

## Realism and Human Kinds

AMIE L. THOMASSON  
*University of Miami*

It is often noted that institutional objects and artifacts depend on human beliefs and intentions and so fail to meet the realist paradigm of mind-independent objects. In this paper I draw out exactly in what ways the thesis of mind-independence fails, and show that it has some surprising consequences. For the specific forms of mind-dependence involved entail that we have certain forms of epistemic privilege with regard to our own institutional and artifactual kinds, protecting us from certain possibilities of ignorance and error; they also demonstrate that not all cases of reference to these kinds can proceed along a purely causal model. As a result, realist views in ontology, epistemology, and semantics that were developed with natural scientific kinds in mind cannot fully apply to the kinds of the social and human sciences. In closing I consider some wider consequences of these results for social science and philosophy.

Three elements of a realist philosophical world-view seem to go together: The ontological view that there are kinds of things that exist and have their nature independently of human beliefs, representations, and practices; the epistemological view that acquiring knowledge about such kinds is thus a matter of substantive discovery in the face of possibilities of gross error and ignorance; and the semantic view that reference to these kinds proceeds via a causal relation to an ostended sample, so that the extension of the term is determined by the real nature of the kind rather than by our associated beliefs and concepts, enabling us to refer to the kind despite our possible ignorance and error regarding its nature.

A general realist position, however, requires only that the realist hold that there are *some* things and kinds that exist independently of the mental—not that everything is independent.<sup>1</sup> Thus many realists are willing to accept that, along with independent natural kinds and objects, there are also (e.g.) institutional objects and artifacts that neither exist nor have their natures independently of all human beliefs, representations and practices. The idea that such objects differ ontologically from the realist's paradigm independent objects is usually taken as an unimportant triviality; e.g., Michael Devitt notes "The world that the Realist is primarily interested in defending is independent of us

---

<sup>1</sup> *Pace* Crawford Elder [1989], who argues that the realist must hold that beliefs about *all* components of the world may be massively false, and so must either deny that culturally generated kinds are constituted by social beliefs or deny their existence.

except in one uninteresting respect. Tools and social entities are dependent on us..." [1991, 249]. Similarly, John Searle describes realism as the view that if there had never been any representations, "Except for the little corner of the world that is constituted or affected by our representations, the world would still have existed and would have been exactly the same as it is now" [1995, 153].

I will argue, however, that given the ways in which realist ontological, epistemic, and semantic theses hang together, this is far from uninteresting. For although realist epistemological and semantic views are widely held to apply universally, they have been developed with independent natural kinds and objects in view. Focusing on two central sorts of object in the human world, namely institutional entities and artifacts, I will argue that as a consequence of the specific ways in which they differ ontologically from the natural objects and kinds of the realist's paradigm, realist epistemological and semantic views cannot apply universally or fully to them. As a result, we will require a substantively different ontology, epistemology, and semantics if we are to make sense of the objects studied by the social and human sciences as well as those of the natural sciences.

The plan for this paper is as follows. In section one I outline some principles of a widely held realist paradigm. I will not be concerned to defend or attack this view, either as a preferred statement of realism or as a view appropriate to the objects and kinds of the natural sciences; I hope merely to draw it out in a way that shows how the ontological, semantic, and epistemological realist theses for particular kinds of object are interrelated. In sections two and three respectively I will consider the cases of institutional and artifactual kinds (and their members), examining their forms of mind-dependence, and the differences this makes to our epistemic relation to these kinds and the reference of the associated terms. Finally, in sections four and five I discuss some wider consequences of these results for social science and philosophy.

### 1. The Realist Paradigm

The minimal core of ontological realism is the position that something exists independently of all mental states.<sup>2</sup> In its minimal form, realism is compatible with a sort of cheese-paring metaphysics that allows that what there is entirely lacks structure apart from what our beliefs or concepts

---

<sup>2</sup> The sort of dependence relevant to realism [Searle 1995, 156] and throughout this paper, is what is commonly called "logical" dependence, knowable a priori by analyzing the relevant concepts, not a mere causal or nomological dependence based in laws of nature (and discoverable a posteriori). Similarly, below when I speak of the dependencies of institutions and artifacts on collective acceptance or individual intentions, I mean these to be not causal claims about the creation of these entities, but logical claims, e.g., that the very idea of something being money presupposes collective agreement about what counts as money, and that the very idea of something being an artifact requires that it have been produced by someone with certain intentions.

impose upon it.<sup>3</sup> This version of ontological realism is, however, often thought to be too weak to support what might be called “epistemological realism”, the view that facts about the world are genuinely discoverable through substantive investigation subject to possibilities of confirmation and error, not just imposed by or to be read off of our concepts or beliefs. For if there were no pre-existing characteristics or structure of the world to discover, all supposed discovery would be mere imposition. A more robust realism is thus often thought to require that there be a world that not only exists, but also has a certain structure independently of the mental.<sup>4</sup> In that case, there must be diverse kinds that exist and have their boundaries quite independently of how our concepts and representations might happen to divide things up, in particular, independently of what we believe about the conditions relevant to drawing those boundaries. Such a picture allows hope of genuine discovery of facts about the structure of the universe and the nature of kinds in it that are not the simple imposition of our concepts on reality.

As noted above, the realist need only hold that there are some entities and kinds of entities meeting the above criteria, not that everything does, and so accepting the existence of entities or kinds that do not meet the above conditions is no threat to realism. Thus as well as the general (absolute) thesis of realism, it is useful to describe a relative thesis of ontological realism asserting that a certain kind, *K*, of objects meets these criteria.<sup>5</sup> In parallel to the general definition, there are two dimensions to an ontological realist thesis regarding objects of kind *K*. The first dimension, independence, may be expressed by an Independence Principle, asserting that things of kind *K* exist independently of the mental, that is, that it is possible that there are things that are of kind *K* and that there are no mental states whatsoever.

The second dimension lies in the idea that the kind *K* has natural boundaries; that is, that it is not merely a division artificially imposed on the world by human concepts. The idea that a kind *K* has natural boundaries may be expressed in ignorance and error principles. If a kind *K* has natural boundaries, then the conditions that determine whether or not something is of kind *K* are independent of whether or not those conditions are accepted by anyone. As a

---

<sup>3</sup> Michael Devitt [1991, 17] calls this “weak” or “fig-leaf” realism, and argues that it is of little interest.

<sup>4</sup> As Devitt puts it “A tiny move in the right direction is to say that the world consists not just of something, but of a structured set of entities” [1991, 17]. This realism is presupposed by, but still weaker than, the realism Devitt advocates as the view that “Tokens of most current common-sense and scientific physical types objectively exist independently of the mental” [1991, 23].

<sup>5</sup> On this understanding, denying a relative ontological realist thesis regarding things of kind *K* should not automatically be equated with denying that there are entities of kind *K*, or else we could not distinguish, e.g., conceptualism or constructivism from eliminativism. For arguments that one should accept that there are mind-dependent entities of various kinds, see my [1999], especially chapters 9 and 10.

result, for all conditions determining the nature of the kind K, it is possible that these remain unknown to everyone. Call this the “Ignorance Principle”. In addition, if a kind K has natural boundaries, since these boundaries are not determined by human beliefs about those boundaries, any beliefs (or principles accepted) regarding the nature of Ks could turn out to be massively wrong. Call this the “Error Principle”. The “could turn out” here is epistemic, not metaphysical—for of course if we happen to have a true belief about the nature of the kind (e.g., that gold is atomic number 79), as a necessary truth about the kind’s nature, it could not turn out to be wrong. Nonetheless, for all we know, we could turn out to be mistaken about that and any other propositions we accept regarding the nature of gold and all other kinds with natural boundaries. Say that a kind has natural boundaries just in case both the ignorance and error principles apply to it.

As the above principles make obvious, the issue of whether a certain kind K has natural boundaries is closely related to epistemological and semantic theses regarding the kind. A realist epistemology regarding a kind K requires that facts about the nature of the kind are not determinable by reporting or analyzing anyone’s concepts, but are potentially substantive discoveries subject both to possibilities of confirmation and error. Such an epistemology for a kind K thus presupposes that K has natural boundaries regarding which everyone’s beliefs are subject to massive error, and with a nature awaiting discovery precisely because it is opaque to everyone.

