CHRIS SWOYER

STRUCTURAL REPRESENTATION AND SURROGATIVE REASONING

ABSTRACT. It is argued that a number of important, and seemingly disparate, types of representation are species of a single relation, here called *structural representation*, that can be described in detail and studied in a way that is of considerable philosophical interest. A structural representation depends on the existence of a common structure between a representation and that which it represents, and it is important because it allows us to reason directly about the representation in order to draw conclusions about the phenomenon that it depicts. The present goal is to give a general and precise account of structural representation, then to use that account to illuminate several problems of current philosophical interest – including some that do not initially seem to involve representation at all. In particular, it is argued that ontological reductions (like that of the natural numbers to sets), compositional accounts of semantics, several important sorts of mental representation, and (perhaps) possible worlds semantics for intensional logics are all species of structural representation and are fruitfully studied in the framework developed here.

We use representations in nearly all our reasoning about the world. There are so many types of representation that a single account probably cannot do justice to them all. Still, I shall argue, a number of important and seemingly disparate types of representation are species of a single relation that can be described in detail and studied in a way that is of considerable philosophical interest. I shall call this relation *structural representation*. My aim here is to explain what structural representation is and to show why it is philosophically interesting.

Structural representation enables us to reason directly about a representation in order to draw conclusions about the things that it represents. By examining the behavior of a scale model of an aircraft in a wind tunnel, we can draw conclusions about a newly designed wing's response to wind shear, rather than trying it out on a Boeing 747 over Denver. By using numbers to represent the lengths of physical objects, we can represent facts about the objects numerically, perform calculations of various sorts, then translate the results back into a conclusion about the original objects. In such cases we use one sort of thing as a surrogate in our thinking about another, and so I shall call this *surrogative reasoning*.

I shall begin with an informal account of structural representation

Synthese **87**: 449–508, 1991.

and surrogative reasoning, explain why they are important, and then move on to a more detailed and precise account of the former. The account uses *intensional relational structures*, so in Sections 2 and 3 I explain what these are, and show how they are useful in the study of representation. In Section 4 I work through an example to show more clearly what structural representation involves, and in Section 5 I examine the nature of representational artifacts and their relationship to various theses about conventionalism and underdetermination. In Section 6 I present a general account of structural representation. The applications of the account are discussed in the last, and most important, section. There I argue that several philosophically interesting phenomena – ontological reductions (like that of numbers to sets), compositional accounts of semantics, mental models, and (perhaps) possible worlds semantics for intensional logics – are examples of structural representation and are fruitfully studied in the framework developed here.

1. STRUCTURAL REPRESENTATION

We represent things using scale models, road maps, computer simulations, musical notation, Gödel numbers, English sentences, smoke signals, and Braille. The diversity of examples suggests that anything can, with sufficient ingenuity and determination, be employed to represent almost anything else, and the uses we make of representations are nearly as varied. Nevertheless, I think that a central point of *much* representation – one reason why it plays so vital a role in our lives – is that it allows us to reason directly about a representation in order to draw conclusions about some phenomenon that it represents.

This can be important for a variety of reasons: the original phenomenon may be difficult to observe, understand, or manipulate – it might not even exist.[1] Such reasoning may be quite self-conscious, as when a geometer studying the projective properties of geometrical objects begins with a figure (like an ellipse), transforms it to some other figure (like a circle) that shares those features of the original figure with which he's concerned, reasons about the second figure, and then transfers the result back to a conclusion about the first. Other times the medium of representation becomes so familiar that we scarcely realize it's there, as with the detailed shapes of the numerals we write as we struggle to balance our checkbooks. In the measurement of length or voltage, we transform information about the physical magnitudes into numerical information, which facilitates reasoning because of the rich set of mathematical concepts, techniques, and theories our culture provides. In an analog computer we move in the opposite direction, representing numbers by physical magnitudes like voltage or length. But the point in each case is the same: to represent something in a medium that facilitates inference about it.

Not all representations allow detailed reasoning about the things they represent; no amount of pondering the embroidery of Hester's 'A' will reveal the details of her exploits. Still, the point of much representation is to *mediate inferences* about things in the world, and this raises what might be called the *applications problem*. The question how an abstract body of theory like mathematics can apply to concrete reality is a venerable one in philosophy, but if the following account is right, it is just a special case of a more general puzzle: How can *any* representational system – from rudimentary arithmetic to a complex natural language – be successfully applied to the world? How is such representation possible? I believe that the *best explanation* why a mathematical theory applies to the concrete phenomena it does is that it has many of the same *structural features* as those phenomena. It is a central thesis of this paper that *shared structure* of precisely this sort explains the applicability of a wide range of representational systems – including many *non*-mathematical ones – to the things they represent.[2]

In many cases the notion of shared structure is familiar enough. A scale model of a DNA molecule has much of the structure of the molecule itself, because various *relationships* among the model's parts correspond to important relationships among the molecule's parts. But shared structure is also important in cases where it is less obvious just what is involved; in order to see how, it will be useful to consider the measurement of length.

When we measure the lengths of physical objects in meters, we pair the objects with numbers in such a way that the two exhibit a common *pattern*. We can view our measurement scale as correlating each physical object with a unique numerical surrogate or representative in the set of positive real numbers. For example, a meter bar is paired with 1 and a twelve-inch ruler with 0.3048. Furthermore, properties and relations of the physical objects are paired with numerical properties and relations. For example, the relation that two objects stand in whenever one is longer than the other is represented by the greater-than relation on the real numbers, so that one object is longer than a second just in case the number representing it is larger than the number representing the second.

Such similarity of structure explains why the mathematics of the positive real numbers applies to physical objects and their lengths as follows. We begin with empirical facts involving physical objects and use our measurement scale as a bridge to their numerical surrogates or proxies. We then mobilize logic and the mathematical theory of the positive real numbers to infer that further numerical facts obtain. Finally, once our calculations are complete, we make the return trip to a conclusion about the original physical objects and their lengths.[3]

This example suggests a general model for structural representation: the *pattern* of relations among the constituents of the represented phenomenon is mirrored by the pattern of relations among the constituents of the representation itself. And because the arrangements of things in the representation are like shadows cast by the things they portray, we can encode information about the original situation as information about the representation. Much of this information is preserved in inferences about the constituents of the representation, so it can be transformed back into information about the original situation. And this justifies surrogative reasoning, since if we begin with true premises about the object of representation, our detour through the representation itself will eventually wind its way back to a true conclusion about the original object.[4]

My aim here is to develop and refine this intuitive picture, but before turning to this I want to emphasize two things. First, the interest of the notion of structural representation is not that it fully captures our *ordinary* sense of representation – it doesn't, and it's not intended to. Structural representation is not a necessary condition for representation in the ordinary sense of the word, since with sufficient perseverance – or perversity – we can use anything to represent virtually anything else, and in many cases the two things won't have any interesting structural similarities at all. And it is not sufficient for ordinary representation, since if you can find one structural representation of something, you can usually find many. Still, if one thing is a structural representation of a second, it has the potential to be used in surrogative reasoning about it, and so an account of structural representation will help us understand how one very important kind of representation is possible. Furthermore, as we shall see in Section 7, a number of philosophically interesting phenomena that are not usually regarded as representations at all turn out to be specimens of structural representation, and so our account will allow us to examine them in a novel framework that suggests fresh approaches to their study.

Second, if we avail ourselves of relations that are easily come by, e.g., relations-in-extension or Goodmanesque relations, then we can view almost anything as a structural representation of almost anything else. The situation is analogous to that in group theory. If we are willing to allow just *any* binary relation with the right formal properties to serve as group addition, then virtually any collection of things can be regarded as a group. But the reason why group theory is so useful is that there are many cases where some relation of independent interest (like the rotation of a geometrical figure) turns out to have the structural features of group addition; indeed, this often explains why a collection behaves in ways of interest to us. Similarly, I shall argue, structural representation is important because there are various pairs of things – numbers and sets, syntax and semantics, Kripke model structures and the actual world – that behave in interesting ways precisely because relations of antecedent and independent significance in one member of the pair have the same structure as relations of antecedent and independent significance in the other.

Structural representation is a term of art, and so we cannot begin with a definitive picture of which examples of representation are structural and which are not. Instead, guided by an intuitive picture of shared structure and a handful of prototypes, like the measurement of length, I shall develop an account of structural representation, and we can then ask which phenomena fit the characterization and which do not. I shall reserve the notion of *surrogative reasoning* for reasoning about a structural representation in order to draw inferences about what it represents, and so the use of analogies and metaphors in reasoning needn't always be surrogative.

Our discussion thus far suggests the following desiderata for an account of structural representation:

(1) It should solve the applications problem, explaining the applicability of structural representations to the things they represent;

(2) Since virtually any medium can be used to provide a structural representation of anything else, it should be medium-

indifferent, insensitive to the idiosyncrasies of particular media (the medium is *not* the message);

(3) It should clarify how we can represent merely possible situations;

(4) It should be precise enough to allow a careful and detailed study of the formal aspects of structural representation;

(5) It should be general enough to subsume a number of different types of representation.

In the next section I shall begin developing the informal account sketched above into a more general and precise characterization of structural representation that satisfies these conditions.[5]

2. ABSTRACTION AND INTENSIONAL RELATIONAL STRUCTURES

Any actual system or situation is likely to have a great many features, and systematic and detailed reasoning about it will require us to focus on some at the expense of others. Indeed, intellectual progress often goes hand in hand with such abstraction, the discernment of a few properties like force, energy, and information as behind-the-scenes causes of the phenomena around us. The logician's concept of a *relational system* furnishes a useful device for dealing with many aspects of abstraction – including some of those important in representation – in a precise way. As traditionally conceived, a relational system is a set-theoretic object comprising a domain of individuals and one or more relations-in-*extension* on that domain. Such relational systems have many important uses, but their treatment of properties and relations as extensional entities (sets and relations-in-extension, respectively) poses several difficulties for their use in a study of representation.

One problem is that several important examples of representation, including the propositional attitudes and the semantics of natural languages, are suffused with intensionality, and so are not easily accommodated by extensional machinery. Furthermore, representation itself is tinged by intensionality, since a representation can represent something as having one property without representing it as having every other property that happens to be co-extensive with it. Suppose that exactly the same objects are blue and have a density of 7 gm/cm^3. Readings on a machine built to detect colors are plausibly thought to represent the color of such an object, rather than its density, and it seems clearer still that 'blue' represents only the color, and '7 gm/cm^3' only the density. The use of traditional extensional relational systems would make it difficult to accommodate this fact; genuine properties and relations will make it much easier.

The introduction of intensional entities like properties and relations has its costs and would be ill-advised unless we had independent reasons to suppose that such things existed. Elsewhere I have argued that we have a variety of reasons for thinking that they do, the most important being that they provide the *best explanation* of a number of phenomena, among them natural laws, causation, measurement, and the modalities (Swoyer [1982b], [1983], [1987], esp. pp. 240–43). I shall not defend such entities further here, except insofar as the present account adds one more item to the list of things that they help us explain, namely representation.[6]

In order to work genuine properties and relations into the picture, I shall employ a slight modification of traditional, extensional, relational systems which I shall call *intensional relational systems* (*IRS*s, for short). These differ from their extensional kin in containing properties and relations that are not constructed out of sets of objects, possible worlds, or anything else at all. *IRS*s can be at least as rich as their extensional brethren, sporting a rich type structure or a supple non-typed one. But such complexities aren't needed to illustrate the essential points about structural representation, and so I shall employ simpler intensional relational systems here.

To this end, let us think of an *IRS* as an ordered quadruple:

$$A = \langle I^A, {}^f\mathfrak{R}^A, {}^s\mathfrak{R}^A, {}^\vee \rangle,$$

where I^A, ${}^f\mathfrak{R}^A$, and ${}^s\mathfrak{R}^A$ are non-overlapping sets. Intuitively, I^A is a domain of individuals, ${}^f\mathfrak{R}^A$ a domain of first-order relations (including one-place relations, or properties), and ${}^s\mathfrak{R}^A$ a domain of second-order relations. I shall call the set of *all* the relations in the system the *full domain of relations*, $\mathfrak{R}$ for short, and the union of the domain of individuals and the full domain of relations the *total domain* of the system. For convenience I shall require that at least two "adjacent" domains be nonempty, i.e., that at most one of I and ${}^s\mathfrak{R}$ be empty. And I shall drop superscripts and related paraphernalia when ambiguity won't result.

In addition to its order or level, each relation in $\mathfrak{R}$ has a fixed rank or number of argument places. I shall let both the order and the rank of individuals be 0. We can then say that the order o and rank r of an item in any domain in an *IRS* determine its *type* $\langle o, r\rangle$; for example, *loves* is a first-order, two-place (type $\langle 1,2\rangle$) relation. Finally, the extension assignment, $^{\vee}$, is a unary function on $\mathfrak{R}$ that assigns extensions to all of the relations in this set (it is needed because genuine properties and relations, unlike relations-in-extension, don't come with built-in extensions). The extension of an n-place relation is a set of n-tuples of entities from level l-1 (so that, for example, the extension of the relation *loves* is a set of ordered pairs of individuals). In short, the extension of an n-place relation is just the familiar set-theoretic object that extensionally-minded philosophers take to *be* that relation. The identity conditions of genuine properties and relations are not determined by the things that happen to exemplify them, and so distinct properties and relations may have precisely the same extension, a fact we accommodate by allowing the extension assignment function to be many-one. And we say that two *IRS*s have the same *similarity type* just in case there is a one-one, onto function from the full domain of relations of the first to that of the second which maps each relation of the first to a relation of the same type in the second.[7]

We can now make the notion of *shared structure* precise. To simplify exposition I shall confine my attention to first-order relations, but precisely analogous points hold for higher-order relations as well. The intensional relational systems $\boldsymbol{A}$ and $\boldsymbol{B}$ have the same structure – they are *isomorphic* – just in case they are of the same similarity type and there is a one-one, onto function, c, from the total domain of $\boldsymbol{A}$ to the total domain of $\boldsymbol{B}$ that preserves both the type and the structure of all the relations in $\boldsymbol{A}$. This means that $\boldsymbol{A}$ and $\boldsymbol{B}$ are isomorphic just in case there is a one-one, onto, type-preserving function c such that

$$(\boldsymbol{PR}) \qquad \langle i_1, \ldots, i_n\rangle \in {}^{\vee}R \text{ if and only if } \langle c(i_1), \ldots, c(i_n)\rangle \in {}^{\vee}c(R),$$

for each n-tuple of individuals $\langle i_1, \ldots, i_n\rangle$ in I^A and every first-order n-place relation R in ${}^{f}\mathfrak{R}^{A}$.[8] The intuitive force of (***PR***) is that a group of individuals (taken in a given order) from the first relational system stand in the relation R exactly when their surrogates in the second system (taken in the same order) stand in the surrogate of R. When this occurs, I shall say that the function c *respects* the relation R.

If a mapping has all of the features of an isomorphism except being onto, it is an *isomorphic embedding*; such mappings yield an isomorphic copy of one *IRS* in another. Various other relaxations in the requirements imposed on the correlating function *c* will be discussed below, but in order to have a tentative working account, I shall *provisionally* identify the structural representation of one *IRS* by a second as the isomorphic embedding of the first in the second. This gives us an account that will be easy to generalize in various ways as we proceed. But it must be stressed that it is often anything but easy to show that such an account applies in a particular case, since the embeddability of one *IRS* in another is by no means an automatic consequence of their having the same similarity type. In addition to this, the corresponding relations of the two systems must have the same structural or formal features. To ensure that they do, we must introduce axioms to constrain the behavior of the relations in each system, then demonstrate that the axioms really do impose the same structure on the two *IRS*s by using them to prove the existence of an isomorphic embedding of one in the other. In many cases this can be quite difficult, and in some it is an open question whether it is even possible.

