What is Money? An Alternative to Searle's Institutional Facts

JP Smit, Filip Buekens and Stan du Plessis

Abstract

In *The Construction of Social Reality* (1995), John Searle develops a theory of institutional facts and objects, of which money, borders and property are presented as prime examples. These objects are the result of us *collectively* intending certain natural objects to have a certain status, i.e. to 'count as' being a certain social objects. This view renders such objects irreducible to natural objects. In this paper we propose a radically different approach that is more compatible with standard economic theory. We claim that such institutional objects can be fully understood in terms of actions and incentives, and hence the Searlean apparatus solves a non-existent problem.

1. Introduction: Searle's view

In *The Construction of Social Reality* (1995) John Searle develops a theory of social institutions, of which money, borders and property are presented as prime examples. He explains how these entities come into existence, gives an account of their nature and puts forward a hypothesis about the conditions required to create them. We believe that Searle's theory, while ingenious, is wrong on all counts.

Searle's key claim is that institutional facts are genuine ingredients of our ontological framework, and that they are not reducible to brute facts. The distinction between brute facts and institutional facts, due to Anscombe (1958), is easy to grasp and fairly unproblematic: brute facts are those that exist independently of human intentionality, such as mountains, rivers, snow and temperature. They exist and are facts whether we regard them as such or not. Social objects, social facts and social institutions depend for their existence on human intentionality. A simple example is a screwdriver or a snowball: they are material objects but would not be screwdrivers or snowballs unless conscious agents regarded them as such. One particular subcategory of social objects or facts is Searle's category of 'institutional facts', of which money is his paradigmatic example (and the one we will focus on in our critique of

Searle's theory). The hallmark of money, according to Searle, is that it cannot perform its function unless some form of *collective acceptance* is in play.

Searle gives us three essential features of institutional facts: (i) collective intentionality, (ii) the assignment of function, and (iii) constitutive rules (1998: 124). The key feature of collective intentionality is that the relevant intentions for creating institutional facts must be expressed as 'we-intentions', which reflect, according to him, the sense in which we do something together. Collective intentionality is a 'biologically primitive phenomenon' that is exhibited in us and in many other species (1995: 24, 38). Collective intentionality manifests itself as the unique capacity of individuals to form *we-intentions*: "The crucial element in collective intentionality that each person has is derived *from* the collective intentionality that they share" (1995: 24-25). We are agnostic about the irreducibility of we-intentions¹. However, we do maintain that they are not necessary for an understanding of the origin or nature of institutional facts.

The second feature, the assignment of function, involves assigning a social function to some natural object; a piece of paper with certain physical features and origins is assigned the function of money and functions that way due to collective acceptance by the relevant actors. The assignment of function is itself based on the third feature, constitutive rules, which create the very possibility of certain types of institutional behaviour – behaviour that involves the use or employment of institutional facts. Unlike regulative rules that modify existing behaviour (like setting a speed limit, or rules of politeness), constitutive rules do not modify existing behaviour, but create the possibility of new actions. The form of these constitutive rules is '*X* counts as *Y* in *C*', where *X* refers to the object being given an institutional status, *Y* to the status itself and *C* to certain conditions under which the status applies. The use of this formula goes back to Searle's deservedly famous work on the nature of speech acts (Searle 1969: 52). Combined, these three features are, according to Searle, sufficient to demarcate a realm of institutional facts that are conceptually irreducible to natural facts.

While many have refined the Searlean approach and applied it to cases (Smith 2003; Smith et al. 2008; Lagerspetz 2006), our aim is to question a basic assumption: is the Searlean apparatus of collective intentionality leading to irreducibly institutional facts at all necessary

¹ See Tuomela and Miller (1998) for an argument that collective intentions can be conceptually reduced to a complex pattern of individual intentions and beliefs.

for understanding things like money, borders, etc? Our strategy is to pose a challenge to his account by presenting an alternative, more naturalistic account of the coming into existence of his paradigmatic examples of institutional facts, via an appeal to *incentivisation*. Essentially, we will be arguing that our strategy explains the same facts, but without postulating a new ontological realm, and hence should be preferred. In doing so we proceed in much the same way as Searle did. The bulk of this paper consists of discussing cases (traffic lights, simple borders, money) where our reductive analysis holds. The success of our reduction of standard Searlean examples like borders and money, the intuitive plausibility of our view and the lack of any positive reason to suspect that our analysis won't hold for all institutional facts lead us to suspect that such analysis can, in fact, account for all such objects. This is so even though, like Searle, we simply cannot provide a detailed analysis for each of the hundreds (thousands?) of institutional objects. Using Occam's razor, we then conclude that our simpler analysis is much more likely to be correct than the ontologically complex Searlean alternative².

Note that we will not challenge the coherence of the Searlean idea of collective intentionality. We do deny that it is a necessary condition for the existence of institutional facts, but we do not deny that it sometimes can be involved in the construction of such facts. We are simply agnostic on the matter. Since collective intentionality is not itself a non-reducible institutional fact this agnosticism is consistent with our view. This brings us to the crux of what we do deny: we deny that institutional facts cannot be reduced to natural facts.

An analogy may help to clarify the relation between our view and the Searlean view of institutional facts. We commonly talk about the sun setting and rising. No-one takes such talk to indicate ignorance of the fact that the earth does, in fact, orbit around the sun. The correct theory of cosmology explains what is really going on and why it is sometimes useful, for our purposes, to talk of the sun setting and rising. In a similar way we have no objection to the way people normally talk about borders, money and the like. We think that Searle's theory, however, takes our common way of talking too seriously when it posits a new ontological realm. We try and explain our common way of talking about such objects without postulating such a new realm, but also in a way that shows why our way of talking about such matters is

² We do, however, plan to extend our account of institutional facts in future papers and will attempt to show that our view can account for a wide array of social phenomena in a natural way.

generally useful. This means that we will still, in this paper, talk about 'social facts', 'social objects' and the like. This is simply for ease of expression as we do think such language can ultimately be reformulated into talk concerning physical objects, actions and incentives, as will be argued below. But we see no more reason to stop talking in these terms than we see reason to disapprove of talking about the sun rising and setting.