Similarly, causal theories of reference are based on the idea that there is a kind with pre-existing boundaries that can determine the extension of the term regardless of speakers’ beliefs and concepts regarding the kind. The idea that a causal theory of reference can apply universally to all cases of reference to a kind thus presupposes that there is a kind with natural boundaries that determine the extension of the term independently of anyone’s concept(s) regarding the kind.<sup>6</sup>

But one must be cautious, for properly speaking, the issue of what epistemic relation or theory of reference is appropriate for a given kind is relative not only to the kind, K, of entities, but also to a certain group, G, of people attempting to acquire knowledge of or refer to the kind. The above claims are only that if a realist epistemology or a causal theory of reference for a kind K is accepted universally with regard to *all* groups G, then that presupposes that K has natural boundaries. Nonetheless, causal theories of reference and a real-

---

<sup>6</sup> Moderated causal theories may (to avoid the qua problem in establishing reference) require that speakers have at least a high-level concept such as “same species as these”, but nonetheless, the extension is still determined independently of speakers’ substantive beliefs regarding what makes things of the same species as these, and so similarly presupposes that there is (e.g.) a species with boundaries independent of anyone’s having a substantive concept of the nature of this species. For discussion of the need for moderated causal theories, see Devitt and Sterelny [1999, 90-93].

ist epistemology may be appropriate for *some* groups G with regard to a certain kind without that kind having natural boundaries. The possibility of members of a group G making substantive discoveries about a certain kind presupposes that it exist and have its nature independently of G members' beliefs and concepts regarding its nature. That, however, does not require that it exist and have its nature independently of *everyone's* beliefs and concepts. Similarly, causal theories of the reference of a term "K" as used by a group G presuppose that the extension of "K" is determined by pre-existing boundaries of the kind with which the members of G are in causal contact, not on G members' associated beliefs or concepts regarding the nature of the kind. But this does not ensure that there is a kind with pre-existing boundaries independently of what *everyone* accepts regarding the nature of the kind. So epistemological realism and causal theories of reference may apply locally (relative to some groups G), but not universally, to kinds that lack natural boundaries.

The independence dimension and the natural boundaries dimension of ontological realism for a given kind are relatively independent, at least in principle. It is possible that there be artificially gerrymandered kinds with boundaries determined solely by what conditions we accept, but which conditions nonetheless can be fulfilled even in worlds with no mental states. Conversely, it is possible for things of kind K to depend necessarily on certain mental states, but not on the acceptance of certain beliefs about the nature of kind K, in which case it is in principle possible that a kind have natural boundaries in spite of its lack of independence.<sup>7</sup> As a result, it is in principle possible that a realist epistemology or a causal theory of reference be held universally for a kind K without accepting independence regarding that kind. Nonetheless, more typically independence and natural boundaries stand or fall together, as they do in the cases to be considered below, where the existence of things of kind K depends on the acceptance of certain conditions regarding the nature of Ks.

## 2. Institutional Kinds

I will begin by discussing those entities in the human world that are most obviously heavily dependent on human representations, namely institutional facts (e.g., that George W. Bush is president, that this is a \$1 bill) and institutional objects, concrete or abstract (a drivers' license, a law). John Searle's discussion of institutional facts in *The Construction of Social Reality* provides a useful starting place for discussing the ontological status

---

<sup>7</sup> Mental state kinds themselves might be plausibly considered examples of such kinds. For mental states certainly do not exist independently of the mental, yet at least according to many, mental state kinds are to be distinguished by the causal or functional roles of the state, regarding which everyone may be in ignorance or error. (Thanks to William Lycan for this example).

of institutional facts, although as we shall see it will require modification in certain ways.

While Searle [1995, 26] defines social facts as any facts involving collective intentionality, institutional facts, on Searle's analysis, are those social facts involving "collectively impos[ing] a function on a phenomenon whose physical composition is insufficient to guarantee the performance of the function, and therefore the function can only be performed as a matter of collective acceptance or recognition" [1995, 124]. Institutional facts clearly depend on mental states to (collectively) endow the objects involved (e.g., Bush) with new powers (e.g., the power to veto legislation). Thus the Independence Principle clearly fails for institutional entities. The difficulty, however, lies in specifying precisely the way in which that principle fails and determining whether or not that failure has any interesting consequences for our knowledge of and reference to institutional entities.

### 2.1 *The Ontology of Institutional Entities*

On Searle's account institutional facts are created by the collective acceptance of certain constitutive rules regarding the kind of fact in question. Constitutive rules are rules of the form "X counts as Y in context C", which themselves "create the possibility" of facts of kind Y [1995, 28]. Thus, for example, the collective acceptance of the rule "bills issued by the Bureau of Engraving and Printing(X) count as money(Y) in the United States(C)" creates the possibility of money, and endows the bills with powers (e.g., to serve as legal tender in business transactions) that they otherwise would have lacked. The existence of a social kind Y, on this view, thus requires the existence of Y-regarding collective intentions that take the form of constitutive rules for the existence of Ys. As a result, institutional facts exhibit what Searle refers to as "self-referentiality", that is, e.g., for a certain sort of thing to be money, it is necessary (and sufficient) that it be the sort of thing that is used as, regarded as, or believed to be money [1995, 32].<sup>8</sup>

In some cases the institutional facts must be created token by token through acceptance that that very individual has that institutional property (e.g., for any token person to be president, he/she must be sworn in and

---

<sup>8</sup> Searle actually claims that the feature of self-referentiality applies to all social concepts (not just institutional concepts). But if one accepts his definition of social facts as any facts involving collective intentionality, this is clearly not true. For some social facts (e.g., the bigotry of a certain population group) may involve collective intentionality (i.e. involving beliefs about the respective abilities of people of different races) without it being true that some group is bigoted if and only if they are believed to be bigoted. Dependence of a social kind K on collective intentionality does not entail the dependence of the kind K on intentional states *involving K-regarding beliefs*. Nonetheless, this feature clearly does apply to institutional facts, as considered by Searle, since in these cases the institutional fact consists of powers that exist if and only if they are collectively believed to.

regarded as president; for any couple to be married, they must accept that they are getting married and that must also be accepted by the marrying official and witnesses). In other cases, institutional facts may be created “wholesale” by simply accepting general rules such that anything that has the relevant features (e.g., is a bill of this pattern issued by the Bureau of Printing and Engraving) counts as having the institutional feature (being a \$1 bill). But the difference between these cases is merely superficial. In both kinds of case, the institutional facts are created not *simply* by accepting that some individual thing (X) counts as Y in the given context, but by accepting constitutive rules that specify general conditions, such that anything that fulfills them is (counts as being) of the relevant institutional kind. The difference is merely that in the token creation cases the specified conditions (X) include the requirement that the object must itself be accepted as having the relevant institutional feature (Y).

Thus we can say that institutional facts can exist only if there is collective acceptance of general principles (constitutive rules) that outline sufficient conditions for something to “count as” having the relevant feature.<sup>9</sup> (This places them in contrast to facts regarding natural features such as being a tree, since plausibly, e.g., it may be a fact that this is a tree without any collective acceptance of any principles regarding trees or anything else.) Once given that collective acceptance, anything that fulfills those conditions “automatically” acquires (or counts as having) the relevant institutional feature. So, e.g., money can exist only if the relevant society collectively accepts certain principles about what “counts as” money; having accepted those principles, anything with the features specified in the principles is money. Searle only lays out what he calls the basic “formula for constructing social reality”, “X counts as Y in context C”, but this neither captures the general character of

---

<sup>9</sup> H. L. A. Hart [1952] argues that legal concepts such as contract or property cannot be analyzed in terms of necessary and sufficient conditions, first because of the importance of precedent in establishing whether or not a particular case involves e.g., a contract, and second because application of the concept is always “defeasible” (e.g., if one has signed under duress). He uses this to argue that applications of legal concepts do not *describe* a pre-existing state of affairs definable in terms of physical necessary and sufficient conditions (a conclusion fully in accord with the above position), but rather *ascribe* a certain legal status by judgments based on accepted rules for such ascription. This is not in conflict with the above analysis, though the relevant accepted sufficient condition (which indeed ascribes rather than describes the nature of contracts) in this case must be stated carefully. Just as it is accepted that two people are married if they are declared married in the following circumstances..., so (in English law) it is accepted that something is a contract if it *would be* regarded as a contract by judgments in accord with the following rules (something is a contract if...unless...and provided it is sufficiently similar to preceding cases judged to be contracts...). (I say “would be” rather than “has been” regarded in this way to accommodate what Hart calls the “timeless” nature of judicial rulings.) Sufficient conditions still must be accepted for contracts to exist at all, though that does not require that the notion of a contract be reductively definable in terms of physical necessary and sufficient conditions.

the rules involved nor the dependence of institutional (Y-type) facts on human intentions. To exhibit these features we can lay out the following principle outlining the conditions for the existence of a concrete institutional fact, where K is an institutional kind.

**Dependence Principle 1 (Concrete institutional entities):**

DP1: Necessarily, for all x, x is K if and only if there is a set C of conditions such that it is collectively accepted that (for all y, if y meets all conditions in C, then y is K), and x meets all conditions in C.