The use of *IRS*s in dealing with actual cases of representation also involves an element of idealization, for most of the things that we represent, and most of our representations of them, are not literally intensional relational systems. Moreover, there is a risk of confusion here, since an *IRS* may itself be regarded as a sort of model or representation of the real-life situation we use it to study. To keep the distinctions straight, I shall say that an intensional relational system is an *IRS-model* of a real-life situation or thing when it satisfies two conditions. First, the relational system contains at least some of the same individuals and relations as that situation, and, second, an *n*-tuple of objects is in the extension of a relation in the relational system just in case those objects, taken in that order, stand in that relation in the real-life situation. Strictly speaking, the relationship of structural representation holds only between an *IRS*-model of a real-life situation and an *IRS*-model of a representation of that situation. But although a real-life situation and its *IRS*-models are distinct, an *IRS*-model can portray the situation much more directly than any extensional relational system can. This is so because the former will include at least some of the very same properties and relations present in the situation itself, rather than extensional stand-ins for them. Indeed, it comes about as close to the original situation as a mathematically tractable creature can. Hence, it is often quite natural to treat real-life systems as though they were *IRS*s, and to speak of one system as a structural representation of another.[9]

It is also true that real-life surrogative reasoning about a phenomenon is unlikely to make explicit use of the axioms for an *IRS* used to model it. Such axioms are better viewed as representing the competence underlying surrogative reasoning, helping to explain *what* it is that people do in a given bit of surrogative reasoning (e.g., what function they compute), but leaving open *how* they do it (e.g., what algorithm they use to compute a given function).

I have treated properties and relations as primitive entities having no internal complexity, but it is possible to add operations to an *IRS* that build more complex relations out of simpler ones, and in dealing with some kinds of representation it is useful to do so. For example, it might be thought that if a domain of first-order relations contains the two properties *F* and *G*, then it should also contain the *conjunctive* property, *being (both) F and G*. And if it contains the two-place relation *L*, perhaps it should also contain the property, *bearing L to something*. To accommodate such views, we can expand an *IRS* to include a set, *Op*, of operations corresponding to connectives (e.g., conjunction, negation, infinitary disjunction), quantifiers (e.g., existential quantification), or operations from the algebra of relations (e.g., conversion, reflection). Each new relation-building operation we add would require additional axioms specifying its domain and range, as well as its mode of interaction with the extension assignment. For example, our conjunctive operation would naturally be thought of as mapping pairs of relations of a given type to a relation of the same type, and something would be in the extension of the property *being (both) F and G* just in case it was in the extension of *F and* in the extension of *G*. All this can be done quite precisely, and if we add operations that close off open argument places – as those corresponding to quantifiers do – it is even possible to obtain 0-place relations, that is, propositions, and to take their extensions to be truth values. Such operations generate additional ontological commitments, however, and so I shall include them in *IRS*s only when there is a particular need to do so.[10]

3. REPRESENTATION REVISITED

In everyday life our ability to represent sequences of events and alternative possibilities is crucial to planning and deliberation. In science we are frequently interested in the possible states and histories of various systems, which are often represented by points in, and trajectories through, a state space. Hence, it is not surprising that a key feature of many representational systems is their ability to depict a range of times and possibilities. The *IRS*s described thus far are *synchronic*, providing static snapshots of things frozen in time, but they can be enriched to form what I shall call *multi-track IRS*s by adding an ordered set of *times*, a set of *tracks* (to represent various options or alternatives), and an extension assignment that assigns an extension to each relation at every track at every time. Of course in some cases, e.g., sample spaces in probability, we need only one time in our system, and in others, e.g., those involving just the *actual* history of something, only one track.

These modifications require additional axioms to ensure that the set of times and tracks have the right sort of structure. For example, the ordering relation on times should be endowed with at least some features like irreflexivity and transitivity that we normally attribute to the earlier-than relation. It needn't have all such features, though, since many representations depict only a few moments, for example, ones before and after some experimental manipulation or the adoption of a new diet. The axioms governing the relations in a multi-track system must also do more than their counterparts in a synchronic *IRS*, since they govern the structural configuration of properties and relations through time and in alternative situations. In what follows I shall focus on synchronic *IRS*s, but most of the discussion is straightforwardly generalized to multi-track *IRS*s by adding the appropriate times and tracks, making the corresponding changes in the extension assignment, and redefining notions like isomorphism in the obvious way (as sameness of structure at all times and tracks).[11]

The provisional account of structural representation as the isomorphic embedding of one intensional relational system in another satisfies all of the desiderata listed in Section 1, except for generality. It explains the applicability of representational systems to what they represent in terms of structure-preserving mappings from one *IRS* to another. It is indifferent to the nature of the representing medium. It helps us see how representational systems can model possible variations on actual situations, since we can employ extension assignments that assign extensions to relations different from those which they actually have, but which mirror the facts as they might have been. And it is precise. The account is still not general enough to cover as many cases as it should, however, and in Section 6, I shall introduce a more general sort of mapping to remedy this defect. But before turning to this, it will be useful to illustrate the points discussed thus far with a brief sketch of a paradigm case of structural representation, the measurement of length.

4. EXTENSIVE MEASUREMENT AS A PARADIGM OF STRUCTURAL REPRESENTATION

In order to illustrate several additional points about structural representation, I shall treat length measurement as the assignment of numbers to first-order *properties – lengths –* rather than to *individual objects* (as it was treated in earlier sections). This will help us to see that the first-order structure of an ensemble of individuals is often less important than the *higher-order* structure of the properties and relations which those individuals exemplify. It will also provide an example of the important phenomenon of *trans-type* representation, in which constituents of a representation are of a different type from the things that they represent.

We shall begin by introducing the notion of an *Extensive Property System* (*EPS*). An *EPS* has nothing in its domain of individuals and only two items in its domain of second-order relations, so for readability I shall omit the former and list the members of the latter. This done, an extensive property system has the form: $\boldsymbol{E} = \langle E, >, \bigcirc, ^{\vee} \rangle$. In intended interpretations, E is a set of one-place properties, of the sort W. S. Johnson called *determinates*, specific lengths like the property which (once a numerical scale has been established) is naturally identified as *being five meters long*. Next, $>$ is a second-order, two-place relation that holds between a pair of lengths just in case any object exemplifying the first is longer than any object exemplifying the second, and $\bigcirc$ is a second-order, three-place relation that holds between a triple of lengths just in case the combined length of any pair of objects exemplifying the first two is equal to that of any object exemplifying the third. We can then think of a measurement scale as a function, s, that pairs determinate lengths in E with their *numerical surrogates* in the positive, additive real numbers. And we model the latter by the *IRS* $\boldsymbol{R}^+ = \langle R^+, >, +, ^{\vee} \rangle$, where R^+ is the set of positive real numbers and $>$ and $+$ are the greater-than relation and addition.[12]

This sort of approach to measurement has been developed in detail in the representational theory of measurement (cf. fn. 3), and the present account is heavily indebted to work in that tradition. However, it diverges from much of this work in its rejection of nominalism, extensionalism, and an austerely empiricist epistemology, and these differences will be important in allowing us to generalize this approach to measurement to an account of structural representation. The third difference deserves emphasis, since many traditional accounts of measurement employ primitive terms that are amenable to a reasonably simple empirical, even operationalistic, interpretation. For example, the relation $>$ might literally be defined in terms of the behavior of a given type of ruler. In real life, however, the measurement of such magnitudes may be quite indirect and subtle, as it is for the diameter of a hydrogen atom or the distance across the galaxy, and such epistemological considerations play no special role in the present account.

In order to show that the positive, additive real numbers really do provide a structural representation of lengths, we must show that any *EPS* can be isomorphically embedded in $\boldsymbol{R}^+$. The formal features of $\boldsymbol{R}^+$ are well enough understood for present purposes, so the task is to devise axioms governing $>$ and $\bigcirc$ that do justice to our knowledge about length. Furthermore, the axioms should be qualitative (non-numerical), since on the present account there is nothing intrinsically numerical about the properties and relations in $\boldsymbol{E}$, and the goal is to prove – rather than presuppose – that the mathematics of the positive, additive reals is applicable to them.

The system $\boldsymbol{E}$ can be isomorphically embedded in $\boldsymbol{R}^+$ just in case there is a one-one function s from the domain of determinate lengths to the domain of real numbers that respects structure. This means that for any two determinate lengths, P_1 and P_2, in $\boldsymbol{E}$: (*A*) $P_1 > P_2$ just in case $s(P_1) > s(P_2)$, and (***B***) $s(P_1 \bigcirc P_2) = s(P_1) \bigcirc s(P_2)$.[13] The proof that such a function exists yields what is called a *representation theorem*. This theorem shows that any *EPS* can be isomorphically embedded in $\boldsymbol{R}^+$, i.e., that it can be represented in the positive, additive reals. This ensures that the set of isomorphic embeddings of any *EPS* in $\boldsymbol{R}^+$ is not empty, but it leaves open whether the mappings in this set have anything interesting in common. To show that they are related in some specifiable way is to prove a *uniqueness theorem*. In the case of extensive measurement, the uniqueness theorem tells us that such embeddings are unique up to a similarity transformation, that is, up to multiplication by a positive real number (a conversion factor that simply changes the units). This expresses the fact that nature does not determine a correct unit for length, but that once we select a unit, all of the remaining scale values fall into place.[14]

Axioms allowing the proof of the representation and uniqueness theorems may be found in a number of places (e.g., Krantz, et al. [1971], p. 73), but it is worth mentioning them briefly here in order to illustrate the *sorts* of axioms that are often relevant to structural representation. The following axioms can be formalized in an artificial language (as in Swoyer [1987]), but it is more perspicuous to give them in mathematical English. Moreover, many of the notions that figure prominently in structural representations – e.g., being infinite, being continuous, having the Archimedean property – elude the expressive capacities of elementary logic, so that standard benefits of formalization, like the availability of a complete logic, aren't in the cards anyway.

A system $\boldsymbol{E} = \langle E, >, \bigcirc, ^{\vee} \rangle$ is an *EPS* just in case it satisfies the following six axioms: first, $>$ is a linear order; second, $\bigcirc$ is a function (so that any two properties have a unique sum); third, $\bigcirc$ is associative; fourth, the sum of any two properties is greater than either property alone (this is a positivity axiom); fifth, $>$ and $\bigcirc$ interlock in such a way that summation preserves order (this is a monotonicity axiom); and sixth, the system has the Archimedean property.[15] The requirement that $\bigcirc$ be a function is added primarily for convenience, but the remaining five are axioms are necessary for a representation in the positive, additive reals. The simplest way to show that an axiom is necessary for such a representation is to assume that the representation exists, then to demonstrate that the axiom follows from this assumption.[16] Showing that a set of axioms is sufficient for the proof of a representation and uniqueness theorem is more arduous, since it involves demonstrating that a mapping of the required sort exists, but Hölder's work at the turn of the century establishes that axioms like those mentioned above do the trick. This means that *if* a particular family of properties (like lengths or masses) satisfies these axioms – and it is always an empirical question whether a given family does – we are justified in measuring them on a ratio scale and applying the mathematical theory of the positive, additive real numbers to them. This is so because the axioms guarantee that the family of properties shares much (though not all) of the structure of the positive, additive reals.

A great part of the *philosophical* significance of the representation theorem is that it explains the applicability of mathematics to reality;

more specifically, it explains – and justifies – the applicability of the mathematical theory of the positive, additive real numbers to lengths. And part of the philosophical significance of the uniqueness theorem is that it explains why scales for measuring length that are obtained from each other by multiplication of a positive constant are equally good (and hence why laws involving them are invariant under such transformations). Furthermore, as we shall see in the next section, it also helps us to separate the conventional aspects of a representation from the nonconventional.

Many species of representation are more complicated than extensive measurement and have been studied much less, and so it would be unrealistic to expect such precise results for them to be forthcoming immediately. However, our example of extensive measurement underscores the importance of representation and uniqueness theorems, and thus provides a useful ideal against which to evaluate treatments of more complicated varieties of structural representation.

5. ARTIFACTS AND CONVENTIONS

A map that reproduced every feature of Jamaica at a scale of a mile to a mile would be worse than useless. Distillation and abridgment are essential to representation, but representations typically add as well as subtract, having surplus features that do not correspond to anything in the phenomena they depict. Sometimes it is even tempting to mistake adventitious features of a representation for genuine features of the phenomena it portrays. It is only natural for children to suppose that Greenland is larger than Algeria, since the picture of it on their maps at school is so much bigger. They do not yet realize that the relative sizes of the pictures of land masses on conformal maps don't correspond to their actual sizes, that sizes, unlike shapes, are *artifacts* of such representations. We are all susceptible to such misapprehensions, and they can even be exploited to deceive, enticing us to focus on things like the widths of the bars in a graph, rather than on relevant aspects, like their heights. Despite the ubiquity of representational artifacts, there have been few attempts to provide a general account of them, but measurement theorists have made a systematic effort to deal with *one* species of artifact in their discussions of a technical concept of *meaningfulness*, and we can learn something about artifacts in general from their efforts.

Suppose that we have two metal rods, *a*, which is 100 meters long, and *b*, which is 50. As the uniqueness theorem for extensive measurement shows, we can set up a ratio scale for lengths by picking a unit (like a meter), whereupon all of the other scale values fall into place. Hence, the claim that *a* is twice as long as *b* would remain true even if we switched to some other system of units, and so it reflects an objective, *scale-independent* feature of the two rods. The situation is quite different for temperatures measured on interval scales like the Fahrenheit scale. Suppose that the temperature of rod *a* is 100°F and that of *b* 50°F. The uniqueness theorem for such measurement tells us that our scale is only unique up to a choice of unit *and* a zero point. Had we instead used the Celsius scale, which encodes the same objective information about temperatures as the Fahrenheit (in the precise sense that the same sort of representation and uniqueness theorems can be proved for each), we would have found the temperature of *a* to be 37.78°C and that of *b* to be 10°C, which is nowhere near a ratio of two to one. Hence, although the claim that *a* is twice as *long* as *b* reflects an objective fact about lengths, the claim that *a* is twice as *warm* as *b* reflects an artifact or idiosyncrasy of a particular temperature scale, and so ratios of scale values for temperatures have no direct representational significance.

The more nearly unique a scale, the more information about the world it conveys, and so the uniqueness properties of scales provide some indication of the degree to which they are *underdetermined* by the phenomena they represent. But how are we to make this precise? Measurement theorists say that objective properties and relations like ratios between lengths are *meaningful* or objective, whereas things like ratios of temperatures are not, and their basic idea is this. Claims like 'Rod *a* is twice as hot as rod *b*', which change truth value with a change in scale, reflect artifacts or idiosyncrasies of particular scales. This suggests that by focusing on those claims whose truth value remains unchanged, regardless of the scale employed, we might hope to filter out those cases which depend on idiosyncrasies of particular scales, leaving us with claims that are about scale-independent or meaningful features of the world.

A common implementation of this strategy focuses on a special set of mappings from the representing system back to itself. This set forms a group (in the algebraic sense) of what are called the *permissible transformations* of the system, and its distinctive feature is that mappings in it carry the numerical values assigned by one scale to a set of numbers that provide equally good scale values. Hence, if the scale c provides a legitimate representation of some magnitude like temperature or length, then the scale c' will be legitimate just in case there is a permissible transformation g of the representing system $\boldsymbol{B}$ such that $c'(x) = g \circ c(x)$ (where $\circ$ is function composition). We shall say that an n-place relation R in the representing system $\boldsymbol{B}$ is *invariant* under such a transformation just in case, whenever an n-tuple of original scale values $\langle c(\alpha_1), \ldots, c(\alpha_n)\rangle$ is in the extension of R, the n-tuple of new scale values $\langle c'(\alpha_1), \ldots, c'(\alpha_n)\rangle$ is also in the extension of R (and conversely). And since the relations that are invariant under the permissible transformations of the representing system are precisely those that are *not* sensitive to the idiosyncrasies of particular scales, it is natural to conclude that these, and these alone, are surrogates of objective or meaningful relations back in the system that is being represented.

All this will be clearer if we return to our example of length. The permissible transformations of the ratio scale for measuring length in meters simply multiply scale values by a positive real number. Multiplication of our original scale values, 100 and 50, by a number like 1.0936 (which converts meters to yards) will produce a pair of numbers that still stand in a ratio of two to one. Hence numerical ratios are invariant under permissible transformations of the representing system, and ratios of length are objective, scale-independent features of physical objects. By contrast, the permissible transformations of interval scales allow the addition of a real number (as well as multiplication by a positive constant), and as the example with our rods attests, such transformations need not preserve numerical ratios.