While we do not deny the obvious fact that in a metaphysical sense the entities Searle puts forward as prime examples of institutional facts *are essentially* mind-dependent – they could not come into existence in mindless worlds – we explain their mind-dependence in a more natural way that does not involve appeals to collective intentionality, constitutive rules, and status functions, and so we deny that institutional facts are irreducible. Rather, the mind-dependence of such objects is the result of the need for incentivisation, and the fact that talk about incentivisation is warranted only if talk about desires and beliefs - and hence minds - is warranted.

We also wish to be quite clear about the domain of our theory. In this paper we will only be addressing institutional facts, where by an 'institution' we mean formal structures like governments and the like. Hence we are primarily concerned here with things that *ordinarily* exist due to some government actions. We will discuss traffic lights, borders and money. These are things that are typically 'created' by governments, but institutional facts in our sense would also include driving licenses, reserve banks, police officers and a wide variety of other objects.

Two things should be noted about proceeding in this way. The first is that institutional facts as a category will be somewhat vague. It is an essential part of our theory that, while these objects typically are created by governments, they need not be. This will become clear when we turn our attention to money. It is also essential to our theory that the distinction between what is and what is not a government will also be somewhat vague. We develop our theory in terms of these objects for the simple reason that Searle starts his account with 'objects' like these. The second issue concerns what we leave out of the analysis. There are social facts that are not typically institutional, as we will use the term. These include promises, obligations, language, etc. Tackling these non-institutional social facts as well would make it impossible for us to present our views usefully in a single paper. We do think that we can explain promises and the like in a reductive, non-Searlean fashion, but will only attempt to do so in a future paper.

2. Institutional facts versus actions and incentives

2.1 Traffic lights

We will start by considering traffic lights. This simple example should be sufficient to demonstrate the difference between our view and that of Searle and to give the reader an intuitive grasp of our position.

A traffic light seems to be something that is easily analysed in terms of the Searlean formula 'X counts as Y in C'. The X-term would refer to the physical object, namely, the actual physical traffic light. But the traffic light does not fulfil its function due to the purely physical facts associated with it. Rather, this object *counts as* a traffic light only because we collectively regard it as having a certain status. What makes it a 'traffic light', as opposed to so many lights on a pole, is that the green *counts as* a signal that we should go, the red *counts as* a signal that we should stop, etc. This status of the lights obviously does not follow from their physical constitution, but from being objects of collective intentionality, which confers this status.

It is our contention that the above view, while coherent, is radically mistaken. Our examples and analysis will firstly show that collective intentionality is not always needed for institutional facts. We are agnostic on whether, in the final analysis, it is ever required, and similarly agnostic as to whether collective intentionality, as Searle understands it, exists. Our second, and more fundamental, claim is that the content of such facts can be fully analysed in terms of actions and incentives. Hence we do think that such facts reduce to natural facts. Our view is easily summarised by saying that we reject the heuristic '*X* counts as *Y* in *C*'³ in favour of the heuristic '*S* is incentivised⁴ to act in manner *Z* towards *X*'.

³ Searle does not regard the formula '*X* counts as *Y* in *C*' to be the canonical formulation of the nature of institutional facts, but rather, sees it as a useful heuristic that does a decent job of relating what is essential in his theory (Smith and Searle 2003: 300). Our challenge to his view is basic enough that we can proceed by challenging him on the level of this heuristic; its shortcomings do not affect the issues at hand.

⁴ We restrict the definition of 'incentivised' here to include only cases where the incentives are the result of intentional human action or belief. In other words, a falling rock cannot 'incentivise' you for the same reason that it cannot 'murder' you.

Our formula refers to actions and incentives, and in the case of a traffic light it is easy to see what these are. A traffic light is only a traffic light because a set of actions are essentially tied to it. These actions include stopping on red, continuing to drive on green, and proceeding with caution on yellow. These actions are tied to the traffic lights, not in virtue of an 'X counts as Y in C' formula, but in virtue of the fact that a given subject, S, has been *incentivised* to act in the prescribed manner when encountering a traffic light. Laws and customs incentivise such actions by prescribing punishment to any action that does not accord with the laws and customs relating to traffic lights. It is our contention that this is the full story about traffic lights. Once we are aware of the actions associated with traffic lights, and the incentives that tie these actions to the physical object, there is no remaining need for talk about 'irreducibly social objects', 'collective intentionality', and the like. Hence we have a clear instance of at least one institutional object that violates the essence of Searle's view.

2.2 'Institutional objects'

Before explaining the individual components of our view, we need to clarify what we mean by 'institutional objects'. Consider the following passage from Searle:

The notion of a social object seems to me at best misleading, because it suggests that there is a class of social objects as distinct from non-social objects ... rather, what we have to say is that something is a social object only under certain descriptions and not others, and then we are forced to ask the crucial question, what is it that these descriptions describe?

[T]here is only one object which is both a piece of paper and a dollar bill, but the fact that it is a piece of paper is not the same fact that it is a dollar bill, even though they are both facts about one and the same object (Smith and Searle 2003: 302).

In the above quotation Searle explicitly dismisses the claim, often put forward as an interpretation of his theory, that there is a distinct class of social objects. Rather, talk about social objects is just talk about social facts, and the descriptions describing these facts ultimately function only as a different way of potentially picking out the same natural object. In this regard, we fully agree with Searle. Although we will occasionally talk of 'institutional objects', we do not mean to suggest that there is an autonomous realm of institutional objects. All we mean is that there is an object which forms part of a fact that is 'institutional' in the

sense used by Searle. But we will portray these facts as a subset of natural facts, and not as a set of irreducibly institutional facts, as Searle does.

The traffic light example can again serve to illustrate our position on this issue. There are not two things in the world, an 'institutional' traffic light and a 'non-institutional' one. Rather there is just the one traffic light. The reason why we can, as 'shorthand', refer to it as an 'institutional object' is due to the fact that it has the properties necessary in order to be a part of an institutional fact. Searle considers these properties to be something like 'being the object of collective intentionality' and 'having a status conferred upon it that is irreducible to natural facts'. In our view, these properties are not necessary for traffic lights, borders, money and the like. Rather the property that distinguishes 'institutional objects' is 'having an action tied to it through incentives'. Simply put, we view talk of 'institutional objects' simply as talk about natural objects, where these objects can be picked out by a description like 'X that S is incentivised to act in manner Z towards'. To return to our original example, traffic lights are 'institutional objects' in virtue of the fact that they can be picked out by the description 'X that S is incentivised to stop at if green, etc.'