The above principle applies only to cases involving imposing institutional facts on extant entities, so that some particular thing x meets conditions C and thus acquires the institutional feature K. This limitation results from the fact that we have begun from Searle's work on institutional facts, for Searle maintains that all social facts are created simply by imposing new social or institutional features on extant material objects:

Such material objects as are involved in institutional reality, e.g., bits of paper, are objects like any others, but the imposition of status-functions on these objects creates a level of description of the object where it is an institutional object, e.g., a twenty dollar bill. The object is no different; rather, a new status with an accompanying function has been assigned to an old object...[1995, 57]

This, however, does not seem to cover all kinds of institutional entities. For there seem to be institutional entities that are not based on applying new facts to particular material objects, but rather in creating new abstract institutional objects, such as laws or corporations. Laws, for example, are not created by directly applying a new institutional feature to an extant material object, but by allowing that under certain conditions, undertaking concrete activities (writing characters, raising hands, etc. in the right context of a legislative body) counts as creating a new entity, a law. The law created is not itself identical with (or materially constituted by) any of these concrete activities or instances of it, so it cannot be accounted for merely in terms of new properties applied to old material objects.

To account for these among the institutional entities, we need a broader principle to express the dependence of institutional objects on the acceptance of certain constitutive rules without requiring that those rules be applied to pre-existing objects, endowing them with new properties. The following principle applies in these cases:

**Dependence Principle 2 (Abstract Institutional Entities):**

DP2: Necessarily, there is some x that is K, if and only if there is some set of conditions C such that it is collectively accepted that (if all



conditions in C are fulfilled, there is something that is K) and all conditions in C are fulfilled.

## 2.2 *The Epistemology of Institutional Entities*

In both the abstract and concrete case, entities of institutional kinds K fail to meet the Independence Principle. Moreover, K-type entities are not merely dependent on mental states in general; they depend on human representations in quite specific ways, namely on the collective acceptance of certain principles that describe sufficient conditions for the existence of Ks. Thus clearly, here the Independence Principle fails. While this is unsurprising, what has gone rather unremarked upon is the fact that the particular way in which the principle fails in these cases (based on principles DP1 and DP2 above) leads to the failure of the Ignorance and Error Principles, bringing about the need for quite different understandings of reference and epistemology than those accompanying the standard realist view.

### **Case (1): Concrete Institutional Entities**

It is easy to see that the Error Principle fails for concrete institutional entities, understood on the above model. For suppose we collectively accept that, for any y, fulfilling certain conditions C is sufficient for y to be K. In such a situation, it could not turn out that we are wrong, that being C is not really sufficient for being K. For the dependence principle DP1 ensures that if we accept those conditions, and they are fulfilled, then the entity in question is thereby a K. In short, from DP1 above it follows that:

#### **Epistemological Principle 1 (Concrete Institutional Entities):**

EP1: Necessarily, for all sets of conditions C, if we collectively accept that (for all y, if y meets all conditions in C, then Ky), then for all x, (if x meets all conditions in C, then Kx).<sup>10</sup>

Our acceptance of a set of conditions C as sufficient for being K is constitutive of what conditions suffice for being K, so what conditions there are is

---

<sup>10</sup> This and other principles for institutional entities are expressed in the present tense, since a thing belongs to an institutional kind at a time only provided the ongoing acceptance (at that time) of the relevant constitutive rules. For, e.g., cowry shells 'count as' money in a given society only as long as people continue to collectively accept that cowry shells have a certain exchange value. If there is a sudden shift to counting pieces of gold rather than cowry shells as monetary units, then cowry shells cease to have the institutional feature of being money. (Of course the historical fact remains that they *were* ancient monetary units, and about that all current users of money may be mistaken.)

determined by what conditions we accept. As a result, we could not turn out to be mistaken—our acceptance of the set of conditions *C* declaratively establishes the conditions for being *K* rather than attempting to describe pre-existing and independent conditions for being *K*. So the Error Principle fails: Any conditions we accept as sufficient for the existence of *Ks* must be free from error, and thus it is not the case that whatever conditions we accept could turn out to be massively wrong.

Similarly, given the above Dependence Principle 1, there can be no universal ignorance of all conditions relevant to the nature of *Ks*, *provided there is at least one K*. For the existence of some *x* that is *K* requires the collective acceptance of a set of sufficient conditions for the existence of *Ks* (and as we have seen above, these must hold).

### **Case (2): Abstract Institutional Entities**

The same pattern applies in the case of abstract institutional entities. For from DP2 it follows that:

#### **Epistemological Principle 2 (Abstract Institutional Entities):**

EP2: Necessarily, for all sets of conditions *C*, if we collectively accept that (if all conditions in *C* are met, then there is some *y* such that *Ky*), then (if all conditions in *C* are met, then there is some *x* such that *Kx*).

Thus the Error Principle again fails, for if we collectively accept any set of conditions *C* as sufficient for there being a *K*, it could not turn out that we are wrong. For, as the dependence principle DP2 ensures, given that acceptance, if the accepted conditions are fulfilled, there is some *x* that is *K*. The conditions collectively accepted about sufficient conditions for the existence of a *K* are protected from error since it is that acceptance that establishes the relevant conditions.

Similarly, DP2 entails that, if there is something of kind *K*, there cannot be complete ignorance regarding the nature of *Ks*. For the existence of something of kind *K* presupposes that there are some principles accepted regarding (at least) sufficient conditions for the existence of *Ks*, and (by the above argument) these must be true.

In sum, the ontological status of institutional entities has significant epistemological consequences. If we understand institutional entities as dependent on the acceptance of certain constitutive rules laying out (at least) sufficient conditions for their existence, and existing provided something fulfills these conditions, we cannot conceive of investigations into the nature of our own institutional kinds as completely a matter of substantive and fallible discov-

ery. Whereas natural kinds (on a realist view) can exist even if no one knows of their existence or any facts about their nature, institutional kinds do not exist independently of our knowing something about them. Similarly, whereas, in the case of natural kinds, any substantive principles any individual or group accepts regarding the nature of the kind can turn out to be wrong, in the case of institutional kinds those principles we accept regarding sufficient conditions for the existence of these entities must be true.<sup>11</sup> We are guaranteed freedom from complete ignorance and are preserved from error in many of our beliefs regarding the nature of institutional entities precisely because the principles accepted play a stipulative role in constituting the nature of the kind.<sup>12</sup>

This special epistemic relationship should not be exaggerated, however. First, freedom from error holds only for sufficient conditions accepted; it is not true, however, that any merely necessary condition we accept must in fact hold. Thus not all beliefs we may hold about the nature of institutional entities are free from error on this understanding of the ontology of institutional entities. Of course other beliefs we may happen to hold (rather than accepting as criterial) regarding institutional entities may likewise turn out to be false. Thus the above conclusions certainly do not demonstrate that we are entirely protected from error in our complete set of beliefs regarding these things.

Secondly, all that is ensured is a freedom from *universal* ignorance; the very existence of a thing of kind K ensures only that there is *some* true principle that is collectively accepted, not that all true principles are accepted. For example, the dependence principles DP1 and DP2 themselves constitute conditions for the existence of members of the kind K, but these principles may not themselves be accepted by anyone. Thus, as Searle points out [1995, 118], although institutional facts are created by the collective acceptance of constitutive rules, that fact may itself not be recognized or accepted by the

---

<sup>11</sup> Certain beliefs about natural kinds might be *unlikely* to turn out false, e.g., beliefs in a placeholder principle such as that “the nature of these kinds is whatever experts eventually claim it is”. But in this case, unlike the case of institutional kinds, the logical possibility always *remains* that this (along with the expert’s beliefs themselves) could turn out to be wrong since in this case, on a realist view, such beliefs do not constitute, but merely attempt to report on, the nature of the kind. (If one replies that anyone who got it wrong is not *really* an expert, so that the layperson’s placeholder belief and the *real* expert’s substantive beliefs are guaranteed freedom from error, this is not substantive knowledge about the kind but mere knowledge of a tautology.)

<sup>12</sup> Some might urge that in fact we seldom have explicit cognitive awareness of the relevant principles for institutional kind membership, we just have the practice of accepting certain sorts of things and rejecting others as putative kind members. I have spoken in terms of the acceptance of principles in order to make the logical relations clearer, but the basic points can easily be preserved in a less explicitly cognitivist scheme. The results in that case would be that even if (on a realist view) certain kinds of massive error in treating entities as members of a certain kind are possible for natural kinds (e.g., treating whales as fish), the same is not true for practices involving institutional kinds (e.g., treating cowry shells as money). (Thanks to Terence Horgan for raising this issue.)

people in the relevant society, who may believe instead that the king or laws of the land are established by God rather than by their own collective acceptance of certain principles regarding conditions for something being the king or a law.

