harvard University Press.
Bourdieu P. (2005). *The Social Structures of the Economy*. Cambridge, UK: Polity.
Callon M. (Ed.) (1998). *The Laws of the Markets*. Oxford: Blackwell Publishers.
Cash M. (2013). Cognition without borders: "Third wave" socially distributed cognition and relational autonomy. *Cognitive Systems Research* **25–26**, 61–71.
Clark A. (1997). Economic reason: the interplay of individual learning and external structure. In Drobak J. and Nye J. (eds), *The Frontiers of the New Institutional Economics*. San Diego, CA: Academic Press, pp. 269–290.
Clark A. (1998). *Being There: Putting Brain, Body, and World Together Again*. Cambridge, MA: MIT Press.
Clark A. (2001). Reasons, robots and the extended mind. *Mind & Language* **16**, 121–145.
Clark A. (2018). *Supersizing the Mind: Embodiment, Action, and Cognitive Extension*. New York: Oxford University Press.
Clark A. and Chalmers D. (1998). The extended mind. *Analysis* **58**, 7–19.
Craver C. (2007). *Explaining the Brain*. New York: Oxford University Press.
Dekker E. (2022). How cognitive institutions and interpretative rationality enable markets with infinite variety. In D'Amico D.J. and Martin A.G. (eds), *Contemporary Methods and Austrian Economics*. Leeds: Emerald Publishing, pp. 151–167.
Dekker E. and Remic B. (2019). Two types of ecological rationality: or how to best combine psychology and economics. *Journal of Economic Methodology* **26**, 291–306.
Dekker E. and Remic B. (2024). Hayek's extended mind: on the (im)possibility of Austrian behavioural economics. *Journal of Institutional Economics* **20**, e19.
Dengsø M.J. and Kirchhoff M.D. (2023). Beyond individual-centred 4E cognition: systems biology and sympoiesis. *Constructivist Foundations* **18**, 351–364.
Denzau A.T. and North D.C. (1994). Shared mental models: ideologies and institutions. *Kyklos* **47**, 3–31.
Dequech D. (2009). Institutions, social norms, and decision-theoretic norms. *Journal of Economic Behavior & Organization* **72**, 70–78.
Di Paolo E.A. and Potapov K. (2024). Evald Ilyenkov and the enactive approach. *Studies in East European Thought* **76**, 439–463.
Dreyfus H.L. (2006). Overcoming the myth of the mental. *Topoi* **25**, 43–49.
Frolov D. (2023a). Post-Northian institutional economics: a research agenda for cognitive institutions. *Journal of Institutional Economics* **19**, 175–191.
Frolov D. (2023b). *Digital Capitalism and New Institutionalism*. Abingdon and New York: Taylor & Francis.
Frolov D. (2024). The economics of cognitive institutions: mapping debates, looking ahead. *Journal of Institutional Economics* **20**, e28.

Gallagher S. (2013). The socially extended mind. *Cognitive Systems Research* **25–26**, 4–12.
Gallagher S. (2017). *Enactivist Interventions: Rethinking the Mind*. New York: Oxford University Press.
Gallagher S. (2018). Embodied rationality. In Bronner G. and Di Iorio F. (eds), *The Mystery of Rationality. Mind, Beliefs and Social Science*. Cham: Springer, pp. 83–94.
Gallagher S. (2020a). *Action and Interaction*. Oxford: Oxford University Press.
Gallagher S. (2020b). To follow a rule: lessons from baby logic. In Weinstock D. and Levy J. (eds), *Interpreting Modernity: Essays in Honor of Charles Taylor*. Montral: McGill-Queen's University Press, pp. 21–34.
Gallagher S. (2023). *Embodied and Enactive Approaches to Cognition*. New York: Cambridge University Press.
Gallagher S. (2025). Habit, sedimentation and institutions. *Cogent Arts & Humanities* **12**, 2480879.
Gallagher S. and Crisafi A. (2009). Mental institutions. *Topoi* **28**, 45–51.