2.3 Actions and incentives

We claim that all institutional facts are essentially tied to a set of actions. Traffic lights are objects that I am incentivised to stop at, etc; borders are things I am incentivised not to cross; money is something I am incentivised to acquire for exchange, etc. We do not think that institutional objects *give rise* to these actions and incentives; rather, we think that they simply are *natural objects individuated by* these actions and incentives.

In simple cases, it will be possible to state these actions without using concepts that are in some sense 'institutional', and hence, it is easy to show that we can define an 'institutional object' in terms of natural facts. This is why we chose the example of traffic lights; traffic lights are physical objects of a certain shape and colour that I have been incentivised to stop at, etc. The situation is similar with some borders; borders are simply objects (lines) that I have been incentivised not to cross. However, in order to give a full account of the actions normally tied to borders, we would initially need to avail ourselves of 'institutional concepts' like 'passports', 'consulates', etc. The situation is even worse with money: a simple definition of money would have to involve reference to buying and selling, actions that are clearly being

defined in terms of 'institutional facts'. It is our position that all these terms can be fully cashed out in terms of actions and incentives. We will not tackle this topic fully in this paper.

The notion of an incentive being used here is the standard economic one describable in terms of the relative desirability of a set of possible options, but one thing should be noted: that the source of the incentives does not matter when conceptualising the nature of the 'institutional object' itself. Take the case of a border, simply defined as a 'line I am incentivised not to cross' (as we argue below). None of the concepts in this simple definition are institutional. The *source* of the incentives are not part of the concept of 'border', and hence different people can view the same border, and have the same conception of it, despite the relevant incentivisation coming from different sources and having different force. The relevant incentivisation can be strong or weak, morally legitimate or illegitimate, and based on custom, law, or moral beliefs, without this making it less of a border. What a threat or promise does is to change the set of possible outcomes, or their relative probability. What the adoption of a certain moral belief does is simply to change the relative valuations of these possible outcomes. In other words, if A believes that to cross the line would be immoral, and B believes that, if he crosses the line, he will get shot, then they are appropriately incentivised. The fact that the source and strength of their incentives differ has no impact on whether they believe a border exists. These are beliefs that they have *about* the border, not beliefs that change the concept 'border'.

Our basic view of how the logic of incentivisation works is very simple. An actor (or actors) act(s) in such a way that another actor (or actors) is (are) incentivised to act in a specific way. The actions of the incentivising actors and the incentivised actions that result will differ for each social fact, but the deep logic is the same for all such facts. When such incentivisation is stable enough the result is the kind of pattern of activity that we call a 'social fact'. When this process is directed at treating an *object* in a specific way the result is a Searlean 'institutional fact'. We will illustrate these, and related issues, more clearly below.

2.4 Borders

2.4.1 Searle on borders

Searle (1995: 39-40) discusses borders as a simple case of how institutional facts can be created. He asks us to imagine a tribe that builds a wall around its territory to keep foreigners

out (39). The physics of the wall itself is enough to keep foreigners out; it functions as a border due to the physics of the object itself, and as such, it does not yet constitute an institutional fact. Next, Searle considers what would happen if the border gradually wore away so that only a line of stones remained, and the physical properties of the object itself were no longer sufficient to guarantee territorial integrity. In such a case, it is possible that the wall would still be recognised as a boundary by neighbouring groups, and that this would affect their behaviour (39). It would now have ceased to act as a physical deterrent, but would still function to indicate the boundaries of a territory. This, however, would now be due to the symbolic nature of the line of stones, and not to the physics of the object involved. We are now dealing with an institutional fact that can be described according to the formula '*X* counts as *Y* in *C*'. In this case, the line of stones would count as a border, and critically, the idea of 'being a border' cannot be cashed out in terms that do not make reference to institutional facts.

2.4.2 Alternative borders

Let us take a very simple scenario, reminiscent of the 'Robinson Crusoe–style examples so beloved by economists. Alex and Bob wash up on a desert island. They quarrel, and in a fit of pique, Alex says: "If you enter this half of the island I will beat you up". Alex accompanies his claim with a demonstration indicating that he is referring to territory extending between two rocks on opposite sides of the island, cutting the island in half. Bob, angrily, replies: "And if you enter this half of the island, I will beat you up". Assume, for the sake of argument, that this is a credible threat from both sides, and thus it incentivises both to stay on 'their' side of the island⁵.

It would be difficult to deny that a border had been set up on the island. But this has happened in a way that violates the essence of Searle's view. Firstly, the requirement for collective intentionality has not been met, as *all of the relevant thoughts and claims can be expressed using the singular 'I'*. Secondly, nowhere is reference needed to any irreducibly social facts, objects or properties. Both actors can understand the situation fully by using concepts like 'line', 'crossing', 'probability of getting beaten up', etc. The expressed intentions of the actors, if credible, incentivise the two actors not to cross the line, and a border is created. We

⁵ In other words, the expected utility of gaining access to land is less than the expected disutility of injury, etc., and this is true for both parties.

need nothing beyond an understanding of the incentives and beliefs of the two actors in order to grasp the situation fully.

Note, once again, that we impose no constraint on the source of the incentives that can establish an institutional fact. In the above example, Alex could unilaterally build a wall to make it difficult for Bob to set foot on 'his' half of the island, or build a wall and issue a threat, etc. Alex could also have unilaterally issued a threat, without Bob necessarily responding in kind⁶. If any of these actions have the upshot of incentivising them to not stray onto the other's land, then a border has been created. Alex and Bob could even act in the most 'Searlean' way imaginable, get together, and collectively decide to let an indicated line act as a 'border' between them, to be respected by both. If this agreement incentivised them to not stray onto the territory of the other, a border would have been created. Whether they followed this agreement out of an implied threat enforcing it or due to a moral belief that agreements are to be followed does not matter for the creation of the border⁷. All that matters are the resultant incentives⁸. The important thing to note in such a case is that this somewhat artificial construal of the 'Searlean' methodology is not *constitutive* of the very creation of the border. Rather, it is simply another method of making a border come about. And it only succeeds in doing so because it incentivises the participants appropriately.