Finally and most importantly, these epistemic results apply collectively to (and only to) the group whose collective acceptance is referred to in the dependence principle for the kind in question. Since collectively accepted principles are the basis for this epistemic privilege, it is primarily collective privileged knowledge that results; how this relates to individual knowledge of group members depends on what the right account is of the relation between individual and collective acceptance, an issue that cannot be settled here.<sup>13</sup> Other societies may be in complete ignorance or error in all of their beliefs regarding the nature of our institutional kinds, and any knowledge such outsiders may acquire of the nature of these institutional kinds is a matter of substantive and fallible discovery. Since the nature of these kinds does not depend on any of their own beliefs or concepts regarding that nature, their epistemic relation to these institutional kinds parallels the relation everyone has to natural kinds, despite the fact that the former lack natural boundaries.

### *2.3 Reference to Institutional Entities*

As is clear from the above, for our own institutional kinds *K*, at least some of the conditions that determine whether or not something is a *K* must be accepted, and conditions we collectively accept as sufficient thereby are so. As a result, a purely causal theory of reference cannot apply to institutional kind terms as used within the group whose collective acceptance creates and maintains those kinds. Reference to institutional kinds cannot be made without any concept, enabling the extension to be determined solely by the independent nature of the kind. Nor, in such a context, can reference to institutional entities even proceed along the lines of a moderated causal theory that (to avoid the qua problem) recognizes the need for at least minimal concepts such as “whatever is of the same species as this (is a lion)”. In such a context, there is nothing of the institutional kind to refer to unless certain substantive principles are accepted regarding (at least) sufficient conditions for the existence of something of the kind, and those principles cannot turn out to be mistaken. That concept collectively accepted by the group functions at least

---

<sup>13</sup> If a group may collectively accept some principle without every member of the group accepting it, then any given member of the group could fail to know the principles relevant to being a member of the kind. Nonetheless, surely most members must accept the relevant principles (at least qua group members) to secure collective acceptance, and in any case the collective knowledge possessed marks a sharp difference from the case of natural kinds. Note, however, that a group member may know (accept) that cowry shells are money without knowing the higher-level facts that they are money only because there is collective acceptance of the fact, and that there is such collective acceptance.

in part stipulatively to establish the nature of the kind, and thus the extension of the term. Here sense does help determine reference, not by substituting a nominal nature for a metaphysical one, but rather by establishing the metaphysical nature that in turn determines the extension of the kind.

As above, however, these results only apply within the community upon which the institutional entities in question ontologically depend. Once an institutional kind is established in a given community, reference to it by those outside the sustaining community can proceed along a moderate causal model. A foreigner could point at a pile of dollar bills (suspecting their institutional nature) and coin a term to refer to “whatever has the same institutional features (as these things do in their native community)”,<sup>14</sup> but that cannot be the sole means of reference to institutional kinds.

### 3. Artifactual Kinds

A much broader category of everyday objects in the human world is that of artifacts. Although some kinds of artifacts may be institutional (as, e.g., drivers licenses and dollar bills are), many more (tables, teacups, screwdrivers) do not require the collective acceptance of constitutive rules laying out (at least) sufficient conditions for their existence, and so issues regarding their ontology, epistemology, and the reference of their terms must be separately examined. In this section I focus attention precisely on those non-institutional artifacts.

Not just anything created through human efforts is an artifact. As Risto Hilpinen has pointed out [1992, 59-60], artifacts properly so-called are *intended* products of human actions, to be distinguished from other artificial entities such as scrap material that may also be produced through human activities. Thus artifacts, as such, lack independence, for their existence depends on that of certain human intentions.

More precisely, an object is an artifact in the strict sense of the word “only if it is intentionally produced by an agent under some description of the object”, where at least one of these descriptions must be a *sortal* description [Hilpinen 1992, 60]. Many words that we might pre-reflectively classify as artifactual kind terms such as “path” and “village”, although they plausibly can only literally apply to certain products of human activity, need not refer only to artifacts, since some things in their extension may not be things intentionally created under some description. A path, for example, may be created through a series of human intentions to take the shortest route across a field between two settlements (not through an intention to make anything, though this can also occur), and a village may grow out of diverse human intentions to create many houses, shops, and streets without anyone having

---

<sup>14</sup> Following Hilary Kornblith’s parallel argument [1980, 114] for the case of artifacts, that a Martian could successfully refer to our doorstops without a substantive concept.

an overarching intention (under any description) to create the collective object.<sup>15</sup> In those cases where such collectively produced objects do not result from an intention to create them under any sortal description, Hilpinen argues, they are, strictly speaking, not artifacts at all, although they are the (indirect) products of human activity [1992, 66].

I will limit my discussion here to those artifactual kinds corresponding to what may be called “essentially artifactual” terms—that is, terms that necessarily have in their extension all and only artifacts, considered as such (as *intended products of human action*). Thus terms such as “village” may have many artifacts in their extension without being essentially artifactual, since non-artifacts are also in the extension; other terms such as “seven inch aluminum cylinder” may have only artifacts in their extension, but as long as that is not necessarily the case, these are not essentially artifactual terms. Similarly, terms such as “seconds” (referring to mis-made products, not units of time) are not essentially artifactual, since although they can only refer to artifacts (intentionally produced under *some* description), they do not refer to them qua intended products, but qua an unintended feature of those products. For referents of any essentially artifactual term “K”, being K must be an intended property of the object.<sup>16</sup>

### 3.1 *The Ontology of Artifacts*

We have seen that if there were no human intentions, there could be no artifacts, however it is less clear whether their dependence on human intentions prevents artifacts from being understood as kinds with natural boundaries that may determine whether or not something is of that kind independently of all beliefs and intentions regarding the nature of the kind. Indeed, there has been substantial controversy surrounding this issue [Schwartz 1978, 1980; Kornblith 1980; Nelson 1982; Elder 1989] largely because of its relevance to the issue of whether or not the causal theory of reference can apply to artifactual kind terms. (I return to that semantic issue in 3.3 below.)

Among those who defend the idea that artifactual kinds do have a nature capable of independently determining the extension of the corresponding terms, by far the most popular view is that the nature of artifactual kinds lies in a common function. This view may be interpreted either as referring to a common set of *causal/functional capacities* possessed by the object, or a *common intended* (or “derived proper”) function.<sup>17</sup> The first interpretation is

---

<sup>15</sup> The path example and the term “essentially artifactual” (below) are Hilpinen’s, from conversation. The village example and the problem of collectively produced cultural objects are discussed in Hilpinen [1992, 66].

<sup>16</sup> For discussion of intended properties see Hilpinen [1995, 138-9].

<sup>17</sup> The term “derived proper” function is Ruth Millikan’s. In the case of artifacts, the derived proper function and intended function coincide, as Millikan notes “artifacts have as derived proper functions the functions intended for them by their makers” [1999, 205].

not promising, for if we attempted to group artifacts together in terms of their actual causal/functional capacities, broken or deformed screwdrivers could not be classed as members of their kind, and metal snow-coasters shaped like woks and enjoying the same causal capacities would be classed as belonging to the same artifactual kind as those cooking devices. Classifications based on actual causal/functional capacities do not come close to providing appropriate extensions for artifactual kind terms.

Classifications based on *intended* function seem to have a much greater plausibility, for they can classify working and non-working screwdrivers together based on their intended function, and similarly distinguish woks from snow coasters. But even on this preferable interpretation, functional analyses cannot provide a general analysis of the nature of artifactual kinds. First, some artifactual kinds (such as works of art) may have no (or no common) function. Second, some tokens of particular artifactual types (e.g., some chairs) may not be intended to serve the standard function of the kind. Something can be a boat or a chair even if its maker desires that it never be placed in water or sat upon (perhaps intending it only 'for show') [Bloom 1996, 5-6; Kobes 2000]. In such cases, the artifacts in question would not have the relevant intended function, although they would be classified as proper boats/chairs according to our ordinary uses of the terms. While the first objection suggests at worst that some artifactual kinds cannot be functional, the second has the far more damaging implication that even in the apparently best cases for the functional view, sharing an intended function cannot be essential to the nature of an artifactual kind.

A second proposed view of the nature of artifactual kinds is what Paul Bloom [1996] calls the "intentional-historical" theory, on analogy with Jerrold Levinson's [1979, 1989, 1993] "intentional-historical" theory of art (according to which the essential feature of art is that the object created be intended to be regarded as earlier art is or was correctly regarded). Bloom (in contrast to Levinson, who explicitly distinguishes the concept of art from concepts of other artifacts [1979, 234]) argues that the core of all artifact concepts may be captured in terms of the creator's intentions to create something belonging to a certain extant kind. Thus, e.g., chairs are "all and only those entities that have been successfully created with the intention that they belong to the same kind as current and previous chairs (or, equivalently, that they be chairs)", and in general "We construe the extension of an artifact kind X to be those entities that have been successfully created with the intention that they belong to the same kind as current and previous Xs"[1996, 10].