Examining the group of permissible transformations of a representing system in order to draw conclusions about the objectivity of properties and relations in the phenomena it portrays will only work in cases where we know which transformations are permissible, and why. In a handful of cases it was clear early on which transformations of the representing medium were permissible, and some types of representations were even classified in terms of them. For example, Stevens classified the more common scale types (ordinal, interval, and ratio scales) in terms of their permissible (scale) transformation (fn. 14). Something similar occurred earlier in Klein's Erlanger program, in which a number of well-known geometries (e.g., projective, equiform, and Euclidean geometry) were classified in terms of their permissible (coordinate) transformations. Unfortunately, disagreement remains about the proper characterization of permissibility in general, but by way of illustration we may consider one of the more widely accepted accounts, according to which the set of permissible transformations of the representing system is simply its group of *automorphisms*, i.e., its isomorphisms back onto itself.

Being an automorphism is sufficient for being a permissible transformation, since by definition each automorphism respects all the structure of the representing system. If c is an isomorphic embedding of the relational system $\boldsymbol{A}$ in $\boldsymbol{B}$ and g is an automorphism of the latter, then $g \circ c$ will also be an isomorphic embedding of $\boldsymbol{A}$ in $\boldsymbol{B}$, and so it too will provide a structural representation of $\boldsymbol{A}$ in $\boldsymbol{B}$. In the case of the more well-known scales and geometrical systems, like those studied by Stevens and Klein, being an automorphism is also necessary for permissibility, but the claim that this is so in general is somewhat controversial. Fortunately, the exact characterization of the group of permissible transformations is not critical here. What does matter is the idea that a group of permissible transformations, the general nature of which is usefully *illustrated* by the automorphism group, is likely to provide the best general way of separating those features of a structural representation with representational significance from those that are merely artifacts.[17]

These points bear on structural representation in the following way. Once we learn that some phenomenon is a species of structural representation, we often discover that various problems involving it are really problems about artifacts and meaningfulness, and this can prompt us to look to current work on meaningfulness for guidance in dealing with them. Doing so can be particularly useful when the phenomenon in question is not typically recognized as a species of representation, since it can suggest questions about it that would otherwise have escaped our attention. Examples of just this sort are discussed in Section 7.

Representational artifacts arise from *conventional choices* of some particular representation over others, and much of the philosophical significance of a uniqueness theorem is that it helps separate those aspects of a representation that are conventional from those that are not. Thus far, we have concentrated on conventions that arise *after* we have settled on some particular representing system like the positive, additive reals. These conventions involve our choice of one from among various legitimate alternatives for representing a phenomenon in *that fixed system*. But an equally important choice must be made about *which* system to use as a representation in the first place, and a second type of convention enters here. We commonly represent lengths in the positive, additive reals, but several other numerical systems would work just as well. For example, if we have a ratio scale c that embeds some particular extensive property system in the positive, additive reals, we can use it to concoct a second scale, c^*, that isomorphically embeds that very same *EPS* in the positive, *multiplicative* reals (let $c^* = \exp c$, and represent $\bigcirc$ by multiplication). The two scales will encode exactly the same information. Each thus provides an equally accurate and complete representation of the same objective facts, and differences between them will simply be representational artifacts. In short, we must distinguish what I shall call *systemic conventions*, which derive from the conventional choice of a representing relational system, from *mapping conventions*, which derive from the conventional choice of a mapping or representing function *once* a particular representing system has been selected. The former gives rise to *systemic artifacts*, the second to *mapping artifacts*.

Many artifacts, like the ratios of scale values in the measurement of temperature on a non-absolute scale, involve relations in the representing system that fail to portray relations in the domain of representation, but artifacts can also involve individuals. Often there will be individuals in the representing system that are not surrogates of anything in the phenomenon being represented. For example, we sometimes use the real numbers to represent countable collections of things, and in such cases, many of the numbers won't stand for anything at all. As with artifacts involving relations, it is possible to mistake individual artifacts for representationally significant features of the representation. Thus, we commonly represent facts about physical space in R^3. And, partisans of relational theories of space might well conjecture that realists about substantival space have been led astray by a systemic artifact, mistakenly supposing that each number is a surrogate of something – a point of physical space – when in fact it is not. This suggests the possibility that various forms of realism arise from mistaking artifacts of a representation for features with genuine representational significance. From this perspective, for example, Quine's claim that we are committed to the existence of sets, because quantification over them is required in the formalization of science, appears to confound a systemic artifact of one way of representing physics (namely, in mathematics reconstructed in some set theory or another) with the things physics itself is about.[18]

A brief example will illustrate the generality of these distinctions and suggest connections to recent discussions of conventionalism. In Section 7b we shall see how to treat the syntax of a language as a relational system whose domain of individuals includes words and whose domain of relations contains syntactic operations (e.g., sentential connectives) for combining words into sentences (cf. Montague [1974]). Such machinery allows us to view Quine's thought experiment about the radical translation of an alien language, A, into a home language, H, as a mapping of A to H that respects all (and only) the objective facts about meaning in A. These involve the stimulus meanings of A's observation sentences, stimulus analyticity and contradictoriness, the meanings of any truth-functional connectives A might contain, and a bit more that needn't concern us here (Quine, [1960], Section 15). Thus, from the current perspective a translation manual μ is a structural representation of A in H. It preserves the stimulus meaning of observation sentences, for example, as well as truth functions (e.g., for each binary truth function † of A and its surrogate ‡ in H, $\mu(\chi_1 \dagger \chi_2) = \mu(\chi_1) \ddagger \mu(\chi_2)$).

A claim about the alien language A reflects an artifact of a particular mapping just in case it would have had a different truth value had we used a different translation manual, μ^*. For example, according to one legitimate translation manual '*gavagai*' means 'rabbit', but according to an equally correct manual – one which respects all the same facts about meanings in A – it means 'undetached rabbit part'. If we take $\boldsymbol{G}$ to be the group of permissible transformations of the representing system (here English, construed as an *IRS*), this means that our claim about an alien sentence χ would shift truth value if we first mapped χ to English using μ and mapped the resulting English sentence to another English sentence using some transformation g in $\boldsymbol{G}$ (i.e., if we used the mapping $g \circ \mu$ to translate χ), in a manner analogous to the shifts encountered with transformation of scales in Section 4. Although Quine doesn't present his account in this way, doing so helps explain his otherwise cryptic remark that the totality of a speaker's sentences could be mapped onto itself in such a way that all of his dispositions to assent to, and dissent from, sentences remained invariant, yet the mapping was 'no mere correlation' of more-or-less equivalent sentences ([1960],

p. 27). In fact, such mappings are just the *permissible transformations* of the representing language.[19]

Radical translation involves both mapping conventions and systemic conventions; in Quine's words, specifying the ontology of a theory is 'doubly relative', first, to choice of background language and, second, to the choice of a translation manual from the target language to that background language ([1969], 54ff.). Just as our claims about ratios of temperatures reflect mapping artifacts arising from a conventional choice of a particular scale, the translation of some foreign term as 'rabbit' is a mapping artifact arising from a conventional choice of some particular translation manual. And just as our use of the positive, additive reals to represent length is a systemic convention, so is our choice of English as background language.

In the framework of structural representation, Quine's semantical conventionalism emerges as a claim that phenomena which previous thinkers took to be objective are really just artifacts of particular representations. It is an interesting question whether other prominent versions of conventionalism, like Grunbaum's doctrine of the conventionality of the metric (1973), also boil down to a similar sort of claim, e.g., that the metrical features of a manifold are just artifacts stemming from a conventional choice of one particular representation of spatio-temporal phenomena over others that are equally correct.[20]

However this may be, as long as we don't mistake artifacts for the real McCoy, they should not be a source of dismay. We are stuck with them, and they can often be turned to our advantage. A good example of the exploitation of a *mapping* artifact is von Neumann's identification of the less-than relation on numbers with the membership relation on sets, a maneuver that greatly facilitates many constructions and proofs in set theory (cf. Section 7a). It is also possible to exploit *systemic* artifacts; indeed, we frequently select representing systems because they give rise to exploitable artifacts. In order to have access to various mathematical concepts and techniques, for example, we often represent phenomena in numerical systems that have a much richer structure than the phenomena they are used to represent. Many assumptions about differentiability, the continuous distribution of random variables, and the like are plausibly viewed in this way. In Section 7 we will apply these points about representational artifacts and conventions to some concrete cases of structural representation, but first we need to develop a more general account of that notion itself.

6. GENERAL CHARACTERIZATION OF STRUCTURAL REPRESENTATION

The provisional account of structural representation in terms of isomorphic embeddings allows us to examine a number of issues in a reasonably simple way, but it is not sufficiently general. Given the intuitive motivations discussed in Section 1, the following sorts of situations should count as structural representations, even though they aren't so reckoned by the provisional account. As before, ***A*** is the *IRS* that is to be represented, ***B*** the representing system, and *c* a mapping from the former to the latter.

(i) In some cases of representation, the requirement that the two relational systems be of the same similarity type is too restrictive. For example, we might want primitive relations of ***A*** to be represented by defined relations of ***B***, or defined relations of ***A*** to be represented by primitive relations of ***B***.

(ii) In some cases things of one type are represented by things of another. In our treatment of lengths in Section 4, first order *properties* (determinate lengths) are represented by *individuals* (positive real numbers). We finessed this by treating both as members of the lowest-order domains of their respective systems, but this isn't always feasible. For example, we sometimes need to include a domain of individuals in a relational system containing lengths, and in such cases we need a more general provision for *trans-type representation*.

(iii) In some cases of representation, *c* does not respect all of the relations in the original system, but only some. For example, it is a basic geometrical fact that a two-dimensional projection of a sphere cannot depict all of its features without distortion, so when we use flat maps to represent the Earth, something has to give. For sixteenth-century mariners, concerned to convert lines of constant compass bearing (rhumb lines) into straight lines on their maps, Mercator's projection, which misrepresents scale, offered the best compromise; for other purposes equal areas maps, which accurately represent scale but distort shape, are preferable.

(iv) In some cases of representation, relations are respected only under certain conditions (e.g., boundary conditions). For example, a mercury thermometer may reliably represent the temperature if it is neither too hot nor too cold, but it would fare poorly in liquid helium or near the surface of the sun.

(v) In some cases of representation, *c* doesn't respect the relevant relations, but only preserves them in one direction or the other. For example, whenever a certain blood test indicates the presence of steroids, it is correct, but if the amount of steroids is small, the test may fail to detect them.

(vi) In some cases representation cannot involve a function *from* the original system *to* the representing system, since the relevant relation from ***A*** to ***B*** is many-one. For example, in linguistic representation one person may have two names.

(vii) In some cases there may be reasons to include individuals (or relations) in ***A*** that are not paired with anything in ***B***, so the requirement that the representing function be total is too restrictive. For example, ***A***'s domain of individuals might be the set of students in a university and ***B***'s a set of numbers representing their grades in Philosophy 101. Some students aren't enrolled in the course, and so receive no grade.

There is some overlap among these points, but they are worth separating because each suggests a different modification in our provisional account of structural representation.[21] I shall accommodate the first five points by identifying the structural representation of one *IRS* in another with a special sort of mapping from the first to the second. Once this is done, the account will be extended to subsume the sixth point as well. To accommodate the seventh point, we would need to weaken the requirement that *c* be total; this is certainly possible, but here I shall concentrate on the less obvious sorts of changes required to deal with the first six considerations.

In deference to the first consideration, we shall no longer require that ***A*** and ***B*** be of the same similarity type, or that *c* be one-one or onto (even with respect to the two systems' domains of relations). Second, in order to allow trans-type representation, we shall no longer require that *c* respect the level of relations; this allows properties and relations to be represented by individuals (e.g., lengths by numbers), and individuals to be represented by properties or relations (e.g., floors of a building by colors).[22] Third, to accommodate the fact that a representation need not respect all of the relations in the phenomenon that it represents, we shall let Θ be a subset of the relations in ***A***, and will say that c is a Θ-morphism just in case it respects all of the relations in Θ. The intuitive idea is that those relations in ***A*** that are members of Θ are accurately represented in ***B***. Fourth, in some cases properties and relations are accurately represented only under certain conditions; suppose, for example, that a given pan balance can only measure masses less than P. If c is a mapping from masses to reals, it would then be natural to be interested in such conditionals as: *If* $P > P_i$ and $P > P_j$, *then* $P_i > P_j$ just in case $c(P_i) > c(P_j)$. Operations that build more complex relations out of simpler ones (cf. p. 458) make it possible to give intensional definitions of new properties and relations right in an *IRS* itself, and in the present case, we can define a two-place relation that holds between two objects just in case both are less than P (as ordered by $>$) and the former bears $>$ to the second. In this way, we can build the relevant conditions under which a relation is respected into a more complex relation that is respected across the board.

What about point (v)? To say that a mapping from the total domain of ***A*** to that of ***B*** *respects* the n-place relation R is just to say that n things from ***A*** (taken in a given order) stand in R if and only if their n surrogates (taken in the corresponding order) stand in the relation in ***B*** that represents R. However, this biconditional can be split into two conditionals that are of independent interest. When we have the conditional running from left to right, viz., if $\langle x_1, \ldots, x_n\rangle \in {}^{\vee}R$, then $\langle c(x_1), \ldots, c(x_n)\rangle \in {}^{\vee}c(R)$, I shall say that *c preserves R*. And when we have the conditional running the opposite direction, viz., if $\langle c(x_1), \ldots, c(x_n)\rangle \in {}^{\vee}c(R)$, then $\langle x_1, \ldots, x_n\rangle \in {}^{\vee}R$, I shall say that *c counter-preserves R*.

When a property or relation is both preserved and counter-preserved, the resulting representation is an optimal indicator, telling us the whole truth and nothing but the truth about atomic facts involving that property or relation. When a mapping merely counter-preserves a property or relation, say property P, the representation's claims about atomic facts involving P will still be true, but since P is no longer preserved, an object might be P without the representation saying that it is. For example, if a representation says that a is P (i.e., if $c(a) \in {}^{\vee}c(P)$), then a is P, but a might be P without the representation saying so. Hence, the representation delivers *only*, but not *all*, true verdicts about P. The situation is reversed when P is preserved but not counter-preserved;

here, if an object is P, the representation will say that it is, but it might also claim that something is P that really is not. And much as a set of unsound inference rules vitiates a logic in a way that an incomplete set does not, failure of counter-preservation vitiates a representation in a way that failure of preservation does not. When counter-preservation fails, the representation's verdicts about the represented domain will sometimes be *wrong*, and surrogative reasoning based on it can lead to false conclusions, thereby thwarting its very purpose (cf. Mundy [1987a]). This suggests identifying structural representations with mappings that are like Θ-morphisms in all respects, except that they are only required to counter-preserve relations in Θ. However, I think that there are at least three reasons why an account of structural representation should find a place for preservation as well.

First, in some types of representation, particularly linguistic representation, the representing relation runs in the opposite direction from that in the cases encountered thus far, going *from* the representation (language) *to* what it represents (the world). In such cases a representation is truthful just in case relations are preserved.

Second, as contraposition shows, a relation is preserved just in case its negation is counter-preserved, so a mapping that preserves P will provide a good representation of the property being not P. For example, when a pan balance tells us that one object is more massive than another, it is probably right, and so inequality is counter-preserved. But if it fails to tell us that two objects differ in mass, this may simply be because it isn't sufficiently discriminating to detect small discrepancies, and so equality is not counter-preserved. However, the counter-preservation of inequality is equivalent to the preservation of equality (cf. Adams [1965]). Moreover, it is often somewhat arbitrary which relations we take as primitive and whether a particular relation is regarded as a negation or not, and so it is useful to leave a place in our account for preservation.