Gallagher S., Mastrogiorgio A. and Petracca E. (2019). Economic reasoning and interaction in socially extended market institutions. *Frontiers in Psychology* **10**, 1856.
Gallagher S. and Petracca E. (2024). Trust as the glue of cognitive institutions. *Philosophical Psychology* **37**, 216-239.
Gigerenzer G. and Todd P.M. (Eds.) (1999). *Simple Heuristics That Make Us Smart*. New York: Oxford University Press.
Gode D.K. and Sunder S. (1993). Allocative efficiency of markets with zero-intelligence traders: market as a partial substitute for individual rationality. *Journal of Political Economy* **101**, 119–137.
Greif A. and Mokyr J. (2017). Cognitive rules, institutions, and economic growth: Douglass North and beyond. *Journal of Institutional Economics* **13**, 25–52.
Harper D.A. (2010). Numbers as a cognitive and social technology: on the nature of conventional number sequences used in economic systems. *Journal of Institutional Economics* **6**, 167–190.
Hertwig R. and Grüne-Yanoff T. (2017). Nudging and boosting: steering or empowering good decisions. *Perspectives on Psychological Science* **12**, 973–986.
Hindriks F. (2013). The location problem in social ontology. *Synthese* **190**, 413–437.
Hindriks F. and Guala F. (2015). Institutions, rules, and equilibria: a unified theory. *Journal of Institutional Economics* **11**, 459–480.
Hodgson G.M. (2004a). *The Evolution of Institutional Economics: Agency, Structure and Darwinism in American Institutionalism*. London: Routledge.
Hodgson G.M. (2004b). Reclaiming habit for institutional economics. *Journal of Economic Psychology* **25**, 651–660.
Hodgson G.M. (2006). What are institutions?. *Journal of Economic Issues* **40**, 1–25.
Hurley S.L. (1998). Vehicles, contents, conceptual structure, and externalism. *Analysis* **58**, 1–6.
Hutto D.D. and Myin E. (2017). *Evolving Enactivism: Basic Minds Meet Content*. Cambridge, MA: MIT Press.
Johnson S.G., Schotanus P.R. and Kelso J.S. (2024). Minds and markets as complex systems: an emerging approach to cognitive economics. *Trends in Cognitive Sciences* **28**, 1037–1050.
Kaplan D.M. (2012). How to demarcate the boundaries of cognition. *Biology and Philosophy* **27**, 545–570.
Kirchhoff M.D. (2012). Extended cognition and fixed properties: steps to a third-wave version of extended cognition. *Phenomenology and the Cognitive Sciences* **11**, 287–308.
Kirchhoff M.D. (2017). From mutual manipulation to cognitive extension: challenges and implications. *Phenomenology and the Cognitive Sciences* **16**, 863–878.
Kuchař, P. (2025). Competition as socially extended cognition: a Hayekian perspective on market knowledge and institutional evolution. *The Review of Austrian Economics*, 1–23. https://doi.org/10.1007/s11138-025-00695-1
Malafouris L. (2013). *How Things Shape the Mind*. Cambridge: MA: MIT Press.
Mantzavinos C. (2001). *Individuals, Institutions and Markets*. Cambridge, UK: Cambridge University Press.
Mantzavinos C., North D.C. and Shariq S. (2004). Learning, institutions, and economic performance. *Perspectives on Politics* **2**, 75–84.
Menary R. (2007). *Cognitive Integration: Mind and Cognition Unbounded*. Basingstoke: Palgrave Macmillan.
Menary R. (2018). Cognitive integration: how culture transforms us and extends our cognitive capabilities. In Newen A., de Bruin L. and Gallagher S. (eds), *Oxford Handbook of 4E Cognition*. Oxford: Oxford University Press, pp. 187–215.
Menary R. (2010). Cognitive integration and the extended mind. In Menary R. (ed), *The Extended Mind*. Cambridge, MA: MIT Press, pp. 227–243.