⁶ This would be analogous to historical cases where a stronger power simply took a piece of land, banishing the conquered to a fraction of their former territory. For the conquered, it would be quite obvious that a 'border' simply is a matter of 'cross this river and get shot'.

⁷ Searle, of course, is putting forward a theory of the nature of institutional facts, not of their origin. We include this part because i) some seem to interpret Searle as committed to some views regarding the origin of institutional facts (Lagerspetz, 2006), and ii) we think that, if the above logic can explain the origin of such facts and what sustains them, there is no need to suddenly introduce an entirely new theory of the *nature* of such facts.

⁸ This is clear when we consider real word examples; for instance, the three 'Cod wars' between Britain and Iceland from the 1950s to the 1960s followed this pattern: in each instance, Iceland unilaterally declared an 'exclusive economic zone' first to 12 nautical miles (nm) in 1958, then to 50 nm in 1972, and finally, to 200 nm in 1975. Britain did not recognise these expansions as legitimate, and this led to a series of naval confrontations. Eventually, the British government agreed to withdraw all British fishing boats from the extended 200 nm boundary to improve diplomatic relations with its NATO ally Iceland. The expanded exclusive economic zone of Iceland should be understood in terms of unilateral action and incentives (Kurlansky, 1997).

Given the above logic, it is easy to see how institutional facts are sustained. If the incentives associated with the institutional fact remain⁹, then the institutional fact will remain, but if the incentives change the institutional fact will change or disappear. This could happen if Alex became unable to defend his territory, or Bob came to believe that the initial settlement was unfair, or one half of the island was found to contain objects that were valued so highly that the risk of injury could no longer deter the other from straying onto it, etc. In such a case, the border could come to be violated so often, and to such a degree, that over time it became pointless still to regard it as a border. 'Incentive compatibility', as economists call it, implies feedback - rewarding behaviour consistent with the institution and imposing costs on inconsistent behaviour. Where the feedback effect is not compatible with the institutional fact the institutional fact will eventually disappear. Economists have studied the incentive compatibility of institutions very extensively in the burgeoning New Institutional literature as a crucial explanation of the co-operative or conflicting behaviour observed in society (see, for example: Kasper and Streit, 1998; World Bank, 2002). Note that the above understanding of institutional facts also implies that the difference between a single case of being incentivised in a certain way, and a situation where the incentivisation is strong, widespread and durable enough to amount to an institutional fact, is a mere *statistical* difference.

The upshot of the above should be quite clear. One could look at the actions of Alex and Bob, see that they refrained from crossing a certain line, and conclude that there was a 'border', in Searle's sense, between them. From the fact that the line itself could deter them from crossing it in virtue of its physical characteristics, we could feel forced to conclude that the 'border' was an object that was irreducibly institutional. This fact, coupled with the fact that both agents respected the border and seemed to understand it in the same way, could further lead us to conclude that they must have understood themselves as collectively viewing this line in the same way. Hence, all the normal temptations in favour of the view that the border was to be understood as a case of 'X counts as Y in C' were present here. But we trust that the reader is not feeling these temptations as strongly as is normally the case. Our (much simpler) analysis is that unilateral or collective action tied an action to an object through incentivising such

⁹ This only means that there is a *prima facie* incentive of the required type, not that all people are at all times, all things considered, incentivised to act in the relevant way. Take the example of borders: if an enemy army occasionally invades in order to steal cattle, but it remains costly for them to do so, and the defending army can keep this cost (disincentive) in place, the border remains.

action, and that we simply had a case of 'S was incentivised to act in manner Z towards X'. It may well be that readers of Searle could confidently apply Searle's vocabulary to the situation at hand, but that is not an argument for the irreducibility of the concepts of Searle's framework.

Above, we have made much of the simple example of borders¹⁰. Much more has been written on 'money', and this is the subject that is most often cited as being somewhat mysterious. Below, we will illustrate how easily our analysis applies to something like money.

2.5 Money

2.5.1 'Money' as being incentivised to act in a certain way

On Searle's view 'money' is constituted by the fact that a group of people collectively decide that a certain object is to count as 'money', i.e. the meaning of 'money' cannot be cashed out except by using terms that cite irreducibly social facts or properties.

But once again, another alternative is available, since normal game-theoretical considerations can lead to the same outcome. Consider three people (Alex, Bob, Carol) on an island. They engage in barter¹¹ on the island. Imagine they also have cigarettes on the island, despite the fact that no-one smokes, and cigarettes are of no direct use on the island. One day Alex announces that, instead of accepting a chicken in exchange for three shirts, he will accept 20 cigarettes. He makes it clear that he does not intend to smoke the cigarettes, but that, from now on, he will always be willing to accept a certain undetermined number of cigarettes in exchange for commodities. To lend credibility to this commitment Alex might invest some of

¹⁰ In our toy-example the definition 'line that S is incentivised not to cross' was sufficient for an account of the situation. In the real world, of course, borders can do more than that. Provincial or state borders, for instance, have little, if anything, to do with keeping people in or out. Rather they indicate different administrative regions. To accommodate these (and other) uses of the term border we need the broader definition 'line dividing two areas that S is differently incentivised towards'. National borders can then be seen to be a kind of border.

¹¹ 'Barter' here seems to presuppose property rights, which seem to presuppose 'social facts'. We will deal with this issue below.

his resources (time, etc.) in making a fine cigarette case to show his commitment to the 'cigarette standard'¹².

Why would Alex forego part of his consumable wealth in this way? Karl Menger (1892) provided the first modern answer to this question by focusing on the economy of effort a society can achieve by adopting a single commodity as money. While Menger focused on the physical characteristics of money, such as the durability, divisibility and desirability of gold, modern analyses have moved away from citing these physical features as explanatory factors. Brunner and Meltzer (1971) and Alchian (1977) found the origin of money not in the physical features of money that allow it to solve problems such as the double-coincidence of wants, or other search costs, but in the role that money plays to reduce the "information about the attributes of goods available for exchange" (Alchian 1977) and to render more efficient chains of transactions that lower the cost of information asymmetries (Brunner and Meltzer 1971).