Bloom is concerned mainly with the psychological issue of what criteria we actually use in classifying artifacts, not with the ontological issue of what the nature of artifacts is. Nonetheless, it is easy to formulate the corresponding ontological thesis as the view that an item is of an artifactual kind K if

and only if it was successfully created with the intention that it be of the same kind as current and previous Ks. This proposal enables us to include both functionally based and non-functionally-based artifacts, and to include e.g., decorative chairs as members of the kind "chair" provided they are intended to be chairs (even if they are not intended to be sat upon).

Bloom's requirement that the created object be successfully created "with the intention that it be of the same kind as current and previous Xs" is ambiguous, however. Is this intention to be understood transparently, as indexically tied to a certain extension of objects (the Xs) and intending merely to make something of "the same type as these" (whatever they are), or opaquely, as an intention to make something of kind X, where this is understood intensionally as a kind with a certain nature? Although he does not address this question directly, it seems that Bloom wants and needs the former interpretation. For he touts as a main advantage of his analysis that it enables artifactual kinds to be understood analogously to natural kinds, on a Kripkean view, and artifactual kind terms to refer just as natural kind terms do, on a causal/historical analysis [1996, 22-25]. In both cases, he argues, whatever stereotypical features or prototypes anyone has in mind are irrelevant to the real extension of the concept. This would not be the case, however, if makers were required to have a certain contentful intention (and successfully execute that) in order to create an X, for then (as we shall see below) their concept of the nature of an X would be relevant to the extension of the term. If a merely transparent intention to create "one of these" is all that is required, then artifactual kinds may be understood analogously to biological kinds, as determined by historical relations among kind members, not by anyone's concepts or superficial stereotypes of relevant features. If the analogy were complete, that would (despite the dependence of artifacts on human intentions) permit artifactual kinds to have natural boundaries and enable us to use artifactual kind terms successfully without requiring of anyone any (or any correct) substantive associated concept.

But if we understand the relevant intention transparently, then Bloom's suggestion cannot provide either a necessary or sufficient condition for membership in an artifactual kind. It cannot be necessary since we must allow, for example, that prototype knives are also knives, although there is no group of objects such that the maker of the first knife intends this to be of the same kind as those. This prototype problem is multiplied by noting that (unlike biological species) members of the same artifactual kind may be created in entirely independent historical chains (beginning from independent prototypes). If we required that, for anyone to create a knife after the first knife-maker, he/she must intend it to be as the same kind as *that one* (the original



prototype) or *those* (extant knives), we could not allow that separate cultures could all produce knives.<sup>18</sup>

Mere transparent intentions (executed to the satisfaction of their possessor) also cannot provide a sufficient condition for membership in an artifactual kind. For as Bloom himself admits, if a maker merely has the intention to create “one of these” and lacks an adequate concept of what is involved in being of that artifactual kind, he or she may well fail to make an artifact of the relevant kind. Although in general one can “make” a penny into a chess pawn simply through intending it to be a pawn, “someone who does not know the rules of chess could not do this, as she lacks the right understanding of what it is for something to be a pawn” [1996, 18]. She cannot simply make a pawn by having sensory contact with a bunch of pawns and intending this to be “one of them”; a substantive concept is required.

Another of Bloom’s own examples shows that, moreover, a would-be maker must have the *right* concept:

Imagine a madman who creates a tiny pile of dirt, assuming that people will happily sit on it, and he states that this pile was successfully created with the intention to be a chair. Still, we would not view it as a chair...[Our response would be parallel to that to] a 2-year-old who creates a flat disk out of clay and claims that it is a cup [1996, 19-20].

Since Bloom is primarily interested in the psychological issue of how we actually make judgments about whether or not something is of a particular artifactual kind, he claims only that this shows that *believing* that the maker intends to make a K is not sufficient for *believing that* her product is a K. Put in ontological terms, the objection is that the intention to create something of kind K (even if regarded as successful by the creator)<sup>19</sup> is not sufficient for creating something that actually is of kind K.

Bloom handles this objection by noting that, in these cases, the relevant intentions to produce a K are not actually present, because the creator does not have the same concept of Ks as ours:

All of these examples share a certain property, however. When the madman describes a pile of dust as the successful result of an attempt to create a chair, it is clear that his understanding of chairs is quite different from our own, perhaps so much so that it is not actually correct to say that he is in an intentional state that makes reference to the same concepts that we possess. Similarly, we are likely to infer that the child who calls a disk “a cup” does not really know what a cup is, or is perhaps confusing the words “cup” and “plate” [1996, 20].

---

<sup>18</sup> Levinson notices this problem for the case of art and, to avoid it, allows either transparent or opaque intentions to suffice on his theory [1979, 237-8].

<sup>19</sup> It is hard to see what other success criteria could possibly be given here, since the creator may, according to this model, have no contentful concept against which the product can be measured for success.

Thus clearly having a mere intention to create something of the same kind as current and previous Ks, if those are considered merely transparently, is not sufficient to create a K (even if the creator regards that intention as successfully executed).

We may find a better view, however, by pursuing the other interpretation of Bloom's suggestion—that something is of artifactual kind K if and only if it is the product of a successful intention to create something of kind K, where the relevant intentions involve a substantive intensional concept of the nature of things of kind K rather than a transparent reference to an historical sample. This will enable us to handle makers of prototype artifacts and those of reproductions together, for there is no need in either case for their intention to be tied to a pre-existing historical sample, although the latter may be.

Consider first the case of a would-be maker of a prototype of a new kind of artifact, K. For that maker to be able to really intend to make a K, she must have a substantive notion of what a K is supposed to be, determining the success criteria for her creative activity. "What are you making there?", "A whizool", "What's a whizool?" "One of these" is only a bad joke, and (seriously uttered) would lead us to doubt that the supposed creator is really involved in the intentional making of any kind of thing (as opposed to just "messing around").<sup>20</sup> The intention to make something of kind K thus must be based on a declarative intention associating that kind with a number of criteria that would constitute success at creating a K, and involving a number of K-relevant features that the inventor intends to impose on the object in order to succeed at producing a K. This fits in naturally with Hilpinen's emphasis [1992, 64] on the fact that genuine artifacts must have a number of intended properties, thus requiring makers to have a structured intention regarding a number of properties they intend to impose on the object (thus not just a bare intention to create "one of these"). Later makers may acquire the concept K through acquaintance with the initial prototype or with later copies, or they may independently arrive at it, but as long as their intention shares the same content as the original, and they intend to impose the same K-relevant features, they may be said to have the required intention to create a K.<sup>21</sup>

Of course, those features criterially associated with being of the relevant kind K need not be the only intended features of the object (one may intend to create something that toasts bread, and also intend this one to have curvy

---

<sup>20</sup> This does not of course imply that creators must always begin with a complete mental plan for their creation. Naturally ideas can form and evolve in the course of "messing around"; all that is required is that at some stage during the creative activity the maker develop a substantive concept of what is to be made, such that the final product can be evaluated for success.

<sup>21</sup> Of course, later makers may tie this concept to a different word; speakers of different languages may all create chairs provided they share the same concept of a chair.

sides, without the latter being criterially associated with the artifactual kind “toaster” created). It should also be noted that what is involved in the K-relevant features criterially associated with the concept is left quite open and flexible. These K-relevant features may be functional, structural, aesthetic, historical, or of various other sorts or combinations. They also need not involve a strict set of necessary and sufficient conditions (though they may); they may instead be formed as a cluster concept or perhaps via other means. All that is important is that there be some contentful concept that involves (perhaps vague) success criteria for creating an artifact of kind K, which can then direct the intention to produce a K via the imposition of certain K-relevant features.

