

Resolving Arguments by Different Conceptual Traditions of Realization (final version for Phil. Studies)

There is currently a significant amount of interest in understanding and developing theories of realization. Naturally arguments have arisen about the adequacy of some theories over others. Many of these arguments have a point. But some can be resolved by seeing that the theories of realization in question are not genuine competitors because they fall under different conceptual traditions with different but compatible goals. I will first describe three different conceptual traditions of realization that are implicated by the arguments under discussion. I will then examine the arguments, from an older complaint by Norman Malcolm against a familiar functional theory to a recent argument by Thomas Polger against an assortment of theories that traffic in inherited causal powers, showing how they can be resolved by situating the theories under their respective conceptual traditions.

1. Conceptual Traditions of Realization

I begin with three conceptual traditions that are relevant to the present discussion (see Endicott, 2005). There is a representational tradition according to which realization is a *semantic or intentional relation*. Generally speaking, some object realizes (or is the realization of) a term, concept, or other form of representation because the object meets the condition for satisfying that representation. Thus Alfred Tarski said that a set of objects is the “realization” of a language when the objects satisfy the predicates of that language (1956, pp.416-17; see also Lewis, 1980, pp.210-211). The representational tradition’s idea of realization *qua* semantic or intentional satisfaction has been commonly used within areas of logic and philosophy, and it also captures the folk notion expressed in dictionary definitions whereby an idea is “made real” by an object. An idea represents an object, whose existence thus counts as the idea’s realization.¹

¹ A reviewer asked whether the folk notion is intentional as I claim rather than a metaphysical notion whereby a person (such as an artist) makes some object real by using other objects. My view, more cautiously stated, is that the common folk notion is at least *in part* an intentional notion. This can be shown by certain kinds of cases, e.g., when an artist unintentionally produces an object that does not correspond to his or her idea (the artist accidentally knocks over some paint which makes a mess). Since the result is neither intended by the artist nor represented by the artist’s idea, it is not a ‘realization’ of

Then there is a mathematical tradition according to which realization is a *mapping or isomorphic relation*. Generally speaking, some object realizes (or is the realization of) another object because elements of the latter map onto elements of the former. For example, mathematicians speak of points in a geometric plane realizing sets of coordinate numbers by virtue of a one-to-one mapping from the sets of coordinate numbers onto the points in a geometric plane, and computer scientists speak of one machine realizing another by virtue of a one-to-one mapping from the states of the realized machine to the states of the realizing machine (Hakimi, 1962; Arbib, 1969, p.272). The mathematical tradition's idea of realization *qua* one-to-one mapping has enjoyed a wide application, being employed in various areas of mathematics, computer science, and the early philosophical literature on the computational model of the mind.

Finally, there is a metaphysical tradition according to which realization is a relation of *determination or generation*. Generally speaking, one object realizes (or is the realization of) another object only if the former produces or determines the latter. Thus Jaegwon Kim says that the “very notion of a ‘realization’ presupposes nomological connections” between the properties of realizing and realized objects (1993, p.180), which philosophers typically understand to imply that the instantiation of a realizing property supplies a lawfully sufficient condition for the instantiation of a realized property (Lepore and Loewer, 1989, p.179; Tye, 1995, p.41). The metaphysical tradition's idea of realization *qua* determination has been popular in recent philosophy of mind, science, and metaphysics.

Of course specific theories often contain supplemental notions. For example, philosophers who write in the metaphysical tradition typically add that the stated determination applies to synchronic, inter-level cases (Tye, *loc. cit.*; Kim, 1996, pp.221-

the artist's idea in the common sense of that term. Nevertheless, the common notion might involve an additional element of metaphysical production, which is illustrated by a remark from Leibniz that I quote three paragraphs hence about realized things that correspond to ideas in the imagination and are made by God. Leibniz's remark is also relevant to another question the reviewer asks about whether mixed concepts of realization are more basic when compared to the unmixed versions that I describe under the semantic, mathematical, and metaphysical traditions. While mixed versions might be *historically* more basic – witness Leibniz's early usage – the unmixed views are *conceptually* more basic by standing as simpler parts to more complex concepts.

22). Some philosophers even mix ideas from different traditions. Thus, Leibniz describes God as a being “through whom those things which would otherwise be imaginary are realized [L, ‘realiso’]” (Leibniz, 1697/1989, p.45).² So things that are realized correspond to ideas in the imagination (intentionality) and owe their existence to God (metaphysical determination). But mixed concepts aside, theories that belong to different traditions were typically developed to meet quite different goals.

Theories in the representational tradition are meant to provide meaning for a language, or content for a concept, or otherwise contribute to an understanding of intentionality. Theories in the mathematical tradition are meant to provide proof of a numerical equivalence, or supply a simulation of a particular domain, or otherwise create a formally equivalent model. And theories in the metaphysical tradition are meant to address certain philosophical problems, like the mind-body problem, or the object-constitution problem, or the many-sciences problem that concerns the relation between the ontology of special and general sciences. But whether a specific theory meets the goals associated with the tradition to which it belongs, the facts typically addressed by one tradition are both different from and yet compatible with the facts addressed by another tradition.