The latest work on monetary theory has emphasised the potential of money to solve strategic problems such as trust and memory in social interactions with a time dimension: Kiyotaki and Moore (2002) show how money can overcome a lack of trust, while Kocherlakota (1998, 2002) shows how money smoothes trade when contracts are imperfectly enforced and memory is limited. All this shows that there are many reasons to expect the emergence of money in a social setting¹³. It is not surprising that money emerges so readily in many societies and under widely different circumstances, even if the form it takes may differ dramatically (e.g. gold, silver, sea shells, cattle, large stone rings or bits of paper with certain markings on them) (Friedman 1992)¹⁴.

¹² By analogy with the cavernous marble halls built for bank branches prior to the government-backed deposit insurance. These expensive marble halls signalled a commitment by the Bank to prudent business and flagged their intention to be around for many years.

¹³ Software engineers have found that the participants in large scale multi-player computer games (or virtual worlds) such as "Second Life" and "World of Warcraft" cannot suppress trading or the emergence of money in their games, and as a consequence, they now facilitate this by issuing their own currency, such as the Linden dollar (L\$) in "Second life" (Castronova 2005).

¹⁴ The recent literature de-emphasises the physical features of money – which were quite important for what Menger (1892) called the 'saleableness' of money – and focuses instead on the features required to fulfil the other tasks mentioned above. From this more modern

Returning to the example of Alex, Bob and Carol, the unilateral attempt to introduce cigarettes as money would succeed if, and only if, Alex's unilateral action incentivised the other people on the island to act accordingly. If Bob and Carol were trading it would be rational for them to accept cigarettes, since the cigarettes would now 'represent' the possibility of future trades with Alex. This, in turn, would reinforce Alex's commitment to using cigarettes, since the willingness of Bob and Carol to accept them as currency would incentivise him in the same way when dealing with them, which would reinforce their use of cigarettes, which would reinforces his use, and so on.

Any such unilateral action can get this self-reinforcing logic started, and more than one of the islanders might try the same thing. Greenfield and Rockoff (1996) use a tipping-point model¹⁵ to show that under normal circumstances the choice of money is usually stable¹⁶. What matters most in the decision to express your prices in a certain currency, or your decision to accumulate your liquid wealth in a certain currency, is the behaviour of others. Once sufficiently many people use a certain type of money, the outcome of the competition amongst several monies is likely tip towards the dominant standard. It is important that Alex's attempted monetary reform would only work – cigarettes would only become the dominant money – when the tipping point had been reached, i.e. when many decision-makers accept money, confident that they will in turn be able to use it as money.

In this way, a monetary system is created without collective intentionality and without any 'objects' that cannot be cashed out fully in terms of natural facts concerning changed incentives. We will discuss how 'money' fits the formula 'S is incentivised to act in manner Z towards X' when we discuss fiat money specifically, as there is a subtlety involved that will

perspective, it is not surprising to see examples of peculiar assets emerging as money, such as cigarettes in a POW camp (Radford 1945) – or tinned mackerel in US prisons (Scheck 2008).

¹⁵ These tipping-point models have wide use in strategic settings such as arise with competing monies, and were first analysed by Thomas Schelling (1978).

¹⁶ Hyperinflation or the collapse of an external peg for the currency could undermine confidence in a currency and encourage a society to solve the problems of exchange with the help of alternative money. In Argentina, by 2001 a fiscal crisis combined with the external collapse of the currency was sufficient to incentivise the emergence of more than a dozen new kinds of money issues by cities such as Buenos Aires, local governments and even a shopping centre (Catan 2002)

1

only become clear later. But note, once again, that the source of the incentives does not matter. In the above example Alex did not engage in any coercion to set up the system, but he could have set up the system using the threat of force. Alex could have announced that, at the end of every month, he would be expecting one cigarette each from Carol and Bob, and that, if they did not comply, he would punish the person who defaulted¹⁷. As long as there was a non-zero probability that this threat would be made good, Bob and Carol would be incentivised to acquire cigarettes as a means of ensuring future security, and cigarettes would become currency¹⁸. We could even have people collectively agreeing that from now on cigarettes 'should count as money'. 'Money' would result as long as this 'agreement' lead to a non-zero probability that someone would exchange cigarettes in exchange for good or services. This would lead to the cigarettes having this value for all actors involved as they are now incentivised to value the cigarettes in this manner. The important thing to note is that a unilateral change in the action of even, in principle, one individual, can change everyone else's incentives in such a way that an 'institutional fact' is created. What matters is that the appropriate incentives are created, not how they are created.

2.5.2 Fiat money

Reading the literature on fiat money, one gets the idea that it is considered a paradigm case of the mysterious human ability to create institutional objects out of thin air. There seems to be the idea that somehow 'value' is created out of thin air, and that fiat money is 'not backed by anything'. This is a fallacy. Fiat money is not 'not backed by anything'; rather, it is 'not backed by any *specific, concrete* thing at a *fixed rate*'. This can be seen by looking at the

¹⁷ This is basically equivalent to governments insisting that taxes should be paid in the new fiat currency. An early example is the fiat currency created in France by Johan Law during the minority of Louis XV and the regency of the Duke of Orleans, where the main incentivisation mechanism used was a decree that all taxes should be paid in the new currency (Ferguson 2009: 140).

¹⁸ An instructive case is the situation in Zimbabwe in 2009. Even at ludicrous inflation (where the inflation-rate runs into the millions), the currency still exists as money. This is due to the incentives created by government by paying soldiers and other government functionaries in Zimbabwean dollars, banning the use of any other currency, accepting Zimbabwean dollars as payment for income taxes, and making it illegal to refuse to accept Zimbabwean dollars as currency, etc.

example above, where cigarettes were currency, despite it being stipulated *that no one had any practical use* for them. Cigarettes, in the above case, are fiat money.

Fiat money is different from money that, say, is backed by a gold standard, but this difference is not nearly as dramatic as is commonly supposed. In commodity money (gold, cigarettes) the 'intrinsic worth' of the commodity acts as an anchor for the value of the currency. Of course, there is no such thing as 'intrinsic worth' in any economically relevant sense. Gold is only valuable because we desire it for both industrial and decorative uses. This 'intrinsic worth' is fully explicable in terms of the surface and chemical qualities of gold, as well as our desires and beliefs. These desires and beliefs, obviously, can change over time, and so gold is, in principle, no more an 'inherent or constant store of value' than any other object of our desires would be. In the final analysis, the value of gold is 'backed' by nothing more than our preference for it, something which can, and historically has, changed over time. The same is obviously true of 'contract money' or 'token money', i.e. currency that is backed by a commodity at a fixed rate. What goes for gold in this regard would also go for currency on a gold standard. Fundamentally, it is not backed by anything more than our changing preferences.