But, whether in the case of inventors or later makers, such an intention alone is clearly not sufficient for the production of a K, for that intention can go terribly wrong in production. Bloom’s condition thus requires not only that there be such an intention, but that the object be the product of a “successful intention”, where this success condition “expresses the constraint that the entities must turn out as they were intended to turn out” [1996, 10]. We can say that the intention to create a K is successful to the extent that the intended K-relevant features are imposed on the object as a direct result of these creative activities. But although some degree of success is surely required, we should not require that an intention be so completely successful that all of the intended K-relevant properties are embodied in the object as a result of those intentions—this would rule out any even slightly deformed or mis-made products from being artifacts of the relevant kind. If I intend to make a cup (and thus intend to impose such features as that the object be a flared hollow cylinder with a base, capable of holding liquid), yet what comes off the wheel is a shower of damp clay flung onto the walls, I have not succeeded in making a cup at all. But if what comes off the potter’s wheel is a roughly flared hollow cylinder with a base that has a gap that prevents it from holding liquid, I may have succeeded in making a deformed cup. We should require, then, only that the maker’s intentions be *largely* successful if they are to create something genuinely of kind K, but the success need not be complete, for there may be broken or deformed artifacts that are nonetheless (lousy) members of their kind.<sup>22</sup> The requirement that the intention be “largely” successful introduces some vagueness to the definition, but that is only appropriate, for such vagueness is certainly part of the way our ordinary artifactual kind terms work—artifactual kind terms notoriously have vague application conditions (thus making them subject to sorites paradoxes and the

---

<sup>22</sup> This parallels Hilpinen’s general success conditions for the production of an artifact, which require that “an agent produces a genuine artifact only if his activity is *successful* in some respect and to some degree” [1993, 160].

like).<sup>23</sup> There may be borderline (and borderline borderline) cases of cups, knives, boats, etc., in those cases where the execution is somewhat faulty. Any non-revisionary analysis of our ordinary artifactual kind terms and their referents must incorporate this vagueness.

Thus we can draw out a dependence principle for artifactual kinds, emending Bloom's, as follows:

**Dependence Principle 3.1 (Strict Artifactual Kinds):**

DP3.1: Necessarily, for all  $x$  and all strict artifactual kinds  $K$ ,  $x$  is a  $K$  if and only if  $x$  is the product of a largely successful intention that  $(Kx)$ , where one intends  $(Kx)$  if and only if one has a substantive concept of the nature of  $Ks$  that matches that of prior makers of  $Ks$  (if any) and intends to realize that concept by imposing  $K$ -relevant features on the object.

This dependence principle allows in the products of makers of prototypes, later makers, and coincidental makers of the same artifactual kind.<sup>24</sup> It also includes somewhat deformed and mis-made members of artifactual kinds. It should be noted that DP3.1 is far more open-ended than DP1 and DP2, since it may involve concepts of  $K$ -relevant features of all kinds, and needn't specify a simple set of sufficient conditions for being a  $K$ . It also (unlike DP1 and DP2) does not require any *collective* acceptance of  $K$ -relevant conditions, but only individual acceptance by each maker. Each maker may have the individual intention to create a thing with certain characteristics, and all coincidentally make chairs, without there being any collective belief regarding the nature of chairs. Although artifacts depend on intentionality, they need not be the products of collective intentions nor depend on collective beliefs, and thus (if one accepts the common definition of social entities as those involving collective intentionality [Searle 1995, 26]), they need not be properly social objects in the way that institutional entities must be.<sup>25</sup>

---

<sup>23</sup> This feature leads some to deny the existence of artifacts and the referents of other ordinary terms with vague application conditions. See, e.g., Unger [1979], Heller [1988]. I shall leave criticism of this response to the vagueness of ordinary terms (including artifactual kind terms) for another occasion.

<sup>24</sup> In the age of mechanical reproduction, the group of makers need not be limited to those physically forming the objects. Makers should be understood as those responsible for production, whether or not they are directly physically manipulating the object (since that may be done purely mechanically). They may thus include designers of manufacturing machinery and consumers who place orders for products under a quite specific description, even if they are not the ones who execute the orders.

<sup>25</sup> This runs counter Searle, who takes something's being a bathtub or a screwdriver to be a social fact. Since it seems that artifactual kinds need not involve *collective* intentionality, I have hesitated to call them "social kinds" and have spoken throughout instead of problems with accounting for "human" kinds.

I have labeled this the dependence principle for “strict” artifactual kinds because the kinds it marks out all follow a single concept uniformly accepted by the makers. This seems apt for many of the strict artifactual kinds designated by experts (tailors, architects, chefs, etc.) involved in making the artifacts (e.g., double-breasted waistcoat, split-level, Peking duck), with strict criteria that must be known by those in the business and closely reproduced.

Certain layman’s artifactual terms, however, have much looser application. It has been frequently pointed out [Bloom 1996, 14; Kobes 2000] that artifactual kinds corresponding to ordinary terms such as “bottle” or “key” can undergo great changes over a period of historical development, as e.g., bottles have evolved from a classic glass milk-bottle to the ergonomic plastic baby bottles of today, and keys have evolved from metal cylinders with intricate flat ends to include the plastic hotel room keys of today. The variations are no doubt even greater with regard to such artifactual kinds as items of clothing, with e.g., modern dresses no doubt lacking a great many of the properties that would have been considered K-relevant by Victorian dressmakers. Since the above strict definition required that later makers’ concepts of Ks must *exactly* match that of prior makers, that would preclude any innovations and developments regarding the properties considered K-relevant (though it notably would not preclude all changes, since the concept of K shared by makers would typically be highly incomplete and thus allow great variation in other intended features without a difference in artifactual kind).

Thus it is preferable, in dealing with more loosely defined artifactual kinds, to require only that later makers’ concepts of the kind K *largely* match that of earlier makers, allowing room for the concept associated with an artifactual kind to gradually drift, just as we required only that the imposition of the relevant concept be *largely* successful. Thus we can lay out an altered dependence principle to cover both strict and loose artifactual kinds as follows (significant changes from 3.1 are italicized):

**Dependence Principle 3.2 (Artifactual Kinds):**

DP3.2: Necessarily, for all x and all artifactual kinds K, x is a K *only if* x is the product of a largely successful intention that (Kx), where one intends (Kx) *only if* one has a substantive concept of the nature of Ks that *largely* matches that *of some group* of prior makers of Ks (if there are any) and intends to realize that concept by imposing K-relevant features on the object.

Since DP3.1 entails DP3.2, the latter is clearly the weaker principle, and applies to all artifactual kinds, whether strict or loose. Weakening the requirement to specify only that the maker’s concept must “largely match”

that of some prior group of makers enables this principle to allow for a gradual shift in the concept of K-relevant properties shared by makers, while preserving a certain continuity in the concept held by makers. It does not, however, allow for sudden ruptures in the concept and relevant features of the kind, and thus still successfully rules out products of the madman and child of Bloom's examples, for they clearly fail to have a concept that matches that of some group of prior makers of chairs and cups, and thus cannot be said to have the relevant intention. On the other hand, it introduces an additional source of vagueness beyond that in DP3.1, which invites certain dangers and difficulties that must be dealt with. First, problem cases could arise if a maker's intention differed somewhat from that of prior makers, and the execution lived up mainly to the idiosyncratic aspects of the maker's concept, leaving us with an object that might be so distant from immediately prior Ks that it would be doubtful whether it was a K at all. Second, over a long period of time the concept (and corresponding kind) can gradually change so much that it is unclear whether or not we should really count items at the beginning and end of the series as being of the same artifactual kind. Such may be the case, e.g., between early adding machines and the modern computers that have arisen through a continuous history of development from that origin, but that now are counted (in ordinary language at least) as being of different artifactual types.

Finally, another danger that arises is that the concept may change and develop in two or more different directions. Consider the case of knickers. The original artifacts popularized in the 1850s were knickerbockers, loose-fitting men's breeches gathered at the knee. These gradually developed in two directions (as the word was shortened to "knickers"); on the one hand, towards sporting garments and ultimately women's fashion items of similar form; on the other, towards loose knee-length undergarments that gradually shortened to become the women's undergarments of today. Here, although we have the same word, it refers (perhaps ambiguously) to two groups of artifacts radically different in form and use, and we might well hesitate to group these in the same artifactual kind.

What are we to say about such cases? What marks the breaking point of change in artifactual kinds? Our above dependence principle is silent on this issue, because (precisely for this reason, and unlike DP1, DP2 and DP3.1) it lays out only a necessary, not a necessary and sufficient condition, for membership in an artifactual kind K. Thus we are not forced to conclude that an item badly made according to an idiosyncratic concept belongs to its intended kind, that computers and adding machines are of the same kind, nor that modern knee-length trousers are the same kind of artifact as petite undergarments. All this is left open by the definition, which is just as well, since our actual practices in determining what counts as being of one artifactual kind and what

as being two seem to vary greatly in terms of how much change over time or difference at a time can be tolerated while producing artifacts of the same kind. The facts are clear regarding the development of a kind and the shifts or differences in the corresponding concepts; where the line is drawn between kinds is a relatively boring issue. Certainly the differences marked by our terminology and classificatory practices seem to depend largely on our interests (e.g., a tailor might classify jackets in several dozen significantly different categories, where the average consumer would count all as the same artifactual kind), so any strict necessary and sufficient conditions laid out would be unlikely to map our ordinary conceptual and linguistic practices. The above dependence principle does, however, at least lay out a necessary condition that should apply to all artifactual kinds.