Third, surrogative reasoning often depends on the delicate interplay of a number of properties and relations, and in such cases preservation can be important. Imagine that the relational systems $\boldsymbol{A}$ and $\boldsymbol{B}$ each contain a (nonempty) domain of individuals and just two first-order properties, P^A and Q^A in $\boldsymbol{A}$, and P^B and Q^B in $\boldsymbol{B}$. Let c be a mapping from the total domain of $\boldsymbol{A}$ to that of $\boldsymbol{B}$ that pairs P^A with P^B and Q^A with Q^B. Now suppose that c preserves (but doesn't counter-preserve) P^A, and that it counter-preserves (but doesn't preserve) Q^A. Finally, imagine that the lone axiom for ***B*** asserts that all things in the extension of P^B are also in the extension of Q^B. Under these conditions, we can begin in ***A*** with an individual x that exemplifies P^A, and the fact that this property is preserved affords a bridge to the representing system ***B***, so that we can infer that x's surrogate, $c(x)$, exemplifies P^B. Reasoning with $c(x)$, we then deduce that it also exemplifies Q^B. Finally, the counter-preservation of Q^A allows us to make the return trip to system ***A***, and the conclusion that x exemplifies Q^A. Such reasoning would not be possible if both relations were merely counter-preserved, and similar points hold in more complicated cases as well. In short, even where the preservation of a relation is not of direct representational significance, it can be important in surrogative reasoning.

These considerations suggest the following (penultimate) definition of structural representation. When Δ and Ψ are subsets (at least one of which is nonempty) of ***A***'s full domain of relations and c is a function of the sort described in the four modifications proposed at the beginning of this section, c is a *Δ/Ψ-morphism* just in case it preserves all of the relations in Δ and counter-preserves all the relations in Ψ. We then identify the *structural representations of* ***A*** *in* ***B*** with those Δ/Ψ-morphisms from ***A*** to ***B*** in which Ψ is nonempty, i.e., in which at least one relation of ***A*** is counter-preserved. In such a case the image of the set Ψ under c contains the (primitive) relations in ***B*** that are of direct representational significance.[23]

Thus far we have required that the representing relation be a function from the things represented to the medium of representation. But in linguistic representation, something can have more than one name, so that the relevant relation from the object of representation to the representation itself is one-many and, hence, not a function. However, in those languages where each term has only one meaning, there is often an important mapping running in the opposite direction, from language to the world (in formal semantics, this is the interpretation function that assigns denotations to terms). This mapping *is* a function, and it will underwrite surrogative reasoning of a sort. True, we cannot begin with some fact in the world and move to the sentence representing it (since several different sentences might do so). But we can frequently move to *some sentence or other* that does the job, and often it doesn't greatly matter which we pick. As long as we get to *one* of the sentences that represents the original fact, we can reason in language and, once we are finished, make the return trip to a conclusion about the world.

As we shall see, the details of linguistic representation are complex, but my concern here is just to motivate an extension of our present account of structural representation, so that one *IRS* can structurally represent a second when there is a Δ/Ψ-morphism from either to the other. More precisely, I shall say that one *IRS* represents a second just in case there is a Δ/Ψ-morphism from one to the other in which Ψ is nonempty (if the representing function runs from the object of representation to the representation) and Δ is nonempty (if it runs the other way).

Allowing mappings besides isomorphic embeddings to count as structural representations is necessary if our account is to be sufficiently general, but it does complicate our earlier picture of surrogative reasoning, according to which we begin with facts in the domain of representation, pass over to a representation to reason about their surrogates, and then return with a conclusion about the original system. As our recent example with P^A and Q^A shows, neither the route from $\boldsymbol{A}$ to $\boldsymbol{B}$, nor the route back, need exist in all cases, and when either doesn't, uncritical, wholesale surrogative reasoning is illegitimate. Just which bits of surrogative reasoning are possible in a given situation will depend on which relations (primitive and defined) are preserved and counter-preserved, and this, in turn, will depend on the features of the particular mapping. Still, some general morals can be drawn. For example, if a set of relations is preserved, relations defined in terms of them and analogues of conjunction and existential quantification will be preserved (so we can move from claims about these to their surrogates), and relations defined in terms of the analogues of negation and universal quantification will be counter-preserved (so we can move from claims about these back to the original system).

Additional requirements might be imposed on Δ/Ψ-morphisms in order to obtain various species of structural representation. As we'll see in Section 7a, in some cases a mapping will provide a useful representation only if it is general recursive. In other cases it is important that a representational system display counterfactual sensitivity, so that had the structure of the represented domain been different, the structure of the representational system would have differed to follow suit. For a thermometer to genuinely represent the temperature, it is not enough that it always happens to give the right reading, as it would if it were stuck on 50°C and spent its life in water carefully maintained at that temperature. It must also be the case that if the temperature had been different, the thermometer would still have given the correct reading.[24] There is little doubt that so-called *externalist* relations, like counterfactual sensitivity and causation, play an important role in many types of structural representation that we find useful. Still, our frequent use of numbers and sets as media of representation shows that structural representation can occur in their absence, and so we should not build them into our account.

I should not wish to claim that the present account is the last word on structural representation, but I hope I have said enough to support the claim that the story in terms of Δ/Ψ-morphisms is on the right track. It provides a definite working characterization of structural representation that satisfies all five of the desiderata at the end of Section 1, and it displays a variety of types of representation as sharing a common nature in virtue of which they are representations. The real test, however, comes in seeing whether the account helps clarify and explain specific examples of philosophical interest. In the next section I shall argue that it does.

7. ILLUSTRATIONS AND APPLICATIONS

Ideally, the aim in exhibiting a phenomenon as a structural representation is to develop such a precise account of it that we can prove representation and uniqueness theorems. But an examination of a phenomenon in the present framework can be fruitful even when we don't achieve this goal, since it is still likely to suggest fresh questions about the phenomenon and to furnish new concepts and techniques for its study. In order to illustrate the range of the notion, I shall consider four examples of structural representation here, beginning with one that is reasonably straightforward and working my way up to one that is rather more speculative.

7a. *Ontological Reduction*

A variety of things go by the name of *ontological reduction*. Some, like phenomenalism, behaviorism, and methodological individualism, were never developed very satisfactorily. Others, notably reductions of various number systems to set theory and of geometry to arithmetic, were carried out in impressive detail. I shall consider the more successful sort of reduction here, focusing for definiteness on reductions of the natural numbers to sets. We shall find that sets provide structural representations of the natural numbers in a quite straightforward way, but my chief claim is that the philosophical significance of ontological reductions of numbers to sets lies wholly in this fact.

A reduction of the natural numbers to sets should include a (recursive) pairing of numbers with sets that allows us to use the axioms of the reducing set theory to prove that the relevant sets have the right (numerical) properties. Such reductions by Frege, Zermelo, and von Neumann are well known, but they can be interpreted in various ways. Some philosophers, including Frege himself, regard them as discoveries, telling us what the numbers really *are* (in much the way that the identification of water with H_2O tells us what water really is). On such a construal, there can be only one correct reduction. If the natural numbers literally *are* the sets that von Neumann took them to be, they cannot *also* be the sets that Zermelo thought they were, since (with two exceptions) these thinkers disagreed about which set each particular number was. Thus, Zermelo identified 2 with the set $\{\{\emptyset\}\}$, whereas von Neumann identified it with $\{\{\emptyset\}, \emptyset\}$, and it's about as simple a theorem of set theory as one could hope to find that these two sets are distinct. Hence, by the transitivity of identity, 2 cannot be identical with each, and so multiple reductions, construed as identifications, come to grief over the logic of identity.[25]

We might attempt to avoid this conclusion by simply declaring that some particular correlation of numbers with sets is uniquely correct. However, alternative reductions clash at points having little to do with our original beliefs about numbers. Worse yet, no one has ever identified features of any particular reduction that provide much support for the claim that it is the one true story about what numbers really are. Indeed, many of the competing accounts have comparable explanatory value, unifying power, comprehensiveness, simplicity, and all the other virtues routinely cited as canons for theory choice. To be sure, some reductions have features that make them easier to work with than others; for example, von Neumann's identification of the less-than relation on numbers with the membership relation on sets can be extremely useful for many purposes. But it is difficult to see why usefulness should count for much in determining what the numbers really are. More importantly, even if it did, it wouldn't settle the present issue, since different reductions can be useful for different purposes.

When an ontological reduction is construed as telling us what numbers actually are, the problem of multiple reductions seems insuperable. But if we regard accounts like Zermelo's and von Neumann's as providing alternative *representations* of the natural numbers by sets, the difficulty vanishes. Multiple representations simply depict the same thing in different (and often complementary) ways, and differences among them can quite legitimately be regarded as artifacts, as Quine's "don't cares". Furthermore, if ontological reductions are *structural* representations, we can explain why they are so fruitful even if numbers are distinct from sets.

In order to exhibit reductions of the natural numbers to sets as structural representations, I shall treat the numbers as the *Natural Number relational system* $\boldsymbol{N} = \langle N, s, <, ^{\vee} \rangle$, where N is the set of natural numbers, s and $<$ are the successor function and less-than relation, $^{\vee}$ is one-one (for convenience let it be the identity mapping), and the system behaves in accordance with Peano's postulates. For purposes of illustration, I shall represent the system $\boldsymbol{N}$ in von Neumann's sets, though the following points would hold if we used some alternative, like Zermelo's. Let us say that a *von Neumann relational system* is an ordered quadruple $\boldsymbol{V} = \langle V, \sigma, \in, ^{\vee} \rangle$, where V is a nonempty set (the von Neumann classes), σ is a function from V to V such that for all x in V, $\sigma(x) = x \cup \{x\}$, $\in$ is the relation of set membership, and $^{\vee}$ the identity mapping. We can axiomatize $\boldsymbol{V}$ in set theory, using the axiom of infinity to guarantee that there is at least one inductive set (one containing $\emptyset$ and closed under σ), then singling V out as the smallest set of this kind (i.e., as a subset of every inductive set). Thus, V contains all and only the von Neumann classes, $\emptyset$, $\{\emptyset\}$, $\{\emptyset, \{\emptyset\}\}$, $\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$, and so on up, and they are ordered by $\in$.

We can now provide a structural representation of $\boldsymbol{N}$ in $\boldsymbol{V}$ via a function c that pairs the relation $<$ with $\in$, the (functional) relation s with σ, and the individuals 0 with $\emptyset$, 1 with $\{\emptyset\}$, and so on (more generally, $c(0) = \emptyset$ and $c(s(n)) = \sigma(c(n))$). It is then possible to prove a representation theorem showing that c is a Δ/Ψ-morphism in which *all* of the relations in $\boldsymbol{N}$ are preserved and counter-preserved; that is, for every natural number m and n in N, (***A***) $m < n$ just in case $c(m) \in c(n)$, and (***B***) $c(s(n)) = \sigma(c(n))$. And this similarity of structure justifies our acting *as if* the natural numbers were sets when reasoning about numbers.

What about artifacts and uniqueness? The relational system **V** has no non-trivial automorphisms, and the only way for a mapping from its domain of individuals back to itself to preserve both σ and $\in$ is for it to shift everything up the line some fixed number of steps (e.g., a shift of one step maps to $\emptyset$ to $\{\emptyset\}$, $\{\emptyset\}$ to $\{\emptyset, \{\emptyset\}\}$, etc.). We can then define a set-theoretic operation $\oplus$ in **V** that mimics addition (it's just that operation that reductionists take to *be* addition), and any permissible transformation c' of the mapping c will have the form $c' = g \circ c$, where $g(x) = x \oplus k$ (for some set k in **V**). This provides a picture of the permissible transformations of **V** *relative to* the "standard" mapping c. More generally, we can take the set of *these* transformations along with their inverses (which will not be total functions on V, since if g shifts everything up n places, g^{-1} will be undefined for the first n von Neumann sets), and any two isomorphic embeddings of **N** in **V** will be related by one of them.

Features of the standard mapping c that are absent from other legitimate mappings from **N** to **V** are *mapping artifacts*. For example, if each natural number is identified with the set to which c maps it, then the number n will have exactly n members. This is often cited as an anomaly of von Neumann's account, but on the present construal it is simply a mapping artifact of his particular representation. This feature is often convenient, but it has no representational significance, and it will be absent from other, equally good, representations of **N** in **V** (it is missing, for example, from the mapping that pairs 0 with $\{\emptyset\}$, 1 with $\{\emptyset, \{\emptyset\}\}$, etc.). By contrast, once we decide to use **V** as the representing system, it will turn out that on every acceptable representation $m < n$ just in case $c(m) \in c(n)$ just in case $c(m) \subseteq c(n)$. These relationships will be invariant across all representations of **N** in **V**, but since they will be absent from other, equally legitimate, representations of numbers by sets (like Zermelo's), they are *systemic artifacts* of the use of **V** as the medium of representation.

Philosophers like Russell, Goodman, and Quine have taken a more restrained view than Frege, holding that ontological reductions are not *identifications*, but merely *explications*, demonstrations that sets can do the work for which numbers were originally thought to be required. The most commonly cited motivations for explications are to achieve ontological parsimony, to obtain epistemological security, and to replace murky or defective notions by clearer ones. These considerations are not irresistible, however, and I think there are good reasons to construe reductions as representations rather than as identifications or as replacements.

Although ontological parsimony is always welcome, it is difficult to see why economy for its own sake should be of overriding importance in the present case. It would certainly be of interest to learn that there were no abstract objects at all. But ontological reductions of the natural numbers to sets already take set theory at face value and, hence, are committed to an enormous number of abstract objects – far more than the natural numbers – that pose at least as serious metaphysical and semantical difficulties as the numbers themselves. How about epistemological security? Years ago it was hoped that reductions would show that mathematical truth was simply a species of the less mysterious genus of logical truth. But as Gödel's theorem shows, we can at best get a "reduction" of number theory to *second-order* logic, and logical truth here is really no clearer than arithmetical or, indeed, set-theoretic truth (indeed, questions about such perplexing things as the truth of the continuum hypothesis can be expressed as questions about the logical truth of certain second-order formulas). And what of the Quinean quest for clarity (e.g., [1960], 257ff.)? In light of Peano's postulates, the notion of a natural number can't be said to be objectionably imprecise. In fact, it's a good deal clearer than the notion of a set, which, as recent proposals for axioms to settle the continuum hypothesis or debates about the axiom of foundation show, is far from being clear or intuitive (or epistemologically secure).

So-called 'reductions' and 'explications' are often quite valuable, but rarely, I think, for the three reasons just discussed. In general, the best motivations for them are just those that we have found for structural representations in general. The representation theorem for length measurement justifies, and explains the success of, the use of numbers in surrogative reasoning about lengths. Just so, the representation theorem for numbers and sets justifies, and explains the success of, the use of set theory in surrogative reasoning about the natural numbers. Sets provide an extremely effective medium of representation. This is not so because the notion of a set is particularly clear, however, but because set theory is a powerful and well-developed theory rich in concepts, theorems, and techniques for proofs that greatly facilitate surrogative reasoning. Thus, set theory provides a unifying framework, a mathematical *lingua franca*, in which an otherwise bewildering variety of theories can be represented and compared.

Construing ontological reductions as structural representations rather than as exercises in ontology does leave us with the important question of what the natural numbers really *are*. I have adopted a naive realism about numbers here. However, the conclusion that judiciously selected sets can be used to represent numbers because they have all of the structure of the numbers suggests a further step, according to which it is this structure, rather than any individual objects, that forms the proper subject matter of arithmetic.[26] But however this may be, there are many instances of powerful theories that were developed in detail and widely applied in the absence of any satisfactory account of their ontological underpinnings. Familiar examples include differential and integral calculus, probability theory, set theory, and number theory itself. In each case we knew – and arguably still do know – far more about the structure of the things and relations these theories deal with than we know about the things and relations themselves. Thus we know a good deal about the structure of numerical relations like greater-than and addition, but little about what (if anything) the numbers themselves really are. Similar situations arise outside pure mathematics; we quite successfully represent quantum mechanical systems in Hilbert spaces, but it is notoriously difficult to understand the nature of the real-life systems that these represent.

The moral is that for many purposes a good representation of something is more useful than the metaphysical truth about it. But this is not to disparage metaphysics; indeed, it is common to work backward, from a representation, to a more direct account of the phenomenon that it represents. For example, qualitative accounts of various sorts of measurement, probability, and scientific theories have typically come long after the development of their associated mathematical representations. Thus a good representation of the structure of something is often an important first step toward a satisfactory metaphysical account of it. The fact that quantum mechanical systems can be painted onto Hilbert spaces tells us *something* about what such systems are like, and thus provides a clear and definite starting point for an account of the intrinsic nature of such systems.