Fiat money is often thought to be strange, and in some sense this is true. In contrast to gold, the actual object has no, or only trivial, use value. It also seems 'to be tied to nothing'. This, however, is an illusion. It is, in the final analysis, 'backed' by our preferences, or our incentives to act towards it in a certain way, *just like gold*.

Fiat money is normally issued by a government, often with an initial exchange rate pegged to an existing currency. This existing currency, and the possibility of trading currencies, and hence having access to the goods obtainable in that currency at existing prices means that such fiat money is backed by a host of existing price relations. Whether 'collective intentionality' or any special ceremony had been involved is irrelevant. The point is that, if someone else (like government, business, etc.) suddenly promises to exchange these currencies at a given rate, then I am incentivised to obtain the new currency and 'money' results.

It could be objected that the prior currency, i.e. the currency serving as initial peg, could also have been a fiat currency, and so the mystery is not resolved. But excluding this possibility makes no real difference. There are many different ways of creating the incentives that tie the action 'acquire for exchange, rather than for consumption' to a type of object. Even the different ways that governments (and non-governments) have historically done this is but a small set of the possible ways of accomplishing this. In principle, it could be something as simple as government paying all public-sector salaries in this new currency. This would mean that other businesses could poach these workers by paying superior amounts of this currency, which would mean that other businesses would know that the first-mentioned businesses wanted some of this currency. So it would make it rational for them to acquire it in exchange for goods and services, which is why the government workers can use it to acquire goods in the first place, etc. This currency was not really backed by 'nothing', but by the services that the employees in the public sector could deliver, and hence, in the final analysis, by our changeable preferences for these services. A myriad of other methods are possible. Government could demand that all legal transactions be calculated in this currency, or could stipulate that it is legal tender¹⁹, or could prohibit other currencies from being used, or could demand that all legally recognised salaries be calculated in it, or could demand that taxes be paid in it, or any combination of the aforementioned, etc. The upshot of all of this would simply be that it would be a good idea for me to get my hands on some of it. Other people also realise this, and so we get a functional common currency. In all these cases, governments created money, not by creating an 'institutional object', but by manipulating incentives²⁰.

The upshot of the above is that commodity money and fiat money differ in that fiat money is not backed by one commodity, but by *all* exchangeable commodities in an economy, and that the rate at which it is backed can change, rather than being fixed, as with gold²¹. But in the deeper sense, they are backed by the same thing, namely, human preferences. To hanker after

¹⁹ The main incentivisation mechanism is some *legally enforced*, blanket declaration equivalent to the "this note is legal tender for all debts, public or private" printed on US banknotes.

²⁰ How easy it is to do this is borne out by the fact that a lot of currency-laws are not so much aimed at creating currency, but *preventing* it from being created by non-governmental institutions, by banning such institutions, demanding to be paid in one currency only, making it an offense to refuse to honour currency in commercial transactions, etc. Currencies occur so 'naturally' that government has to do a lot of incentivisation to preserve their monopoly.

²¹ The supply factors of fiat money are obviously not physical and technological, as is the case for all commodities, including commodity money. Rather the supply factors relevant here are the rules that limit that discretion of the central bank. This, however, is not a difference that introduces any conceptual difficulties.

a gold-standard if you do not trust your government to refrain from printing a lot of money can be rational, but to desire it due to the belief that money must be 'worth something' is to fundamentally misunderstand the nature of fiat *and* commodity money. And fiat money issued by an industrial superpower should no more drive us towards vague ontological commitments to 'institutional facts' than the simple expression of an intention to accept cigarettes in exchange for goods on an island should.

2.6 Fiat money, electronic money and 'free-standing X-terms'

There remains one problem that needs to be addressed. An objection Searle's account frequently encounters concerns the issue of the ontological nature of electronic money. Searle's heuristic 'X counts as Y in C' seems to presuppose that there will be an object that serves as the referent of the X term. In cases like fiat money this does not seem to be a problem as there are actual dollar bills we can identify as the referent of the X-term. But the situation is much murkier when we consider the case of 'electronic money'. Only a relatively small percentage²² of the money supply in the world exists in the form of notes and coins. Most of it only 'exists' in the form of entries in the account ledgers of financial institutions. This, of course, is a sensible practice, as only a small percentage of the money in the world actually needs to circulate. Central Banks monitor the demand for notes and coins carefully and can respond to a surge in that demand (due to a withdrawal of deposits, for example) by supplying more notes and coins for circulation. Indeed the ability of Central Banks to respond in this manner to a surge in the demand for the most liquid form of money, notes and coins, is an argument for having a Central Bank in the first place.

This does, however, cause a problem for Searle's formula. The problem, originally pointed out by Barry Smith (Smith and Searle 2003: 287) is that there is now no real, physical object for the *X*-term to refer to. Searle originally (2003: 301) replied to the problem by claiming that '*X* counts as *Y* in *C*' is only a heuristic and that the objection does not much matter. He had also previously (1995: 56) claimed that the entries on an account ledger (or, more realistically, the electronic blips on a server hosting the data) can count as money in the same way that notes and coins can. But, as Smith rightly pointed out, the data is not money, rather it

²² The (seasonally adjusted) money stock data for the USA in March 2010 serves as a typical demonstration. Currency in circulation amounted to \$871.6 billion, while the broader money stock amounted to \$8512 billion (Federal Reserve Statistical Release, H.6, 20 May 2010). Hence the currency in circulation was only 10.2% of the broader money stock.

is a record of money (2003: 287). It would be perverse to count the separate records of the same fact as also constituting the fact. Rather it is intuitively clear that such separate entries are records of one fact, but that this one fact does not quite fit Searle's theories. If a ledger gets burned (or a server destroyed) no money is lost, rather we have simply lost one record of a certain fact. Searle now seems to accept that such electronic money is a mere record, but still claims that this does no real damage to his theory, as his schema '*X* counts as *Y* in *C*' was only ever supposed to be a heuristic. Our heuristic '*X* that *S* is incentivised to act in manner *Z* towards' is open to the same objection as Searle's account as they both use an *X*-term in need of a referent. There is, of course, no need for the *X*-term to be a physical object in the ordinary sense²³, but we do need there to be a referent.