### *3.2 The Epistemology of Artifactual Kinds*

The epistemological consequences vary significantly based on whether one is speaking of strict or loose artifactual kinds, although in either case we clearly have a closer relationship to artifactual kinds than to natural kinds. For strict kinds, the existence of anything of the kind K entails that there is a unified concept of Ks accepted by makers. Moreover, that concept is protected from massive error, for DP3.1 ensures that something can be a K only if the intention to make something that meets the shared concept of K is largely successful, and thus only if the product largely realizes the concept. It should be noted, however, that this privileged epistemic position only applies to makers (conceived broadly); other individuals outside the sphere of production may be entirely ignorant or in error regarding the existence and nature of the artifactual kind.

We lack any guarantee of such a close epistemic relationship with loose artifactual kinds, since the conditions relevant to being of such a kind may change so greatly over time and vary so greatly in competing traditions. Nonetheless, even loose artifactual kinds fail to have natural boundaries in the sense required for realism. For that requires that all conditions that determine whether or not something is of kind K are independent of whether or not those conditions are accepted by anyone. But that is not the case even for loose artifactual kinds. For something is a K only if it largely matches a substantive concept of some group of makers (minimally, its own). Thus if a K exists, there cannot be universal ignorance of K-relevant features; someone (at least the maker) must have a substantive concept of Ks, and that thing must largely fulfill that concept. Moreover, if there are prior K-makers, that maker's concept must largely match that of some group of prior K-makers, and that group's shared (perhaps partial) concept of Ks is not subject to massive error regarding the Ks of that time and tradition. This preserves the idea that makers have some privileged knowledge of K-relevant features for Ks of

their own time and tradition, but also (as it should) acknowledges increasing scope for error the further from one's time and tradition one goes.

However, both of these results (for strict and loose artifactual kinds) might be thought to be threatened by frequent arguments that all beliefs about the natures of all artifactual kinds could turn out to be massively mistaken. James Nelson [1982, 362] argues the point by asking us to imagine a case in which pencils are not writing implements (as we had thought) but rather alien devices used to monitor human activities, where the aliens have created the systematic illusion that these things are used and useful for writing.<sup>26</sup> But it does not here "turn out" to be the case that there are pencils although makers are massively mistaken regarding their nature. The mistaken individuals are not makers, nor are the objects pencils. These people mistakenly believe they are the makers of certain artifacts, when in fact the aliens are the true creators of these artifactual kinds. Since these people have made nothing at all, the humans intending to make pencils have not made pencils, nor any other kind of artifact. Since the aliens have an entirely different concept in mind (that of a monitoring device), per DP3, no pencils have been created at all (but only monitoring devices). Thus this is not a case in which makers are massively mistaken about the nature of an artifactual kind they create, but in which certain people mistakenly believe they are makers and thus have mistaken beliefs about which artifactual kind certain objects belong to.

These examples do point out an interesting feature of our epistemic relation to artifacts, however. For although makers have a certain amount of incorrigible knowledge of the nature of the artifacts they create, no one has incorrigible knowledge that he or she is a maker of any particular artifact—cases can always be conceived in which someone is simply fooled into thinking that he or she has created an object. Thus although makers do have some incorrigible first-order knowledge of the nature of strict artifactual kinds, they lack incorrigible second-order knowledge that they have such knowledge. The common examples [Nelson 1982] used to show that we may all be wrong about the functions of common artifacts exploit the second-order fallibility, but do not undermine the first-order incorrigibility of the concept of whomever the makers really are.

### *3.3 Reference to Artifactual Kinds*

These results also have important consequences regarding the reference of artifactual kind terms. Although it is widely accepted that causal theories of reference apply uniformly to artifactual and natural kind terms [Putnam

---

<sup>26</sup> Hilary Putnam [1975a, 244] makes a similar argument that it could turn out that pediatricians are not doctors at all (as we had thought) but Martian spies. Since pediatricians seem to form an institutional rather than artifactual kind, I treat only the pencils case above. Nonetheless, parallel considerations apply to both cases.



1975a, Kornblith 1980, Nelson 1982, Elder 1989, Bloom 1996], this view cannot be maintained for all cases of reference to artifactual kinds. As we have seen, here, as in the institutional case, the kinds fail to have natural boundaries. As a result, the causal theory of reference (whether in pure or moderated form) cannot apply universally—in particular, it cannot apply within the context of creation. It is supposed to be a virtue of causal theories of reference that they enable us to refer to a kind without the need for anyone to have any substantive concept of the nature of the kind, since the term's extension is determined by the natural boundaries of the kind, not by any of our beliefs or concepts regarding those boundaries (except, in moderated versions, for high-level concepts that specify (e.g.) species-level boundaries as the relevant ones). Reference to artifactual kinds, however, cannot proceed without someone (namely, those responsible for the production and reproduction of these artifacts) having a substantive concept of the nature of the kind. In the case of strict kinds, that concept plays a central role in determining the nature of the kind and thus the extension of the term, since anything that is of that strict kind must have most of the features believed by makers to be relevant to membership in that kind. Even in the case of loose kinds, makers' concepts are essential to there being a kind to refer to, and, taken jointly, jointly determine the substantive features relevant to the nature of the kind, and thus play a central role in determining the overall extension of the term. Contrary to standard causal theories of reference, there can be no reference to the kind without someone having a relevant concept that in turn plays a role in determining the term's extension.

As in the case of institutional kinds, these results only demonstrate that not *all* cases of reference to these kinds may proceed along a causal model. Given that there is a kind established and maintained by makers and their intentions, those outside the system may refer causally to toasters as “whatever is of the same artifactual kind as these”, without themselves having any (or any correct) concept of what the relevant conditions are for being a toaster (though those involved in the production and reproduction of toasters must have such a concept). Yet all such causal cases depend on the existence of the primary case involving makers' declarative mental or verbal reference to aid in establishing and maintaining the kind.

#### **4. Consequences for Social Science**

The above results might raise some concern that the failure of these realist views to apply to such objects means that we must treat artifacts and institutional entities as merely imaginary objects or phantasms, and (because of their close epistemic relation to us) rule them out as potential objects for social scientific study and discovery. This would severely constrain the subject matter of social science, given the focus of sociology, political science

and economics on social, legal, political, and economic institutions; as well as the interest of ethnography, archeology and history in artifacts.<sup>27</sup>

But, just as we do not need to maintain that all entities are belief-independent to maintain a general thesis of realism, so we need not conceive of institutional and artifactual kinds as existing and having their natures entirely independently of all beliefs in order to treat them ontologically seriously, as more than mere phantasms or mental constructions (in the way that all kinds are, according to the idealist or constructivist). For although artifactual and institutional kinds depend on human beliefs and intentions regarding their nature, their existence and the way they are also partially depends on real acts e.g., of manipulating things in the environment, and many of the properties of concrete artifacts and institutional objects (weight, flammability, chemical properties...) may be determined directly by those of physical objects without regard for our beliefs about them. Thus, although they do not meet the realist paradigm of entities entirely independent of us for their existence and essence, they certainly are also not mere mental constructs.<sup>28</sup>

Nonetheless, worry might persist that the close epistemic relationship I have described would undermine the possibilities of discovery in social science by making facts regarding the nature of artifactual and institutional kinds transparent to those who create and maintain them. More thoroughly considered, however, the above results do not unduly constrain possibilities of social scientific discovery. First, although the above results do apply to artifactual and institutional kinds, it should not be inferred that the same follows for all ordinary, social, or (partially) mind-dependent objects. As mentioned in section one above, it is perfectly possible for there to be mind-dependent entities that nonetheless have natural boundaries. The reason both institutional and artifactual kinds lack natural boundaries has to do with the specific form of dependence on mental states exhibited by these kinds, namely that they depend on certain people accepting principles *about the nature of the kind itself*, as, e.g., members of the relevant community must accept constitutive rules stating sufficient conditions for the existence of a member of an institutional kind, and makers of artifacts must, in intending to produce an artifact of kind K, hold a substantive view of the nature of Ks. These are what I shall call “transparent” kinds, such that the existence of the kind K entails the existence of K-regarding intentional states.<sup>29</sup>

---

<sup>27</sup> For a detailed discussion of the ontology and epistemology of the kinds of human geography and the consequences for the science of geography, see my [2001].

<sup>28</sup> Devitt discusses the dual dependence of artifacts on beliefs and practices as well as on the independent world, and the differences between artifacts and the constructivist's entities, in [1991, 246-249].

<sup>29</sup> This is closely allied to what Searle [1995, 32-34] calls the “self-referentiality” of social concepts, such that if x is of kind K, then x, or things like x, must be regarded as K.