7b. Linguistic Representation and Compositional Semantics

Talk of 'linguistic representation' is ubiquitous, but can it be made precise and, if so, will it turn out to be a species of structural representation? The answer to both questions is yes, at least for languages that can be given a compositional semantics. In itself, this is scarcely noteworthy, since on the most influential account of the matter, namely Montague's (1974b), a compositional semantics simply *is* one that is based on a structure-preserving mapping (a homomorphism, to be precise) from the syntax to the semantics. Interesting points emerge, however, when this fact is examined in the broader framework of structural representation.[27]

The distinctive feature of a compositional semantic theory for a language is that the meaning of each complex expression of that language is completely determined by the meanings of its component expressions and their syntactic arrangement. Corresponding to each syntactic mode of combination, a compositional theory provides a semantic operation that determines the meanings of expressions combined in that way. For example, formulations of sentential logic often include a syntactic rule telling us that if φ and ψ are sentences, then $\ulcorner\varphi \,\&\, \psi\urcorner$ is a sentence too, and corresponding to this we have a semantic operation (here a truth function) according to which $\ulcorner\varphi \,\&\, \psi\urcorner$ is true just in case both φ and ψ are true.

In compositional theories, there is an obvious sense in which the syntax mirrors the semantics, but in order to exhibit this similarity as a full-fledged case of structural representation, we must treat it as a mapping from one *IRS* to another. This can be done using Montague's elegant theory of meaning, in which we treat both the syntax and the semantics as relational systems, with the former including syntactic operations like connectives and quantifiers, and the latter containing semantic operations corresponding to these ([1974], Section 3). A mapping from the syntactic relational system to the semantic system is provided by an interpretation function that assigns a meaning or semantic value to every simple expression of the language and which then systematically assigns meanings to complex expressions in the usual way. The key here is that the mapping preserve the structure of relations (all of which are operations, i.e., functions), in the syntactic relational system. For example, in the case of each binary syntactic operation F and its associated semantic operation G, this means that the interpretation function I should conform to the schema (***PF***) $I(x\,F\,y) = I(x)\,G\,I(y)$.[28]

By way of illustration, consider a formulation of sentential logic with the stroke as the only primitive connective (since this logic is extensional, I shall suppress the extension assignment in this example). On the syntactic side we begin with the relational system **Syn*** = $\langle Sent, At, |_F \rangle$, where *Sent* is the set of all sentences of a language for sentential logic, *At* is the set of atomic sentences, and $|_F$ is the syntactic operation that carries each pair of sentences φ and ψ to their combination $\ulcorner \varphi\, |_F\, \psi \urcorner$. From this we extract the relational system **Syn** = $\langle Sent, |_F \rangle$. On the semantic side we introduce the relational system **Sem** = $\langle \{t, f\}, |_G \rangle$, where $\{t, f\}$ is a set of appropriate semantic values for the sentences in **Syn** (here truth values), and $|_G$ is the truth function that yields the value t exactly when either of its arguments is f. Finally our representation theorem shows that an interpretation is a homomorphism from **Syn** to **Sem**. This just means that it is a mapping that assigns each atomic sentence a truth value and preserves the structure of $|_F$, i.e., for every pair of sentences φ and ψ in *Sent*, $I(\varphi\, |_F\, \psi) = I(\varphi)\, |_G\, I(\psi)$. Matters become more complicated when we add quantifiers, but thanks to Tarski's inventive use of operations on infinite sequences of objects, this can be done in conformity with schema like (***PF***), and the approach works for a number of more complicated languages as well.[29]

What about artifacts and uniqueness? In purely formal logic, the fact that two expressions have the same semantic value on some particular interpretation is of little interest, but the fact that two expressions have the same semantic value on every interpretation is. This suggests that the semantic value assigned to an expression on some particular interpretations is simply a mapping artifact of that interpretation. No interpretation is privileged, and idiosyncrasies of particular interpretations (like the assignment of the truth value T to the sentence letter 'p') are of no logical significance. What *does* matter, what is logically meaningful, are features common to *all* interpretations (like the assignment of T to '$p \vee q$' on each interpretation in which T is assigned to 'p'). For it is these that figure in the definitions of basic semantic notions like consistency and entailment. In the case of extensive measurement, we have one degree of freedom; once a unit is selected, all the remaining scale values fall into place. In familiar logical systems, we have a countably infinite number of degrees of freedom, since the assignment of a semantic value to each primitive expression of the language is independent of the semantic value assigned every other expression. But it can be shown that there is exactly one homomorphism from the syntax to the semantics that extends an assignment of semantic values to the primitive expressions to an assignment to all expressions. Consequently, once semantic values are assigned to the primitive expressions, all of the other semantic values fall into place. So an interpretation is unique up to an assignment of values to syntactic primitives.

What are the prospects for extending the picture of linguistic representation as structural representation from formal languages, like first-order logic, to natural languages, like English? An obvious obstacle is that the surface structures of English sentences do not line up neatly with systematic semantic accounts. However, this mismatch between actual syntax and formal semantics motivated the introduction of the theoretical notion of *logical form*, and it may well turn out that when English is redescribed at a more theoretical level of logical form, it's so-called *deep structure* will provide a structural representation of various aspects of reality.

The prospects for this would be clearer if the prospects for developing a compositional semantic theory for a natural language (thus redescribed) were clearer. Such theories have been devised for substantial fragments of various natural languages, and this provides some reason to think that compositional accounts are viable. It is true that various objections have been raised against such accounts, but many of these merely show that it is impossible to devise a compositional semantics that has certain additional features. And in at least many cases, it's not obvious that our linguistic theories should have the additional features the objector deems important. Hence, it seems an open, and largely empirical, question whether compositional theories will work for natural languages. All that can be said now is that if a satisfactory compositional semantics can be developed for a natural language, or even for an interesting fragment of one, that language or fragment would provide a structural representation of reality and could be studied in the framework developed here.

Thus far we have thought of language as a representation of the world, but it is possible to reverse this picture and think of an interpretation of a language as a representation *of it*. As before, the representing function runs from syntax to semantics, but on the earlier picture the syntax represents the semantics, while according to the current suggestion the semantics represents the syntax. Something rather like the second picture is at work in Davidson's account of interpretation (e.g., [1984]), and it is very similar to the examples of structural representation in earlier sections, in which the representing function ran from the thing represented to the representation. I shall call the function from expressions of a language to their semantic values a *scheme of reference*. And for simplicity, I shall imagine that we are concerned with a simple, first-order fragment of English that contains just names and one-place predicates which schemes of reference map to objects and to sets of objects, respectively. We could then begin with a scheme c that assigns the intuitively correct objects to names (e.g., George Bush to 'George Bush') and sets of objects to predicates (e.g., the set of gentle things to 'gentle'). In concert with the standard, Tarskian recursive apparatus, c would then assign truth conditions to all the sentences of the language, telling us, for example, that 'George Bush is gentle' is true just in case c(George Bush) $\in$ c(gentle). Indeed, we can even prove a sort of representation theorem, showing that a sentence is true just in case its truth conditions obtain.[30]

This approach yields a picture of interpretation as a sort of measurement. And just as there can be different, but equally good, scales for measuring temperature, an example in Wallace (1977) suggests that there can be different, but equally good, schemes of reference. The set of semantic values for names of our language contains a number of objects like George Bush, and Wallace invites us to consider a one-one mapping, ψ, of this set back onto itself. We can then extend ψ to a mapping that also carries the power set of the universe of objects back to itself in a way that it systematically undoes the changes wrought by ψ with respect to individuals (for example, a one-place relation S is a subset of the universe, and $\psi S = \{x: \psi^{-1} \in S\}$). The mapping ψ is not an automorphism (it doesn't preserve sets like S), but it does respect structure in the sense that any individual x is in S just in case $\psi(x)$ is in $\psi(S)$. Mappings like ψ carry schemes of reference to equally good schemes of reference, in the sense that each assigns the same truth conditions to every sentence of the language (speaking loosely, both will pair any given sentence with the same state of affairs). For example, since c is a scheme of reference, $c^* = \psi \circ c$ will be one as well, and 'Fa' is true just in case $c(a) \in c(F)$, which in turn holds just in case $c^*(a) \in c^*(F)$. Hence, if we think of our universe of objects and the sets and relations-in-extension on them as a representing relational system, it is natural to think of the group of mappings defined in this way as its permissible transformations.

In the case of the measurement of temperature on an interval scale, the objective facts involve things like ratios of differences of temperatures, and these can be represented in equally good ways by scales with different units and zero points. Similarly, on the current picture of interpretation, sentences are the objective units of meaning, and their truth conditions can be represented in equally good ways by schemes of reference which tell different stories about reference and satisfaction. Moreover, just as legitimate scales are related by a group of transformations, so too are legitimate schemes of reference. And since different schemes of reference may be equally correct, claims like those about the referents of proper names, whose truth value depends on the use of one scheme rather than another, will reflect *mapping artifacts*

Whatever the plausibility of such claims, they are interesting here as treatments (perhaps unwitting ones) of the uniqueness problem in semantics and interpretation. More generally, the treatment of linguistic representation as structural representation enables us to examine it from a perspective in which questions about the existence and uniqueness of representations become questions about topics like indeterminacy that are of interest from a semantic point of view. It also locates semantics in a more general context, which allows us to compare it with other forms of structural representation, including some species of mental representation.

7c. Mental Representations

Numerous writers have suggested that many of our thoughts are representations or models of reality. Nearly a hundred years ago, the physicist Heinrich Hertz wrote that when we think, "we form for ourselves images or symbols of external objects; and the form which we give them is such that the necessary consequences of the images in thought are always the images of the necessary consequences in nature of the things pictured. In order that this requirement may be satisfied, there must be a certain conformity between nature and our thought" ([1956], p. 1). Fifty years later the psychologist Kenneth Craik transformed this claim into a scientific hypothesis, urging that thought involves mental representations or models that have the 'same relation-structure' as the things they represent. Craik isolated three steps in the mind's modeling of reality: we 'translate' facts into corresponding mental representatives or surrogates, manipulate these in thought, then make the return trip to a conclusion about the world in the form of a prediction or action ([1943], esp. Ch. 5). These three steps correspond exactly to the steps in our earlier examples of surrogative reasoning, for instance, to our 'translating' facts involving lengths into their numerical surrogates, engaging in mathematical reasoning, then making the return trip to a conclusion about our original objects.

On Craik's view, the action of psychological operations on mental tokens (presumably neural states of some sort) is as much a case of surrogative reasoning as our explicit numerical reasoning about lengths. In the case of measurement, an isomorphic embedding explains the applicability of mathematics to physical objects and their lengths. In the case of mental representation, a hypothesis about the representation's structural similarities to selected aspects of the world aims to explain *the applicability of thought to reality*. Because relational systems were traditionally taken to be extensional, they were ill-suited for developing such ideas about thought (whose very hallmark is intentionality), but with the introduction of intensional relational systems, this is no longer the case.

According to most recent accounts, mental representations are theoretical entities, often inaccessible to introspection, so questions about their nature, and even their existence, are empirical ones to be answered by cognitive scientists. But questions about the structure that various sorts of representations have, *if* current hypotheses about them are right, can be studied independently of empirical investigation, and these will be our concern here. Caution is needed in treating psychological operations on mental representations as a species of surrogative reasoning, however, since we don't reason with mental representations in the same way that we reason with diagrams or maps. When we reason with a diagram, we doubtless represent it in thought. But reasoning with (at least some) mental representations cannot require that we represent *them* in thought, on pain of a vitiating regress. Reasoning about a diagram may involve the manipulation of a mental representation *of* it, but reasoning with a mental representation just is the manipulation (conscious or not) *of that representation*. The activity of at least some mental operations is the end of the line, the place where the representational buck stops, and it is this activity that is constitutive of surrogative reasoning with mental representations.[31]

Some thinkers contend that at least many of our mental representations have the same structure as sentences of natural or formal languages (e.g., Fodor [1975]). If this is right, then we may reasonably hope to devise a compositional semantics for such representations, and so the issues here are not substantially different from those discussed in the previous subsection. Here I will consider two rather different types of mental representation, Philip Johnson-Laird's mental models and Roger Shepard's mental images.

Johnson-Laird's account of *mental models* is striking in its detail, scope and empirical support (1983). Although his concerns are quite different from mine, his mental models involve just the sorts of structural relationships studied above. Indeed, he holds that the feature that distinguishes mental models from other types of mental representation (like semantic networks) is that their structure is "identical to the structure of the states of affairs . . . that the models represent" ([1983], p. 419). Johnson-Laird is concerned with the actual psychological processes of human beings, and so he places additional constraints on mental models that do not hold for structural representations in general. For example, he maintains that the operations involved in the construction and employment of mental models are computable (because he believes thought is computational), and that mental models contain a finite number of elements (because the brain is finite). But such additional requirements in no way conflict with my account, and mental models are an interesting and important species of structural representation.

Johnson-Laird argues that his theory of mental models affords the best available explanation of a number of important psychological phenomena, including our perception of the world, understanding of discourse, and control of bodily movement. But for our purposes, the role that mental models are hypothesized to play in deductive inference is of particular interest, since it is part of an explicit, empirical theory about surrogative reasoning with mental representations. On this account, everyday inference rarely involves the application of syntactic rules like modus ponens or resolution. Instead, we construct a mental model that embodies the information contained in a set of premises (as well as additional general information that seems relevant). We then examine this model and draw a conclusion from it that is not stated explicitly in the premises (and that doesn't discard semantic information). Finally we search for alternative models of the premises that would falsify the conclusion, and if none are found, the argument is judged to be valid. The process is quite fallible, but it should be, since the goal is to explain actual human reasoning, with all its infirmities.

Much remains to be done in developing this account, but Johnson-Laird argues that it does a better job of explaining various features of human inference than its competitors do. For example, it explains why someone confronted with several arguments of the same form fares better with those involving familiar subject matter (it's held to be easier to construct mental models of familiar things). It also predicts that the greater the number of models that have to be constructed to draw a valid inference, the more time it will take us to do so, and the more prone to error we will be, and there is some evidence to bear this out. Finally, a computer program based on this account produces patterns of success and failure with inferences that are quite similar to those of human subjects.

Johnson-Laird's account provides evidence that structural representations play an important role in our mental life, and it suggests that the activities of many psychological operations are in fact instances of surrogative reasoning. On the other hand, our account of structural representation should provide a useful philosophical foundation for Johnson-Laird's theory by situating it in a more general account of representation. Viewing mental models as structural representations also suggests possible modifications in Johnson-Laird's account. For example, his requirement that the structure of mental models be *identical* with the structure of the states of affairs that the models represent not only risks endowing us with more acuity than we actually have, it would also make us inefficient. My conclusion that Jack is taller than Frank, because Jack is taller than Dan, who in turn is taller than Frank, makes no use of my knowledge that Dan is a plumber. In such cases it would be quite inefficient to employ mental models that included such irrelevant information; indeed, it is a plausible conjecture that mental models, like many other representations, incorporate a number of simplifications and even idealizations.

Johnson-Laird's claims about the structural identity of mental models and the things they represent is based in part on his view that a representation should be economical, so that none of its elements (including relations) and none of its structure should lack representational significance. This suggests that he doesn't mean that every (known) aspect of something will be represented in a mental model of it, but merely that every aspect of the model will represent some aspect of the thing. This seems doubtful on his own grounds, however, since he holds that a *general* description of a situation is represented by a single, quite specific, mental model, which stands for all the instances of the general situation. And such specific models are bound to contain various features which don't correspond to features of the more general situation. But a more important reason for suspecting that mental models have features of no representational significance is that virtually all representations contain surplus features or artifacts, and it would be surprising if mental representations turned out to be an exception.

We can accommodate these points while retaining the spirit of Johnson-Laird's account by requiring simply that mental models be structural representations (rather than isomorphisms) that satisfy Johnson-Laird's remaining constraints (e.g., his requirement that they contain only a finite number of elements).[32] This friendly amendment also suggests various empirical questions about mental models. First, what sorts of Δ/Ψ-morphisms (if any) do various types of mental models involve, and which sorts of relations do they preserve or counter-preserve? Second, what sorts of simplifications and idealizations (if any) do mental models incorporate? Third, what sorts of artifacts (if any) do mental models contain, and do people sometimes err because artifacts are mistaken for actual features of the situation being modeled? Fourth, what are the uniqueness properties of mental models; how much convention do mental models involve, and are legitimate models related by an interesting group of transformations?