We have some sympathy with Searle in that we do not think that the objection strikes at anything vital in his or our view. But we also have some sympathy for Smith's argument, as Searle provides no real evidence for his assertion²⁴ that this problem is not that important. Furthermore, we do think that Smith's objection leads to an interesting insight. The problem can, in essence, be stated as follows: paper money is actual money, but electronic money is only a record of money. What, then, is the ontological status of electronic money?

When stated as above we think it becomes apparent that the problem embodies a fundamental confusion, because paper money is not, *by this standard*, money at all. It is itself a mere *record of money*, in exactly the same sense that electronic money is a record of money. The two are on a par; there is no qualitative difference between the actual paper in your wallet and the account entry at a bank. Both should be understood as records, as will be explained below. But before we proceed to the full explanation we would like to point out one fact that should give our account some initial plausibility. No money is actually lost if an account book or server is damaged. This indicates that it is a record, and not actual money. But this is also the case if paper money is damaged. It is standard practice for governments to replace currency that has been accidentally damaged or mutilated. In fact governments standardly create departments that look into cases of mutilated or damaged money and, if the evidence is

²³ As Searle rightly points out concerning the issue of borders, both a sequence of extensionless points and an area of space still count as brute phenomena (2003: 308).

²⁴ Searle still does not really address this issue in his recent *Making the Social World: The Structure of Human Civilisation* (2010).

adjudged to be in your favour, they will replace it. In this way it is treated as a mere record and not as an article of actual value²⁵.

To explain why the practice of replacing currency makes sense we need to remind ourselves of the history of money. Money originates, normally, as commodity-money, i.e. as an object with use-value (like gold) being acquired primarily for exchange. The next stage in the evolution of money is normally a move to 'contract money', i.e. where a bank issues tradable certificates of deposit of commodities like gold. These certificates are often called 'token money' in the economic literature and the name is quite fitting. The certificates serve as records of 'actual money' and as strong *prima facie* evidence of ownership of the recorded amount of the relevant commodity. Token money is uncontroversially understood as a 'record of money', and specifically as a record of the backing commodity (like gold). The conceptual confusion responsible for Smith's problem originates in the idea that a move from token money to fiat money is a move from a currency that serves as a record of a backing commodity to a currency that is 'backed by nothing'. This, we have argued, is a mistake. Fiat money is still a mere record of an underlying 'commodity', but now this commodity becomes something abstract and the relation between the actual note and the commodity that backs it is no longer set at a fixed rate.

To illustrate this, consider again a population on an island that engages in barter. Now imagine that a powerful individual (Alex) manages to somehow create an amount of paper currency (the Lex) and distributes it. He then announces that, at the end of each year, he will be expecting a certain amount (50 Lex) of it back (in a way analogous to real-life taxes²⁶) and

²⁶ Our frequent mention of legally imposed taxes raise the issue of how our view deals with laws in general. We view laws, roughly, as declarations as to how people should act, typically coupled with explicitly stated sanctions for the failure to comply. The incentivisation is then, as with cases like traffic lights and borders, done by *both* official governmental sanctions and non-governmental social sanctions. These non-governmental sanctions, as noted earlier, can arise both from the practical interests and moral convictions of people. One interesting fact about law is that the motivation for such convictions and interests can be either concrete (a condemnation of the specific act) or abstract (a condemnation of the act of law-breaking as such). Space constraints prevent us from discussing these ideas in detail here.

²⁵ In the USA, for example, the Bureau of Engraving and Printing has the authority to replace currency that has been damaged so severely that a bank will not replace it. Institutions and individuals can, and routinely do, have damaged money replaced.

that failure to comply will lead to a lash with a cane. Assume that Alex is powerful enough for this to be a credible threat. Here we have *fiat* currency, but it is obvious that it is not 'backed by nothing'. Rather the backing commodity is the benefit of not getting lashed by a cane. In other words 50 Lex = 1 not-getting-lashed. This incentivisation is sufficient for the inhabitants of the island to have a desire to acquire Lex, and hence for it to be used as a medium of exchange, as explained earlier. What *really* gets traded is avoiding being lashed, and the actual paper is no more than a record of the holder's ability to avoid being lashed. This is, in every philosophically relevant way, *analogous to how token money backed by gold works*. The only difference lies in the fact that the backing commodity is an abstract ability and not a concrete object.

Note also that, even if Alex dies and the threat disappears, the Lex could still be used as currency. Alex's original incentivisation is merely needed as a catalyst to get the process started, not to sustain it²⁷. Once the process is started the normal self-reinforcing logic of money kicks in and the variety of advantages of money, mostly related to how it massively lowers 'transaction costs' (in the widest sense) is enough of to give the society a strong positive reason to continue the practice. More importantly, there are no, or almost no, reasons to unilaterally defect from such a system. Those who have money and other assets would only lose their wealth, and increase their transaction costs, if they did so, and defection by those with no money or assets would be meaningless. In such an event the 'backing commodity' then becomes, ultimately, the ability to acquire goods at massively lower transaction costs. This is the result of the *incentives* at play in a society where some action has served as a catalyst to start such a self-reinforcing system and partly accounts for the stability of currencies referred to earlier.