The possibility is often overlooked, however, that there may be entities (and kinds of entities) depending on mental states of various kinds, without their depending on any beliefs about them (or about that kind itself).<sup>30</sup> Some social kinds such as racism, superstition, etc., do depend on the existence of certain sets of beliefs and intentional behaviors, but may exist without the existence of any beliefs that are themselves *about* racism, superstition, etc. Large scale social facts or those involving statistical generalizations over social facts (e.g., the fact that we are in a recession, or that 70% of Americans support the death penalty) similarly depend on certain mental states (regarding money and financial transactions, regarding the death penalty), but not on any intentional states regarding recessions or statistical generalizations. In such cases, although the Independence Principle clearly fails, these opaque kinds of entities may remain unknown even to those within the relevant society, and thus may require substantive social scientific discovery of their existence as well as nature.

Moreover, even regarding institutional and artifactual kinds, there is much that awaits discovery by the social sciences. First, all facts about these kinds may be opaque to those not involved in the production or maintenance of those entities within the context of the relevant society, leaving historians, ethnographers, archeologists and the like with as much work as ever. But even within a community of insiders involved in sustaining institutional facts or reproducing artifacts, there are many facts to be uncovered by social science. Granted, such individuals may have no need of social science to determine all facts about what it takes to be of the artifactual or institutional kinds they create. But social science does not typically concern itself with such issues as whether or not it is in the nature of a pencil to be a writing instrument or what the necessary conditions are for something to be a dollar bill, but rather with such issues as the impact of the printing press on the growth of religion in Europe, the consequences of mechanical production on urban growth and standards of living, or the importance of the invention of the rifle in achieving military dominance. These are issues regarding causal relations involving artifactual or institutional kinds, as are some of the most famous issues in social science, concerning e.g., the Marxist and feminist claims about (perhaps) unintended and unnoticed oppressive consequences of our practices involving money, division of labor, etc. Such causal facts certainly remain opaque and in need of discovery. There is even room for critique of elements of a society's metaphysical understanding of its own institutions, e.g., in exposing the beliefs of a society that believes that its institutions (king, laws, customs) are established through natural or supernatural powers

---

<sup>30</sup> It is overlooked, e.g., by Searle, who asserts that all concepts naming social facts exhibit what he calls "self-referentiality" [1995, 32].

rather than simply through collective acceptance. The range of social scientific discovery remains as wide as we ever expected.

## 5. Conclusion

Recent analytic ontology, epistemology, and semantics have largely been driven by the desire to provide theories adequate to the natural sciences, and have focused on natural and scientific kinds as the primary objects of concern. This has, no doubt, brought great advantages in certain areas. But the narrow focus also has its price. I have argued that standard realist views of ontology, epistemology, and reference that have been developed with natural and scientific kinds in view cannot apply fully to such ordinary objects in the human world as artifacts and institutional objects. In both cases, such entities fail both the independence and the natural boundaries criteria for ontological realism. As a consequence, both our epistemological and semantic theories must differ substantially from those appropriate for the presumed independent kinds of nature, for certain groups must have substantive concepts of the nature of the kind, where those concepts play a crucial role in determining the extension of the relevant kind terms, and are protected from certain forms of ignorance and error.

This situation is certainly no embarrassment for a general realist thesis, since it is sufficient for that that there be some kinds that exist and have their nature independently of all representations. It is, however, somewhat more of an embarrassment for those who would complacently assume that the widely accepted pictures of ontology, epistemology, and reference developed with the natural world in mind are universally applicable. If philosophy is to help us make sense of the human world—and indeed of those kinds studied by the social and human sciences as well as the natural sciences—we will need to seek a broader picture, for we may require very different theories to handle independent parts and aspects of the world, and those that are in part our own construction.<sup>31</sup>

---

<sup>31</sup> Many thanks to Risto Hilpinen, Bernard Kobes, Hilary Kornblith, William Lycan, Beth Preston, Stephen Schwartz, Deborah Tollefsen, and two anonymous referees from this journal, who read earlier versions of this paper or its predecessors and commented on it in ways that greatly improved the final version. This paper, its parts, or predecessors were also presented at the 2000 Pacific APA, the 2001 Central APA, the University of Hong Kong, the University of Notre Dame, and Texas Tech University, and discussed at the University of Miami. My thanks go out to all those present for very helpful discussions.

## References

- Bloom, Paul (1996). 'Intention, History and Artifact Concepts'. *Cognition* 60, 1-29.
- Cummins, Robert (1975). 'Functional Analysis'. *The Journal of Philosophy* Vol. LXXII No.20, 741-64.
- Dennett, Daniel (1990). 'The Interpretation of Texts, People, and Other Artifacts'. *Philosophy and Phenomenological Research* Vol. 1, Supplement, 177-194.
- Devitt, Michael (1991). *Realism and Truth* (Second Edition). Princeton, New Jersey: Princeton University Press.
- Devitt, Michael and Kim Sterelny (1999). *Language and Reality* (Second Edition). Cambridge, Massachusetts: MIT Press.
- Donnellan, Keith (1966). 'Reference and Descriptions'. *Philosophical Review* 75, 281-304.
- Elder, Crawford (1989). 'Realism, Naturalism and Culturally Generated Kinds'. *Philosophical Quarterly* Vol. 39, 425-444.
- Hart, H. L. A. (1952). 'The Ascription of Responsibility and Rights'. In Anthony Flew, ed. *Logic and Language*. Oxford: Basil Blackwell.
- Heller, Mark (1988). 'Vagueness and the Standard Ontology'. *Noûs* 22, 109-131.
- Hilpinen, Risto (1992). 'On Artifacts and Works of Art'. *Theoria* Vol. 58, 58-82.
- Hilpinen, Risto (1993). 'Authors and Artifacts'. *Proceedings of the Aristotelian Society* 93, 155-178.
- Hilpinen, Risto (1995). 'Belief Systems as Artifacts'. *The Monist* Vol. 78, No. 2, 136-155.
- Kobes, Bernard (2000). Response to Amie L. Thomasson's 'Artifactual Functions and Reference to Artifactual Kinds'. Presented at Pacific Division meeting of the American Philosophical Association.
- Kornblith, Hilary (1980). 'Referring to Artifacts'. *The Philosophical Review* LXXXIX No. 1, 109-114.
- Kripke, Saul (1980). *Naming and Necessity*. Oxford: Blackwell.
- Levin, Michael (1976). 'On the Ascription of Functions to Objects, with Special Reference to Inference in Archaeology'. *Philosophy of Social Science* 6, 227-234.
- Levinson, Jerrold (1979). 'Defining Art Historically'. *British Journal of Aesthetics* 19, 232-250.
- Levinson, Jerrold (1989). 'Refining Art Historically'. *Journal of Aesthetics and Art Criticism* 47, 21-33.
- Levinson, Jerrold (1993). 'Extending Art Historically'. *Journal of Aesthetics and Art Criticism* 51, 411-423.

- Millikan, Ruth (1984). *Language, Thought and other Biological Categories: New Foundations for Realism*. Cambridge, Massachusetts: MIT Press.
- Millikan, Ruth (1993). *White Queen Psychology and Other Essays for Alice*. Cambridge, Massachusetts: MIT Press.
- Millikan, Ruth (1999). 'Wings, Spoons, Pills and Quills: A Pluralist Theory of Function'. *The Journal of Philosophy* Vol. 96 No. 4, 191-206.
- Nelson, James (1982). 'Schwartz on Reference'. *Southern Journal of Philosophy* Vol. 20, 359-365.
- Papineau, David (1979). *Theory and Meaning*. Oxford: Clarendon Press.
- Preston, Beth (1998). 'A Pluralist Theory of Function'. *Journal of Philosophy* Vol. XCV No. 5, 215-54.
- Putnam, Hilary (1975a). 'The Meaning of "Meaning"'. In *Mind, Language and Reality*. Cambridge: Cambridge University Press.
- Putnam, Hilary (1975b). 'Is Semantics Possible?'. In *Mind, Language and Reality*. Cambridge: Cambridge University Press.
- Schwartz, Stephen P. (1978). 'Putnam on Artifacts'. *The Philosophical Review* LXXXVII No. 4, 566-574.
- Schwartz, Stephen P. (1980). 'Natural Kinds and Nominal Kinds'. *Mind* 89, 182-195.
- Searle, John (1995). *The Construction of Social Reality*. New York: The Free Press.
- Searle, John (1996). 'A Taxonomy of Illocutionary Acts'. In A. P. Martinich, ed. *The Philosophy of Language*. New York: Oxford University Press, 141-155.
- Thomasson, Amie L. (1999). *Fiction and Metaphysics*. Cambridge: Cambridge University Press.
- Thomasson, Amie L. (2001). 'Geographic Objects and the Science of Geography'. *Topoi* 20/2, 149-159.
- Unger, Peter (1979). 'There are no Ordinary Things'. *Synthese* 41, 117-154.