Mental images are typically thought of as a special sort of mental representation, in many ways akin to pictures. In an elegant series of experiments over the last twenty years, Roger Shepard and his co-workers have investigated the properties of mental images and their transformations. In a typical experiment, subjects are shown drawings of pairs of angular, three-dimensional objects with differing orientations. Each drawing depicts either the same figure from two different perspectives, or else two different figures that are mirror images of each other. Subjects are then asked to judge whether each pair of pictures displays the same figure from different perspectives or not (e.g., Metzler and Shepard [1982]). Later many of them reported that they began with one figure and imagined it rotating smoothly until it was congruent – or clearly failed to be congruent – with the other figure in the picture. Furthermore, the time they took to decide whether the figures were the same was proportional to the angle through which one of the figures would actually have to be rotated into congruence with the other. Metzler and Shepard concluded that their subjects employed some sort of mental analog of physical rotation in order to transform an image of one figure in a continuous way into an image of the other.

If correct, this account suggests that some surrogative reasoning involves a type of mental operation that is quite different from reasoning in language. Images are not spatial pictures in the head, however, and so they cannot literally undergo spatial rotation. In what sense, then, did subjects' psychological operations correspond to rotations? To answer such questions, Shepard and Chipman introduced a hypothesis about what they called *second-order isomorphism*, according to which there is an approximate parallelism "between the relations among different internal representations and the relations among their corresponding external objects" ([1970], p.1). That is, they hypothesized that there is a structural similarity between an actual rotation of a physical figure, on the one hand, and the mental transformation of its image, on the other.

Such talk is sometimes criticized as unduly metaphorical, but from the current perspective second-order isomorphism is just an instance of structural representation in which higher-order structure is preserved (and counter-preserved), and it could be explained in the following way. The fact that the medium for visual images is so good at encoding information about the geometrical properties of physical objects suggests that it has a structure that derives from relations with at least some of the same formal features as ordinary spatial relations like incidence, betweenness, and congruence. If so, it should be possible to provide an *IRS* model of this medium using what I shall call a Shepard *IRS*.

Geometries are often treated as relational systems which include a set of points and relations like congruence. Geometers frequently step back and talk about various sorts of transformations (like collineations or rigid rotations) of these systems, but such transformations are simply mappings from the set of points back onto itself, and we could just as well expand a geometrical relational system to include them. Similarly, we might think of a Shepard *IRS* as containing a set of individuals, relations among these, and a group of transformations of this set. It is an empirical question just what these individuals would be and whether their structure would be Euclidean, hyperbolic, or the like. But to fix ideas, we might imagine the individuals along the lines of pixels in computer graphics, the relations as ones sharing at least some of the topological and metrical properties of ordinary spatial relations like betweenness and congruence, and the transformations as mappings with at least some of the same structural features as Euclidean transformations like rotations, translations, or reflections.

Three steps would have to be completed to develop this idea. First, in order to secure a grip on the geometrical features of the physical objects that images represent, we would have to adopt some set of axioms for ordinary geometrical relational systems that included various transformations of their sets of points. Second, drawing on current data and theory, we would need to devise axioms for a Shepard *IRS* which determine the structure of such relations as psychological coincidence or psychological rotation. Third, we would have to use these axioms to prove a representation theorem, showing that any ordinary geometrical relational system could be mapped to some Shepard *IRS* system in a way that preserved and counter-preserved appropriate relations; in particular, it should preserve at least some structural features (e.g., continuity) of various transformations of the Euclidean plane. Such a representation theorem would justify our use of imagery in surrogative reasoning about spatial configurations, explaining why, for example, we can represent the current orientation of the piano and the door by images, manipulate these in thought, then translate the result back into a decision about how best to tilt the piano to fit it through the door. Finally, we could test the empirical adequacy of this account by invoking various background assumptions and hypotheses (e.g., that mental rotation occurs at a constant rate), in order to derive predictions about subjects' behavior from the axioms for a Shepard *IRS*. To the extent that these predictions were confirmed, we would have reason to accept the account.

Experiments like Shepard's have inspired much debate over the difference between analog representations (like pictures and visual images), on the one hand, and propositional representations (like sentences of English and LISP), on the other, and a number of criteria have been proposed for demarcating the two. According to one popular account, what distinguishes analog from propositional representations is that the structure of the former is similar to the structure of the things they represent. As we saw in the previous section, however, once we consider sentences at a theoretical level of logical form, it may turn out that they have some of the same structure as the states of affairs that they represent. Shepard himself stresses that when a propositional representation like a matrix is used to provide representations of a figure's orientation before, and after, a rotation, the intervening calculations using matrix algebra do not represent any intervening orientations of the figure at all. With mental rotations, by contrast, there is a structural similarity between the image and a rotating physical object at each of the intermediate stages of the mental transformation. This suggests the hypothesis that what is distinctive about analog representations is that each step in surrogative reasoning with them involves a structural representation of the phenomena that we are reasoning about.

Visual images can also be used in reasoning about situations that are not spatial. In a typical experiment, participants might be told that Tom is richer than Edna, and that Edna is poorer than Dan. Asked to decide who is richest, subjects often represent the people by objects standing in spatial relations (like being to the right of) that involve an order-isomorphism with the richer-than relation (e.g., Huttenlocher, [1968]), a strategy that again involves surrogative reasoning underwritten by a structural representation. But there may be many kinds of mental representations that are not structural representations. For example, much recent attention has been devoted to the *distributed representations* of connectionist accounts of cognitive activities (e.g., Hinton, McClelland and Rumelhart, [1986]). In such representations, specific units of a cognitive system do not stand for specific elements of the thing being represented. Instead, each element is represented by a pattern of activity distributed over many units (and each computing unit is involved in representing many different things). Distributed representations are typically given structural representations in vector spaces, but the representations themselves are not easily viewed as structural representations, since it is difficult to isolate aspects of them to serve as surrogates for the specific, individual constituents of the phenomena they depict. Recently, however, attempts have been made to endow distributed representations with a constituent structure, so that (perhaps at some very abstract level of analysis) a representation of the cat in the vat will contain a representation of the cat that is also present in a representation of the cat's biting the bat. It is too early to tell whether such accounts will succeed, but if they do, they may allow us to bring distributed representations into the fold of structural representation.[33]

7d. Possible-Worlds Semantics: Modal Logic as Measurement

We can show that ontological reductions and compositional semantic theories provide structural representations, and although additional work would be needed to show this for the mental representations discussed in the previous subsection, we found good reason to be optimistic about the prospects for doing so. In this subsection I shall examine a much more programmatic example, that of standard, possible-worlds semantics for alethic modal logic.

When doing metaphysics or the semantics of natural language, we often find possible-worlds semantics quite useful for reasoning about our modal thought and talk, and this raises an applications problem: Why does it work so well? Possible-worlds semantics is sometimes thought to provide a reduction of modality (to extensional logic and an ontology of merely possible entities), and our earlier conclusion that ontological reductions are often best construed as structural representations suggests that perhaps this semantics works as well as it does because it is a structural representation of some sort.

There are two, much more common, construals of possible-worlds semantics. On the one hand, we have modal realists, who willingly accept commitment to the existence of the merely possible individuals and worlds that the semantics appears to invoke. On the other, we have modal formalists, who hold that possible-worlds semantics is just a formal apparatus and, hence, free of ontological commitments. Modal realism has a ready explanation for the applicability of possible-worlds semantics: it is *literally true*; the merely possible entities that it seems to require really do exist. But it achieves this explanation at the cost of a dubious and epistemologically shaky ontology. Modal formalism, by contrast, avoids the ontological commitments of modal realism, but in a way that renders it unable to solve the applications problem; it cannot justify our use of modal logic, or explain why it applies to anything of interest.

The discussion so far suggests a synthesis of this thesis and antithesis that I shall call *modal representationalism*.[34] If successful, it would allow us to avoid the ontological excesses of modal realism, while still applying modal logic in good conscience. Modal representationalism should be particularly appealing to actualists, who hold that only actual things exist (or even could exist), so that what makes a claim like ⌜*a* could have been *P*⌝ true is not something in some merely possible world, but something in the actual world. Actualism doesn't require that modality be primitive and irreducible, but it fits nicely with the conclusion that it is (cf. Swoyer [1984], esp. fn. 11). I won't defend actualism or primitivism here, however, but shall simply try to show how modal representationalism would facilitate their development.[35]

Modal facts have a structure. For example, if it is a fact that *a* is *necessarily P*, then *a* is *actually P*, and if *a* is *actually P*, then *a* is *possibly P*. My hypothesis is that possible-worlds semantics – or, more precisely, the Kripke model structures it employs – provides a structural representation of such facts, and that this is what justifies its use in surrogative reasoning about them. Of course we can't rest content with a blithe slogan that something is *just a representation*. It must be shown in detail that the appropriate structural parallels exist, and one virtue of the present framework is that it makes clear what this would require.

Showing that length measurement is structural representation requires formal accounts of the medium of representation (the real numbers), of the phenomena that are represented (lengths), and a proof of a representation theorem. Analogous steps would be required to develop the sort of modal representationalism envisaged here. First, we would need an axiomatic account of the medium of representation, namely Kripke model structures, that treated them as multi-track *IRS*s (Section 2), i.e., as *IRS*s containing genuine properties and relations that are assigned extensions at different worlds. This would allow us to represent things, like the fact that a specific individual exemplifies a given property in every world, directly in a Kripke system, without any detour through language (cf. fn. 11). Since we would want to be able to deal with *de dicto* necessity, we would also need to include operations (of the sort mentioned at the end of Section 2) allowing us to build properties, relations, and propositions from the properties, relations, and individuals in the system. This would enable us to treat the extensions of propositions as truth values, so that truth could be defined directly in the relational system, again without a detour through language (cf. Bealer [1981]). Thus reconstructed, a Kripke model structure would be a multi-track *IRS* containing a set of tracks or 'worlds', one of which, *G*, would directly represent the actual world. But the things in these 'worlds' could be numbers or pure sets or formulas (as in canonical models of modal systems) or most anything else.

Second, we would need an axiomatic account of the modal structure of reality that does justice to the view that the world contains only individuals, properties, relations, and propositions, none of which are – or are parts of – merely possible worlds. As usual, we will need an *IRS* model of the phenomenon being represented, but where are we to find such a model of the modal structure of reality? We cannot simply adapt the standard Kripkean apparatus, since the goal is to explain why – rather than presuppose that – this machinery can be used to represent modal features of the world. Still, facts about modality can be painted onto Kripke model structures, and this tells us *something* about them; indeed, as we saw in Section 7a, philosophers often work backward, using a representation as an instrument to gain a better understanding of the things that it represents, and this would be a useful strategy here. Thus, we might tentatively select a particular modal logic, transform its standard semantic characterization into axioms for Kripke-style *IRS*s and, finally, use the features of this representation as a guide in devising an account of the modal structure of the actual world, perhaps in the following way.

I shall call the *IRS*s used to model the modal structure of reality or the actual world *modality systems*. A modality system contains individuals, primitive relations, and operations for building (compound) relations and propositions from these, but it would not contain alternative 'worlds' or alternative extension assignments. In addition to the sorts of relation-building operations mentioned in Section 2, it would be natural to add an operation, **Nec**, that maps relations (including propositions) to other relations (their necessitations); for example, it would map the property *being human* to the property *being necessarily human*. Finally, we would need axioms governing this new operation. On the current approach, the most interesting ones would be those specifying how **Nec** interacted with the extension assignment. Letting $\sim$ be the operation that maps relations to their negations, plausible candidates would include $^{\vee}\mathbf{Poss}(P) = {}^{\vee}{\sim}\mathbf{Nec}{\sim}(P)$, $^{\vee}\mathbf{Nec}(P) \subset {}^{\vee}(P)$, $^{\vee}(P) \subset {}^{\vee}\mathbf{Poss}(P)$, and (perhaps) $^{\vee}\mathbf{Nec} \subset (P)\,{}^{\vee}\mathbf{NecNec}(P)$.[36]

This treatment of **Nec** involves a slightly novel picture of relation-building operations. As we saw in Section 2, the extension of the conjunctive property *being P and Q* is *determined* by the extensions of *P* and of *Q*, but on the account envisioned here, the extension of the property *being necessarily P* – **Nec**(*P*) – is not completely determined by the extension of *P*, either in ***M***, or anywhere else (there *is* nowhere else). This suggests that the axioms governing some relation-building operations should place structural constraints on their extensions, rather than completely determining them. For example, although $^{\vee}$**Nec**(*P*) ⊂ $^{\vee}$(*P*) doesn't completely determine the extension of **Nec**(*P*), it *does* require that it be a subset of the extension of *P*. What, over and above this, pins down the extension of **Nec**(*P*)? As long as the constraints provided by the axioms for ***M*** didn't settle the matter (as they would have to, for example, if *P* has the form *Q or not Q*), nothing, at least nothing in the logic, would determine the extension of **Nec**(*P*). But this is just what it means to say that modality is *primitive*.[37]

The final step in the development of modal representationalism would be the proof of a representation theorem, ensuring that each modality system ***M*** could be mapped to some Kripke system ***K*** in a way that preserved modal structure, that is, in such a way that a proposition would be true in ***M*** just in case its surrogate was true in the actual 'world', *G*, of ***K***. We can think of this mapping proceeding in stages, beginning with just ***M*** and *G*. Since we can tailor *G* to fit ***M***, it will always be possible to concoct a mapping *c* that carries the individuals and the primitive relations of ***M*** to surrogates in *G* in a way that respects these relations (so that, for example, an individual in ***M*** will exemplify the primitive property *P* just in case its surrogate in *G* exemplifies the surrogate of *P*). It is also possible to require that *c* be one-one and onto, which means that it will respect all of ***M***'s (non-modal) relations, including those with internal structure (so that, for example, an individual in ***M*** will exemplify the property *bearing L to something* just in case its surrogate in *G* exemplifies the surrogate of this compound property).

The question is whether we can go on to construct a Kripke system of 'worlds' around *G* in such a way that a modal proposition (one constructed using **Nec**) will be true in ***M*** just in case its surrogate is true in *G*. One way to approach this would be to adapt standard rules for semantic tableaux in modal logic, so that, for example, if we have *a* ∈ $^{\vee}$~**Nec**(*P*) in ***M***, we would put *c*(*a*) ∈ $^{\vee}$~**Nec**(*c*(*P*)) in *G*, and then would add a new 'world', accessible to *G*, in which *c*(*a*) ∈ $^{\vee}$~*c*(*P*). Although the basic ideas here are reasonably straightforward, it requires some delicacy to give a plausible account of the detailed interactions of **Nec** and such things as operations that inject quantificational structure into relations. At each stage, the key test would be whether the axioms for modality systems constrained facts about modality in a way that ensured that such an *IRS* could always be mapped to some Kripke system. A representation theorem showing that they do would then explain why things work *as if* the possible-worlds account were true, and so would justify the use of Kripke semantics in surrogative reasoning about modal matters, without requiring the existence of any non-actual things.[38]

One virtue of modal representationalism is that it would allow us to apply lessons learned in our examination of structural representation to the metaphysics of modality. Consider, for example, the problem of representational artifacts. A Kripke *IRS* could contain objects in some of its 'worlds' that were not surrogates of any actual individuals. As a simple illustration, imagine a modality system $\boldsymbol{M}$ that includes just John Kennedy, the two properties, *being a senator* and *being an astronaut*, and the two-place relation, *being the father of*. We might represent $\boldsymbol{M}$ by a Kripke system $\boldsymbol{K}$ containing the three 'worlds', G, H_1, and H_2, where G contains the individual x, and H_1 and H_2 each contain x and a second individual, y. We then introduce a mapping, c, that pairs Kennedy with his surrogate, x, and that pairs the properties and relations in $\boldsymbol{M}$ with surrogates in $\boldsymbol{K}$. In H_1, we let $\langle c(\text{Kennedy}), y\rangle$ be in the extension of the surrogate of the *father of* relation and y be in the extension of the surrogate of the property of *being a senator*, while in H_2 we let y be in the extension of the surrogate of the property of *being an astronaut*. These machinations will make the surrogate of $\boldsymbol{M}$'s proposition that Kennedy could have had a son who was a senator, but who might have been an astronaut instead, true in G. In this example, x is a surrogate for Kennedy, but y is not a surrogate for anything at all – it's just a (systemic) artifact of a structural representation. Like many other artifacts, y plays a computationally important role, greatly facilitating surrogative reasoning about modality, but it has no direct representational significance. In Section 5, I conjectured that certain types of realism were encouraged by mistaking artifacts of representations for features with representational significance, and it is natural to conjecture here that modal realism is similarly abetted by the misapprehension that objects like y are directly representational.[39]

These remarks on modal representationalism are much more speculative than the discussions of earlier examples of structural representation, but if such an approach worked for the alethic modalities, it would be natural to consider extending it to other intensional logics, like tense logic or the logic of belief. After all, it is difficult to regard beliefs, or their objects, as literally involving sets of possible worlds, but it is plausible to suppose that some structural features of facts involving beliefs can be represented by such sets. None of this, of course, is meant to disparage possible-worlds semantics, any more than representational theories of measurement are meant to disparage the real numbers. The existence of qualitative axioms for lengths does not render numbers obsolete for reasoning about length, nor would the existence of axioms for modality systems render Kripke semantics obsolete for reasoning about modality. As with many other structural representations, these two representational systems are familiar, powerful, and computationally tractable, and so they would often be easier to work with than the reality that they represent. But they are still representations, rather than that reality.