The *X*-term in our formula should be understood as referring to the ability to acquire the commodities money can buy. The value of this, ultimate, backing commodity of fiat money, i.e. the abstract ability to acquire goods, will be the result of a mix of the type of government incentivisation discussed earlier (analogous to Alex's declaration) and the goods and services I can acquire (due to the self-reinforcing logic at play) with the currency. This ability is what is fundamentally being exchanged. Actual paper and account entries are records of various institutions' and individuals' abilities to engage in such trade. Paper money is, ultimately, a

²⁷ What happens in such a case, essentially, is that the formal (governmental) incentivising 'backing' disappears and the purely informal incentivising 'backing' remains.

mere record in the same way that electronic money is²⁸ is a record and that the notes associated with 'token money' are a record. Our way of talking about these matters, while practically useful, obscures this fact. This is exacerbated by the 'folk theory' of fiat money in which fiat money is 'money backed by nothing'. Things get even worse when people erroneously claim that fiat money is created by a government declaration that replaces 'one dollar = one unit of gold' with 'one dollar = one dollar'. If such declarations did anything it would indeed be mysterious and disconcerting (as the uncomprehending Marco Polo experienced the creation of fiat money at the mint of Kanbala attached to the court of Kublai Khan²⁹). What those who take our folk theory for granted tend to forget is that such declarations are typically accompanied by a host of complex laws that serve to accomplish the relevant incentivisation in a manner ultimately analogous to Alex's threat of caning, and that this incentivisation is what allows the virtuous cycle of mutual incentivisation to begin.

3. Conclusion

We have argued that a construal of institutional objects as existing in virtue of irreducibly social objects arising in virtue of collective intentionality is mistaken. Proponents of the autonomy of institutional facts consider these objects to rationalise certain actions and have certain consequences. For us, the objects are not distinct from what they rationalise. Rather they simply *are* the incentivisation of the conjunction of these actions, and can be defined and understood in terms of these actions. Institutional objects are constituted by a subject being

²⁹ "In this city of Kanbala is the mint of the Great Khan, who may truly be said to possess the secret of the alchemists, as he has the art of producing money by the following process....the coinage of this paper money is authenticated with as much form and ceremony as if it were actually pure gold or silver; for to each note a number of officers, specially appointed, not only subscribe their names, but affix their seals also...in this way it receives full authenticity as current money, and the act of counterfeiting it is punished as a capital offence. When thus coined in large quantities this paper currency is circulated in every part of the Great Khan's dominions; nor dares any person, at the peril of his life, refuse to accept it in payment. All his Majesty's armies are paid with this currency, which is to them the same value as if it were gold or silver. Upon these grounds, it may be certainly affirmed that the Great Khan has a more extensive command of treasure than any other sovereign in the universe" (Polo 1930).

²⁸ Another way to show this: Imagine a society where everyone is perfectly honest and everyone has a perfect memory. They can dispense with the paper and merely keep track of the numbers in their heads.

incentivised in a certain way, and this incentivisation can be done collectively or individually. As a heuristic, when thinking of institutional facts, Searle offers the formula 'X counts as Y in C'. Our favoured heuristic would be something like 'X that S is incentivised to act in manner Z towards'.

We do not deny that our common way of speaking about institutional facts makes it sound as if we are constantly populating the world with all manner of institutional objects. But we view this as a kind of heuristic, i.e. a way of speaking that enables us to get around effectively in our world, and which causes no great harm when understood as such. It would be unnecessarily cumbersome to replace talk of borders, money, states, etc. with talk of differential incentives, preferences, etc. This would be akin to asking people who speak of 'evolutionary drives', 'biological function' etc. to instead reformulate this type of talk into ontologically respectable talk concerning differential survival rates. The same applies to our institutional facts. In ordinary life, it is fine to talk about money, borders and property. However, when we want to make sure that we are 'cutting reality at the joints', we need to remember that we are fundamentally talking about the incentivisation of actions.

References

Anscombe, G. 1958. On brute facts. Analysis 18: 69 – 72.

Alchian, A. 1977. Why money? Journal of Money, Credit and Banking, 9: 133-140.

Catan, T. 2002. Argentines snowed under by paper IOU's. Pesos or Pacificos? A dizzying array of currencies now fill up the tills. *The Financial Times*. London. 11 April 2002: 4

Castronova, E. 2005. *Synthetic Worlds: The Business and Culture of Online Games*. Chicago: University Of Chicago Press.

Brunner, K. and A. Meltzer 1971. The uses of money: money in the theory of an exchange economy. *American Economic Review* 61: 784-804.

Ferguson, N. 2009. *The Ascent of Money: A Financial History of the World*. London: Penguin Books.

Friedman, M. 1992. *Money Mischief. Episodes in Monetary History*. New York: Harcourt Brace & Company.

Greenfield, R and H. Rockoff 1996. Yellowbacks out west and Greenbacks back east: socialchoice dimensions of monetary reform. *Southern Economic Journal*, 62: 902-915.

Kasper, W. and M. Streit 1998. *Institutional Economics: Social Order and Public Policy*. Cheltenham: Edward Elgar.

Kiyotaki, N. and J. Moore 2002. Evil is the root of all money. *American Economic Review*, 92: 62-66.

Kocherlakota, N. 1998. Money is memory. Journal of Economic Theory 81: 232-251.

Kurlansky, M. 1997. Cod. A Biography of the Fish that changed the World. New York: Penguin.

Lagerspetz, E. 2006. Institutional Facts, Performativity and False Beliefs. *Cognitive Systems Research* 7, 298 – 306.

Menger, K. 1892. On the origin of money. Economic Journal 2: 239-255.

Polo, M. 1930. *The Travels of Marco Polo (The Venetian)*. New York: Liveright Publishing Corp.

Radford, R. 1945. The economic organisation of a P.O.W. camp. *Economica* 12: 189 - 201.

Scheck, J. 2008. Mackerel economics in prison leads to appreciation for oily fillets. *The Wall Street Journal*. New York. 2 October 2008.

Schelling, T. 1978. Micromotives and Macrobehavior. New York: W.W. Norton & Company.

Searle, J. 1995. The Construction of Social Reality. London: Penguin Books.

Searle, J. 1998. *Mind, Language and Society: Philosophy in the Real World*. New York: Basic Books.

Searle. J. 2010. *Making the Social World: The Structure of Human Civilisation*. Oxford: Oxford University Press.

Smith, B. and J Searle. 2003. The construction of social reality: an exchange. *American Journal of Economics and Sociology* 62: 285 – 309.

Smith, B, D. Mark and I. Ehrlich (eds.) (2008). *The mystery of capitalism and the construction of social reality*. Chicago: Open Court

Tuomela, R. and K. Miller. 1998. We-intentions. Philosophical Studies 53: 367 - 389.

World Bank 2002. World Development Report 2002. Building Institutions for Markets. Oxford, Oxford University Press.