Nelson R.R. and Winter S.G. (1982). *An Evolutionary Theory of Economic Change*. Cambridge, MA: Belknap Press.
North D.C. (1990). *Institutions, Institutional Change and Economic Performance*. New York: Cambridge University Press.
North D.C. (2005). *Understanding the Process of Economic Change*. Princeton, NJ: Princeton University Press.
Petit E. and Ballet J. (2025). Bringing emotions into post-Northian institutional economics: a reading inspired by John Dewey. *Journal of Institutional Economics* **21**, e20.
Petracca E. (2017). A cognition paradigm clash: Simon, situated cognition and the interpretation of bounded rationality. *Journal of Economic Methodology* **24**, 20–40.
Petracca E. (2021). Embodying bounded rationality: from embodied bounded rationality to embodied rationality. *Frontiers in Psychology* **12**, 710607.
Petracca E. (2022). Two styles of social neuroeconomics. *OEconomia. History, Methodology, Philosophy* **12**, 95–113.

Petracca E. (2023). Social phenomena as a challenge to the scaling-up problem. *Phenomenology and the Cognitive Sciences*, 1–19.
Petracca E. and Gallagher S. (2020). Economic cognitive institutions. *Journal of Institutional Economics* **16**, 747–765.
Petracca E. and Gallagher S. (2025). Trust and reliance in the cognitive institutions of cryptocurrency. *Mind & Society* **24**, 125–144.
Petracca E. and Grayot J. (2023). How can embodied cognition naturalize bounded rationality?. *Synthese* **201**, 115.
Petracca E. and Grayot J. (2025). Mindshaping and the embodiment of rationality. In Zawidzki T.W. and Tison R. (eds), *The Routledge Handbook of Mindshaping*. Abingdon and New York: Routledge, pp. 212–224.
Remic, B. (2021). Environment as a resource, not a constraint. *Journal of Contextual Economics–Schmollers Jahrbuch* **1–2**, 85–107.
Rolla G. (2021). Reconceiving rationality: situating rationality into radically enactive cognition. *Synthese* **198**, 571–590.
Rupert R.D. (2004). Challenges to the hypothesis of extended cognition. *The Journal of Philosophy* **101**, 389–428.
Searle J.R. (2005). What is an institution?. *Journal of Institutional Economics* **1**, 1–22.
Slaby J. and Gallagher S. (2015). Critical neuroscience and socially extended minds. *Theory, Culture & Society* **32**, 33–59.
Slors M. (2020). From notebooks to institutions: the case for symbiotic cognition. *Frontiers in Psychology* **11**, 674.
Smith V.L. (2007). *Rationality in Economics: Constructivist and Ecological Forms*. Cambridge, UK: Cambridge University Press.
Sutton J. (2010). Exograms, interdisciplinarity and the cognitive life of things. In Menary R. (ed), *The Extended Mind*. Cambridge, MA: MIT Press, pp. 189–225.
Taylor C. (1995). To follow a rule. In Taylor C. (ed), *Philosophical Arguments*. Cambridge, MA: Harvard University Press, pp. 165–180.
Varela F.J., Thompson E. and Rosch E. (1991). *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge, MA: MIT Press.
Veblen T. (1922[1899]). *The Theory of the Leisure Class: An Economic Study of Institutions*. New York: B.W. Huebsch.
Werner K. (2024). Enacted institutions, participatory sense-making and social norms. *Synthese* **203**, 152.
Whitman G. (2025). Countless ways to count: the functional heterogeneity of number systems. *Journal of Institutional Economics* **21**, e32.
Zednik C. (2011). The nature of dynamical explanation. *Philosophy of Science* **78**, 238–263.

Cite this article: Petracca E and Gallagher S (2025). Norms are relational: cognitive institutions, practices, and the 'where' question. *Journal of Institutional Economics* **21**, e39, 1–19. https://doi.org/10.1017/S174413742510026X