My goal here has been to explain what structural representation and surrogative reasoning are, to show why they are important, and to develop an account of them. If I have been on the right track, the account provides a framework in which a number of philosophically interesting phenomena can be fruitfully studied and, in some cases, in which familiar problems involving them can be solved.[40]

NOTES

[1] Representations of non-existent states of affairs sound more mysterious than they are. They will be discussed briefly below, but the basic idea is that many representations can represent the way things *would* have been, had things been slightly different from the way they actually are. Such representations are both common and important. Rather than embarking on a course of action and discovering its consequences the hard way, it is often prudent to do an everyday *Gedankenexperiment* (as in a chess game), a physical simulation (as in a wind tunnel), or a computer simulation (as with models of the greenhouse effect), in order to see what would happen under various *possible* conditions. These vicarious explorations of alternative possibilities play an essential role in planning and decision making.

[2] I shall sketch a defense of my claim about the applicability of mathematical theories to reality below; more detailed defenses may be found in Krantz, et al. (1971), Swoyer (1987), and various presentations of structuralism in the philosophy of mathematics (e.g., Resnik [1981], Shapiro [1983]). I shall talk about the structure of reality as though the notion were unproblematic. Some philosophers contend that our thought or language actually shape this structure, but even if there is some sense in which this is so, we can still talk about the structure of things *given* our concepts and interests. And we can retain fallibilism while talking about the structure that things seem to have in light of our best theories about the world. Although my account will apply to several types of mental

representation, these involve rather special problems, and so I shall set them aside until Section 7c.

3 A more precise account of this is given in Section 4. The picture of measurement at work here is the *representational theory of measurement* developed by Helmholtz, Hölder, Campbell, Stevens and, more recently, Suppes and his collaborators (see e.g., Krantz, et al. [1971], which also contains numerous references to earlier work in the tradition, and Narens [1985]).

4 If this knowledge involves generalizations analogous to scientific laws of coexistence, the conclusion will be one about additional features of the original situation that don't involve change (as in reasoning about distances on a map). If the knowledge involves something more like laws of succession, the conclusion will tell us something about the original phenomenon at some other time (as with calculations to discover the date of the next lunar eclipse). A number of examples of surrogative reasoning will be discussed below; two simple, but detailed, examples will be found in fn. 16. In some cases, e.g., the use of a scale model in a wind tunnel, it may be most natural to say that we reason directly *about* a representation, and thereby reason *indirectly* about that which it represents. In other cases, e.g., the use of numbers in measurement, it may be more natural to say that we use a representation to reason more-or-less directly about what it represents. Structural representations may also mediate inductive inferences, but here I shall concentrate on surrogative reasoning that is deductively valid.

5 Structural representation has a distinguished philosophical history. Leibniz called it *expression*, telling us that one thing "expresses something in which there are relations that correspond to the relations of the thing expressed", so that "we can pass from a consideration of the relations in the expression to knowledge of the corresponding properties of the thing expressed" ([1970], p. 207). Apart from his claim that each monad expresses the entire universe, many of Leibniz's examples are quite similar to the examples of structural representation discussed below: a map of a region expresses the region, a model of a machine expresses the machine, the perspectival projection of a figure on a plane expresses the original figure, speech expresses thought. The view that something very like structural representation underlies such diverse phenomena as the semantics of natural languages and the propositional attitudes forms the core of Wittgenstein's picture theory in the *Tractatus* (1921). If the examples in Section 7b and 7c are correct, Wittgenstein was much closer to the truth than is commonly supposed, and it is natural to conjecture that the glaring defects in his account derive mainly from its accompanying doctrines of extensionalism, logical atomism, the absolute simplicity of objects and (arguably) nominalism, none of which have any part in the present story.

6 I won't assume the existence of any more properties and relations than are needed to accommodate the cases of representation I discuss. I shall reserve the terms 'property' and 'relation' for *genuine* properties and relations, and shall call their extensional substitutes 'sets' and 'relations-in-extension'. In more formal contexts, it is often convenient to treat properties as (one-place) relations, and I shall sometimes follow this practice. I shall also treat properties and relations as universals, rather then tropes or quality instances.

7 Of course an *IRS can* be regarded as a kind of *extensional* relational system. One way to do so is to collapse its domain of individuals and its domains of relations into a single, over-arching domain, add a number of sets to the system (one for each type of entity in original *IRS*), and treat $^{\vee}$ as a (partial) function on the single, new domain. But the *philosophical* uses we shall have for *IRS*s make a Gestalt switch fruitful, reversing this

point of view to exhibit extensional relational systems as a special case of intensional systems. This can be achieved by thinking of extensional systems as containing a tacit or suppressed extension function that is one-to-one (it could even be the identity function). This also facilitates comparisons between the two sorts of relational systems. But from now on, when I speak of a relational system, I shall mean an *IRS*.

[8] The general version of (***PR***) tells us that for every n-place, l-level relation $\underline{\mathrm{R}}$ in A and n-tuple of items, $\langle \tau_1, \ldots, \tau_n \rangle$ of level l-1 in A, $\langle \tau_1, \ldots, \tau_n \rangle \in {}^{\vee}\underline{\mathrm{R}}$ if and only if $\langle c(\tau_1), \ldots, c(\tau_n) \rangle \in {}^{\vee}c(\underline{\mathrm{R}})$.

[9] When we treat a real-life system as an *IRS*, the axioms for it could be construed as definitions of the relations it contains, but it is often better to regard them as empirical claims about those relations (cf. Swoyer [1987], 260ff.). In some cases an *IRS* model of a situation or phenomenon not only disregards some of its relations, but injects a hefty dose of idealization as well. We frequently treat actual things as point masses or ideal speaker-hearers or objects that have perfectly definite lengths, even though we know that there really aren't any such things. This often enables us to provide reasonably tractable structural representations of actual systems in well-understood mathematical systems. In such cases, we can still think of the *IRS* model as a faithful depiction of something actual, though now it is the scientist's idealized version of a real-life system, rather than the system itself, that the representation depicts.

[10] These operations allow *intensional definitions* in which we define new relations directly in an *IRS*, without any detour through language. At the turn of the century, Russell proposed a view of this sort, urging that the class of properties and relations is closed under negation, conjunction, and relative products ([1903], Ch. II, Sections 27–30). Axioms assuring the proper working of the various operations (and families thereof) are pretty much what one would expect; detailed accounts may be found in Bealer (1981), Zalta (1983) and Menzel (1986).

[11] Formally, a multi-track *IRS* is an ordered set $A = \langle I, \mathfrak{R}, \mathbf{Tm}, <, \mathbf{Tr}, {}^{\vee} \rangle$ where $\mathbf{Tm}$ is a nonempty set (whose members are, or stand for, times), $<$ is a binary ordering relation on $\mathbf{Tm}$, $\mathbf{Tr}$ is a nonempty set of tracks, and ${}^{\vee}$ is a three-place function that assigns an extension of the expected sort to each relation at every track at every time. A multi-track *IRS* can do much of the work of a phase space, with the space's *possible states* being correlated with the set of all atomic facts that obtain in any particular track at any particular moment. Multi-track *IRS*s are inspired by the model structures common in intensional logics, but differ from them in containing genuine properties and relations, rather than their extensional stand-ins. This difference matters; on the standard approach to tense logic, for example, the reliance on extensional stand-ins means that a model structure has no intrinsic features capable of representing change. There is no sense in which something can have a given feature at one time and lack it another, since within the structure itself we cannot identify the same feature or property at different times. The structure relevant to change is only adventitiously injected into the model structure when we interpret a language over it. Yet many representations have enough intrinsic structure to represent change (or alternative possibilities) directly, without a side trip through language, and multi-track *IRS*s allow us to do justice to this.

[12] The function s does not preserve type, since determinate lengths are properties and their numerical surrogates are individuals (this situation can also be reversed, with properties representing individuals, as in the color coding of floors in a building). The account in Section 6 will accommodate such trans-type representation in a general way. Here I shall finesse the problem by treating determinate lengths and real numbers as the

entities of the lowest orders in their respective *IRS*s. Although an *EPS* contains no domain of individuals, this is easily added, and the indirect ascription of scale values to objects in it can be explained in terms of the values assigned to the lengths they exemplify. The treatment of determinate magnitudes as properties goes back at least to Aristotle's discussion of the category of quantity in ch. 6 of the *Categories* (e.g., $I^{b}26$ and $6^{a}19$–20), where such things as *being two cubits long* are treated as properties. It is explicitly defended by Russell in (1903), chs. 19–21. In Swoyer (1987) I argued that there are good philosophical reasons to view measurement as the assignment of numbers to properties, rather than to individuals (cf. Mundy [1987b]), but my reason for treating it this way here is to illustrate various points about structural representation.

[13] For readability I shall use infix notation, writing '$P_1 > P_2$' for '$\langle P_1, P_2\rangle \in {}^{\vee}>$', and shall treat '$P_1 \bigcirc P_2$' as a singular term denoting the unique property that is the 'sum' of P_1 and P_2 (its existence is guaranteed by the axioms for an *EPS*). When a relation R is a function, Section 2's schema (***PR***) for the preservation of relations, $\langle i_1, \ldots, i_n\rangle \in {}^{\vee}R$ iff $\langle c(i_1), \ldots, c(i_n)\rangle \in {}^{\vee}c(R)$, is often replaced by the schema for the preservation of functions, $c(F(e_1, \ldots, e_{n-1})) = c(F)(c(e_1), \ldots, c(e_{n-1}))$, so I shall use (***B***) in place of the longer '$\langle P_1, P_2, P_3\rangle \in {}^{\vee} \bigcirc$ iff $\langle c(P_1), c(P_2), c(P)_3 > \in {}^{\vee}+$'.

[14] This means that for any two isomorphic embeddings s and s', there is some positive real number α such that $s' = \alpha s$ (in converting from meters to feet, $\alpha = 1.0936$). In contrast to the ratio scales involved in extensive measurement, interval scales (like the Celsius and Fahrenheit scales for temperature) often involve a representation in the numerical relational system $\langle R, \underline{R}^4\rangle$, where R is the set of real numbers, and $\underline{R}^4xyzw$ just in case $x - y \geqslant z - w$. The uniqueness theorem here shows that interval scales are only unique up to a positive linear transformation (one of the form $s' = \alpha s + \beta$, where $\alpha > 0$; in converting from Celsius to Fahrenheit, $\alpha = 9/5$ and $\beta = 32$). Ordinal scales (like the Mohs scale for hardness of minerals), which simply preserve order, are only unique up to strictly increasing monotonic transformations. I shall discuss the philosophical significance of these matters in the next section.

[15] The monotonicity axiom tells us that for all P_1, P_2, and P_3 in E, $P_1 > P_2$ just in case $(P_1 \bigcirc P_3) > (P_2 \bigcirc P_3)$ just in case $(P_3 \bigcirc P_1) > (P_3 \bigcirc P_2)$. And the Archimedean axiom tells us that for all P_1, P_2, P_3, and P_4 in E, there is some natural number n such that $P_1 > P_2$ only if $nP_1 \bigcirc P_3 > nP_2 \bigcirc P_4$. Part of the intuitive force of this is that for any property we pick in E, we can get a 'larger' property (one bearing the $>$ relation to it) by combining any other property with itself a finite number of times (with the metaphor of combination being spelled out by the inductive definition of the *property* nP: $1P = P$ and $(n + 1)P = nP \bigcirc P$ (cf. Swoyer [1987], p. 271).

[16] This procedure is an excellent source of simple, yet detailed, examples of surrogative reasoning. (***A***) assures us that a representation exists only if there is a one-one function s such that $P_1 > P_2$ just in case $s(P_1) > s(P_2)$, and this provides a bridge between facts in E and their surrogates in $\boldsymbol{R}^+$. Thus, assume that $P_1 > P_2$. We can translate this into the numerical information that $s(P_1) > s(P_2)$, then employ the mathematical theory of the positive, additive reals to conclude that for any number P_3 that is the surrogate of any property in E, $s(P_1) + s(P_3) > s(P_2) + s(P_3)$. Finally (***A***) and (***B***) together allow us to transfer this information back to the original system $\boldsymbol{E}$ to conclude that $P_1 \bigcirc P_3 > P_2 \bigcirc P_3$. Again, suppose that $P_1 > P_2$ and $P_2 > P_3$. (***A***) tells us that $s(P_1) > s(P_2)$ and $s(P_2) > s(P_3)$, which by simple mathematical reasoning yields $s(P_1) > s(P_3)$, and (***A***) again underwrites the return trip to $\boldsymbol{E}$ and the conclusion that $P_1 > P_3$.

[17] By way of example, the automorphism group of $\boldsymbol{R}^+$ is the set of similarity transformations; by contrast, the only automorphism of the positive, additive *integers* is the identity mapping. Other candidates for the permissibility group include the endomorphisms of a relational system and its isomorphic embeddings in itself. For brevity, my discussion of meaningfulness oversimplifies several points about a complex topic; e.g., although in many cases permissibility can be explained in terms of function composition, as we have done here, other cases require a more abstract characterization (Roberts and Franke [1976]). Matters will become even more complex in the next section, when we allow other sorts of mappings besides isomorphic embeddings to underwrite structural representations, but my goal here is just to provide enough feel for the issues so readers can follow the discussion of artifacts below. For an early treatment of the issues, see von Neumann and Morgenstern (1944), pp. 20–25. A good discussion of current accounts of meaningfulness and an attempt to justify the common (but rarely defended) practice of imposing uniqueness condition on representations is Mundy (1986).

[18] In Section 7d I shall argue that modal realism is encouraged by this sort of mistake. Of course this strategy of arguing against specific versions of realism opens the door to more general arguments that all sorts of things various realists think exist are mere artifacts of some mode of representation. I am not endorsing such an approach here, but its possibility underscores the need for a philosophically principled way to distinguish the objects and relations that a theory is genuinely committed to from those it is not. The view that genuine properties and relations play a causal role in the world seems to me one promising place to begin (cf. Swoyer [1982b]).

[19] Quine originally spoke of all dispositions to verbal behavior, but (in 1969a) he accepted Harman's proposal (1969) to speak instead of dispositions to assent and dissent. All this is clearer when the home language and the target language are the same, so that any (non-homophonic) member of $\boldsymbol{G}$ will carry all of a speaker's *own* sentences to other sentences in a way that preserves patterns of assent and dissent. The point of Quine's remark about equivalent sentences is that constraints on translations are weak enough to allow counterintuitive pairings.

[20] From the present perspective, Quine's and Grünbaum's positions involve the sort of defense of antirealism – with respect to meanings and metrics, respectively – discussed in fn. 18 and the accompanying text. Of course, the fact that the various examples of conventionalism mentioned above have a common form does not mean that they are equally plausible.

[21] Points (i) and (vii) were noted in Mundy (1986) Section 2.

[22] For manageability, I shall require that c map items of a given type in $\boldsymbol{A}$ to items of a single type in $\boldsymbol{B}$, that items of different types in $\boldsymbol{A}$ be mapped to items of different types in $\boldsymbol{B}$, that c pair items in adjacent domains with items in adjacent domains (thus preserving order of levels), and that c preserve rank in all cases except those in which it carries relations to (or from) a domain of individuals.

[23] The sets Δ and Ψ may contain defined relations. They will be closed under some operations but not others; for example, Δ is closed under analogues of conjunction and existential quantification, Ψ under analogues of negation and universal quantification. To show that there is a Δ/Ψ-morphism from any *IRS* satisfying a given set of axioms to a given system $\boldsymbol{B}$ is to prove a sort of representation theorem. This approach counts a mapping as a structural representation even if it counter-preserves only a few of the relations in $\boldsymbol{A}$, but it would be arbitrary to require that some particular percentage of

relations be counter-preserved. Besides, it is possible for just a few features of a representation to have representational significance, as in elaborate ciphers and codes, which may be peppered with false leads.

[24] A similar picture lies at the heart of several recent accounts of knowledge, according to which a necessary condition for x's knowing that p is that if p hadn't been true, x wouldn't have believed that p (e.g., Dretske [1971], Nozick [1981]). This is really just a counterfactual version of the requirement that a structural representation counter-preserve relations, so that if the object o hadn't been in extension of the property P, it wouldn't have been represented as being in the extension of P. The requirement that relations be preserved is reminiscent of Nozick's fourth condition for knowledge, namely, that if p had been the case (even if other things had been a bit different), x would still have believed that p.

[25] Benacerraf (1965) made this problem prominent; analogous difficulties arise for many other sorts of reductions. One might attempt to avoid the problem by adopting a theory of relative identity, but I believe that Perry's criticisms of this approach (1970), while not definitive, are completely sound in spirit. In a more radical vein, Goodman (1978) has argued that even though alternative reductions are incompatible, both might still be true; I have argued against this sort of approach (in 1988). Philosophers often speak of the reduction of number theory to set theory, but I shall speak of the reduction of numbers to sets, since the putative ontological significance of such reductions is to tell us what numbers themselves are.

[26] The present framework provides a natural one for developing such a structuralist account of mathematics (cf. Swoyer [1987], Section 4). According to structuralism, any countably infinite (recursive) set can be arranged to form an ω-sequence that can play the *role* of the natural numbers. It is the structure common to all such sequences, rather than the particular objects which any happens to contain, that is important for arithmetic. Since relations in *IRS*s don't come with their extensions built into them, such relational systems allow us to separate relations from their extensions. Hence, we could focus on a particular sort of intensional relational system – what we might call a *Natural-number IRS* – that can be exhibited in various concrete realizations. This system would have an empty domain of individuals, but would contain properties and relations like *being less than*, *being a successor of* and *being the first member*. We could then employ operations like those discussed in Section 2 to define further number-theoretic properties and relations. Finally, a concrete realization would be obtained by adding a domain of individuals and assigning them as extensions to the properties and relations in the structure.

[27] A mapping *from* the representational system (syntax) *to* what it represents (semantics) runs in the opposite direction from the structure-preserving mappings in our previous examples of structural representation. We saw in Section 6 that this is to be expected in linguistic representation, since different words can be used to represent the same thing, and this prompted us to extend our account of structural representation to accommodate mappings from the representation to what it represents.

[28] Montague's relational systems are not *IRS*s, but they are easily transformed into them. In order to assure that his systems were algebras, i.e., that all of their relations were operations, he included the set of all syntactic strings in his syntactic system and assigned meanings to each, but we needn't worry about such subtleties here.

[29] I have given very simple examples in order to illustrate the nature of compositionality; these can be handled by extensional relational systems, and I have done so here to minimize complexities. However, languages containing intensional idioms are not easily

handled in an extensional framework. For examples of treatments of some of the above points in terms of *IRS*s, see Mundy (1987b) and Swoyer (1987). The use of *IRS*s to deal with more complicated languages may be found in Bealer (1981), Zalta (1983), and Menzel (1986).

[30] Of course there is more than one person named 'George Bush', which suggests that many schemes of reference will not be functions after all. Both the Davidsonian and the Montagovian must somehow come to terms with this fact. One solution in the case of proper names is developed in Burge (1973). Davidson himself is primarily interested in the interpretation of the *speakers* of a language, but I think that most of the points made above apply to this as well; further discussion of this may be found in Swoyer (1987), Section 4.

[31] Hence, surrogative reasoning may often involve *multiple levels of surrogates*. We might use a ruler as a surrogate for an object, numbers as surrogates for the marks on the ruler, and some sort of mental representations as surrogates for the numbers.

[32] Like Johnson-Laird, I have assumed that mental models, and indeed structural representations generally, are consistent. Actual representations and models are not always logically impeccable, however, and ways of injecting inconsistency into representations without contaminating them wholesale would be worth exploring (as would the study of the possibility that the extensions of some relations in a structural representation are fuzzy).

[33] A final application of the current framework to mental representation is this. Just as claims about lengths of objects are sometimes analyzed as involving a *relation* between an object and a number, claims about belief are sometimes analyzed as involving a relation between a believer and a proposition (or some other proposition-like entity). Mundy (1987b) and Swoyer (1987) have argued that this conception of measurement is mistaken; the length of an object is a one-place, qualitative property, though the *structure* of the family of such properties allows us to use numbers to classify them. Recently, several philosophers have suggested that relational accounts of belief involve a similar mistake (e.g., Field [1981], who credits the idea to David Lewis). However such proposals have been left as suggestive analogies, rather than being developed to the point where we could usefully discuss such things as their representation and uniqueness problems. In Swoyer (1987), I argue that intensional relational systems provide a natural device for developing such views. Among other things, they allow us to treat beliefs as monadic properties with an internal structure that is quite similar to that of propositions, and this might be used to explain why we can so successfully use propositions to classify them.

[34] An account of Kripke semantics as representational has been developed independently by Menzel (1990). Although our motivations are similar, our approaches have the representing relation running in opposite directions, and they differ in a number of other respects as well.

[35] Some versions of actualism attempt to find respectable stand-ins (like individual essences or maximally consistent states of affairs) for merely possible individuals and merely possible worlds. I think that many of the objections to their full-blooded realist counterparts also tell against such stand-ins, and so I shall adopt a more austere actualism that avoids them.

[36] The first of these introduces the notion of possibility in terms of **Nec**'s dual, **Poss**, the second and third capture the ideas that if something is necessarily P then it is actually P, and if actually P, then possibly P, and the fourth is a version of the characteristic axiom for S4. Additional operations, including analogues of the conditional and quantifi-

ers, would be needed to state more complicated axioms. If we identify the truth values *true* and *false* with the set-theoretic surrogates, {∅} and ∅, these axioms work for propositions as well as relations.

[37] The situation here is not as different from Kripke semantics as it might seem. Suppose that '*H*' is a primitive predicate and that '□*Hk*' is true in a traditional, Kripke model. This means that '*Hk*' happens to be true in every world in the model, but this is just as much a primitive and inexplicable fact about this particular model as the fact that Kennedy is necessarily human is about ***M***. Since modal representationalism is nonreductive, it would not offer *non*-modal truth conditions for modal propositions, although it would allow homophonic theories of truth which contained a necessity operator in the metalanguage.

[38] A variant on this approach would attempt to show that for each modality system, an extensional Kripke model structure could be constructed in such a way that a sentence of a standard modal language would be true in ***M*** (when interpreted over it in the obvious way) just in case it was true in the Kripke structure (when interpreted over it in the usual way).

[39] The Kennedy example is McMichael's (1983). Once we regard Kripke structures as representations, there is no reason why a given individual need exist in more than one 'world'; an actual object could as well be represented by different surrogates in different worlds, all of whom were related by some 'world-line' relation. This would be of particular interest in cases where world lines split or merge from world to world, as in the logic of belief. It would also let us adapt features of Lewis's counterpart theory, without having to construe it in a realistic way.

[40] I am indebted to David Armstrong, Neera Badhwar, Hugh Benson, John Biro, Monte Cook, Rick Kirkham, Adam Morton, Brent Mundy, Scott Shalkowski, the referees for *Synthese*, and to the philosophers at the University of Iowa and Texas A&M University, where I read earlier versions of this paper. I am also grateful to the University of Oklahoma Office of Research Administration for a summer grant for this project.

REFERENCES

Aristotle: 1963, *Aristotle's Categories and de Interpretatione*, trans. J. L. Ackrill, Clarendon Aristotle Series, Oxford.

Adams, E. W.: 1965, 'Elements of a Theory of Inexact Measurement', *Philosophy of Science* **32**, 205–28.

Armstrong, David: 1973, *Belief, Truth and Knowledge*, Cambridge University Press, Cambridge.

Armstrong, David: 1978, *Universals and Scientific Realism, Vol. II. A Theory of Universals*, Cambridge University Press, Cambridge.

Bealer, George: 1981, *Quality and Concept*, Clarendon Press, Oxford.

Benacerraf, Paul: 1965, 'What Numbers Could Not Be', *Philosophical Review* **74**, 47–73.

Burge, Tyler: 1973, 'Reference and Proper Names', *Journal of Philosophy* **70**, 425–39.

Craik, Kenneth: 1943, *The Nature of Explanation*, Cambridge University Press, Cambridge.

Davidson, Donald: 1984, *Inquiries into Truth and Interpretation*, Clarendon Press, Oxford.

Dretske, Fred: 1971, 'Conclusive Reasons', *Australasian Journal of Philosophy* **49**, 1–22.

Field, Hartry: 1981, 'Mental Representation', Reprinted with a postscript in N. Block (ed.), *Readings in the Philosophy of Psychology: Vol I*, Harvard University Press, Cambridge, MA, pp. 78–114.

Fodor, Jerry: 1975, *The Language of Thought*, Thomas Y. Crowell, New York.

Goodman, Nelson: 1978, *Ways of World Making*, Hackett, Indianopolis.

Grünbaum, Adolf: 1973, *Philosophical Problems of Space and Time*, 2nd ed., D. Reidel, Dordrecht.

Harman, Gilbert: 1969, 'An Introduction to 'Translation and Meaning': Chapter Two of *Word and Object*', in Donald Davidson and Jaakko Hintikka (eds.), *Words and Objections; Essays on the Work of W. V. Quine*, D. Reidel, Dordrecht, pp. 14–26.

Hertz, Heinrich: 1956, *The Principles of Mechanics*, English translation, Dover, New York.

Hinton, G. E., McClelland, J. L., and Rumelhart, D. E.: 1986, 'Distributed Representations', in D. Rumelhart and J. McClelland (eds.), *Parallel Distributed Processing, Volume I: Foundations*, MIT Press, Cambridge, MA, pp. 77–109.

Huttenlocher, J.: 1968, 'Constructing Spatial Images: A Strategy in Reasoning', *Psychological Review* **75**, 550–60.

Johnson-Laird, Philip: 1983, *Mental Models*, Harvard University Press, Cambridge, MA.

Krantz, D., Luce, R., Suppes, P. and Tversky, A.: 1971, *Foundations of Measurement, Vol. I*, Academic Press, New York.

Leibniz, Gottfried Wilhelm: 1970, 'What is an Idea?', translated in Leroy E. Loemker (ed.), *Gottfried Wilhelm Leibniz: Philosophical Papers and Letters*, D. Reidel, Dordrecht, pp. 207–08.

McMichael, A.: 1983, 'A Problem for Actualism about Possible Worlds', *Philosophical Review* **92**, 49–66.

Menzel, Christopher: 1986, 'A Complete Type-Free 'Second-order' Logic and Its Philosophical Foundations', Report No. CSLI-86-40, Center for the Study of Language and Information, Stanford University.

Menzel, Christopher: 1990, 'Actualism, Ontological Commitment, and Possible World Semantics', *Synthese* **85**, 355–89.

Metzler, J. and Shepard, R. N.: 1982, 'Transformational Studies of the Internal Representation of Three-Dimensional Objects', reprinted in R. N. Shepard and L. A. Cooper, *Mental Images and their Transformations*, MIT Press, Cambridge, MA, pp. 25–71.

Montague, Richard: 1974, 'Universal Grammar', in Richmond Thomason (ed.), *Formal Philosophy: Selected Papers of Richard Montague*, Yale University Press, New Haven, pp. 222–46.

Mundy, Brent: 1986, 'On the General Theory of Meaningful Representation', *Synthese* **67**, 391–437.

Mundy, Brent: 1987a, 'Faithful Representation, Physical Extensive Measurement Theory and Archimedean Axioms', *Synthese* **70**, 373–400.

Mundy, Brent: 1987b, 'The Metaphysics of Quantity', *Philosophical Studies* **51**, 29–54.

Narens, Louis: 1985, *Abstract Measurement Theory*, MIT Press, Cambridge, MA.

Nozick, Robert: 1981, *Philosophical Explanations*, Belknap Press, Cambridge, MA.

Perry, John: 1970, 'The Same F', *Philosophical Review* **70**, 181–200.

Quine, W. V. O.: 1960, *Word and Object*, MIT Press, Cambridge, MA.
Quine, W. V. O.: 1969a, *Ontological Relativity and Other Essays*, Columbia University Press, New York.
Quine, W. V. O.: 1969b, 'Replies', in Donald Davidson and Jaakko Hintikka (eds.), *Words and Objections; Essays on the Work of W. V. Quine*, D. Reidel, Dordrecht, pp. 295–97.
Resnik, Michael: 1981, 'Mathematics as a Science of Patterns: Ontology and Reference', *Noûs* **15**, 529–50.
Roberts, F. S. and Franke, C. H.: 1976, 'On the Theory of Uniqueness in Measurement', *Journal of Mathematical Psychology* **14**, 211–18.
Russell, Bertrand: 1903, *Principles of Mathematics*, Cambridge University Press, Cambridge.
Shapiro, Stewart: 1983, 'Mathematics and Reality', *Philosophy of Science* **50**, 523–48.
Shepard, Roger and Chipman, Susan: 1970, 'Second-Order Isomorphism of Internal Representations: Shapes of States', *Cognitive Psychology* **1**, 1–17.
Suppes, Patrick: 1973, 'Some Open Problems in the Philosophy of Space and Time', in P. Suppes (ed.), *Space, Time, and Geometry*, D. Reidel, Dordrecht, pp. 383–402.
Swoyer, Chris: 1982a, 'Belief and Predication', *Noûs* **15**, 197–220.
Swoyer, Chris: 1982b, 'The Nature of Natural Law', *Australasian Journal of Philosophy* **60**, 203–23.
Swoyer, Chris: 1983, 'Realism and Explanation', *Philosophical Inquiry* **5**, 14–28 .
Swoyer, Chris: 1984, 'Causation and Identity', in P. A. French, T. E. Uehling, Jr., and H. K. Wettstein (eds.), *Midwest Studies in Philosophy: Vol IX: Causation and Causal Theories*, University of Minnesota Press, Morris, MN, pp. 593–622.
Swoyer, Chris: 1987, 'The Metaphysics of Measurement', in John Forge (ed.), *Measurement, Realism and Objectivity*, D. Reidel, Dordrecht, pp. 235–90.
Swoyer, Chris: 1988, 'Relativism and Representation', *Philosophy and Phenomenological Research* **49**, 151–55.
Von Neumann, J. and Morgenstern, O.: 1944, *Theory of Games and Economic Behavior*, Princeton University Press, Princeton.
Wallace, John: 1977, 'Only in the Context of a Sentence do Words Have any Meaning', in P. A. French, T. E. Uehling, Jr., and H. K. Wettstein (eds.), *Midwest Studies in Philosophy: Vol II: Studies in the Philosophy of Language*, University of Minnesota Press, Morris, MN, pp. 305–25.
Wittgenstein, Ludwig: 1961, *Tractatus Logico-Philosophicus*, Routledge & Kegan Paul, London.
Zalta, Edward: 1983, *Abstract Objects: An Introduction to Axiomatic Metaphysics*, D. Reidel, Dordrecht.

Dept. of Philosophy
University of Oklahoma
455 West Lindsey Street
Norman, OK 73019-0535
U.S.A.