Consider a sample overlapping domain for two theories, one in the representational tradition and the other in the metaphysical tradition. A term or concept can be satisfied by an object while that same object (or one of its properties) is determined by its more basic engineering (or some its engineering properties). A model is depicted in Fig. 1 below, where ‘*D*’ is a type of representation or designator, ‘*s*’ is the subject designated, ‘*F*’ is a nonbasic property, ‘*G*’ is a more basic property, the horizontal arrow ‘←’ is realization *qua* intentional satisfaction, and the vertical arrow ‘↑’ is realization *qua* metaphysical determination.

² Leibniz adds that the usage is “barbaric but graphic” (loc. cit.). Why “barbaric”? Translators Daniel Garber and Roger Ariew state in a footnote that ‘realiso’ is corrupt Latin. I asked Doug Jesseph several years ago about the text, and he explained that the classical Latin ‘res’ had no verbal form, and that even the adjectival form was medieval in origin. Hence ‘realiso’ was a contrived verb that was convenient to express a thing’s being or becoming real.

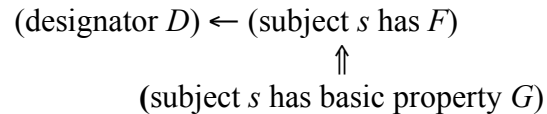


Fig. 1 *Representational and Metaphysical Traditions of Realization.* I point the arrow toward the designator to indicate that while D represents F , F thereby and conversely realizes D . In the representational tradition, realization is the converse of representation.

The phenomenon of reference to a composite object provides a familiar interpretation of the model, as when someone successfully refers to a man or a machine. There is an intentional relation between language and object (pace the semantic tradition), and there is a distinct but compatible determinative relation between that object and its more basic engineering (pace the metaphysical tradition). Similar remarks apply if one were to use mathematical mappings in the place of the semantic relation. A thing can be mapped onto an object while certain properties of the object are determined by its more basic properties. I take these points to be so obvious that they would be hardly worth saying, were it not for the fact that the idiom of realization is ambiguous by covering the different kinds of cases just mentioned, making it possible to confuse the goals associated with one theoretical tradition for those of another.

Granted, some theories might involve special considerations that create an incompatible situation, say, if pace Berkeley, the theory of representation deems material objects to be unthinkable or otherwise unrepresentable while the metaphysical theory involves determination between the properties of a material object. But I am only concerned with a select group of theories, specifically, those involved in arguments which purport to show that one theory should be chosen over another when in fact the theories are not genuine competitors. I start with an argument by Norman Malcolm that places a functional-role notion against a common-sense notion of realization. I then turn to an argument by Thomas Polger that places a family of metaphysical theories against an abstract notion of machine realization.³ I aim to show that the theories in question are

³ A reviewer asked why I omit arguments by Kim, Shoemaker, or Gillett. But, as far as I know, they do not provide clear and explicit arguments that can be resolved by appealing to different but compatible theoretical traditions. Consider Gillett's (2002) worry over flat versus dimensioned theories of realization. Both types of theory fall under the metaphysical tradition, since they are both theories of determination meant to address the

compatible, and I will use the common facts illustrated in Fig. 1 as a model for understanding why that is true.

2. Malcolm Against the Functionalist View of Realization

Norman Malcolm calls into question a familiar second-order functionalist concept of realization. More specifically, although Malcolm rejects David Armstrong's first-order functionalist view that expresses the mind-brain type-identity theory, he says: "one can get a grip on his view. The same cannot be said for those functionalists who subscribe to the causal theory but wish to avoid the identity thesis" (1984, p.97). He then articulates his complaint:

For one thing, their use of the terms 'realize' and 'realization' is baffling. In ordinary language we speak of a plan, a hope, a desire, an intention, as having been realized. ('His long-standing desire to see the Taj Mahal was at last realized'; 'Her hope to become chief editor was realized'.) In this use 'realized' comes to the same as 'fulfilled,' 'achieved,' 'satisfied.' The realization of a desire is the occurrence of what was desired. The realization of a man's desire to see the Taj Mahal would be *his seeing the Taj Mahal* ... How different is the use of 'realized' by the functionalist! According to them, a man's desire to see the Taj Mahal would be realized even if he never got to see the Taj Mahal; it would be realized by one of his neural states (ibid., pp.97-98).⁴

Here Malcolm says that the nonreductive functionalist view is "baffling" when measured against the ordinary language usage of 'realized' whereby realization is the satisfaction of a plan, hope, desire, or intention. By this he means to imply, among other things, that the functionalist view should be *faulted* because it does not fit the common

same philosophical problems, such as the mind-body problem. The theories differ primarily on the details about the determiner – is it a coincident object with a more basic property (a flat view) or the parts of an object with their more basic properties (a dimensioned view)? The choice is *not* between a theory of intentionality versus a theory of determination by more basic engineering, like the cases examined in the present paper. Parenthetically, I believe that flat and dimensioned functional theories are compatible, but not for the reasons given here. See Endicott (forthcoming).

⁴ One might wonder why first-order functionalists get a pass, since their use of 'realized' also departs dramatically from the common language usage – one does not *identify* a man's desire to see the Taj Mahal with his seeing the Taj Mahal!

usage. Malcolm then underscores the difference by pointing out that, on the functionalist view, the desire to see the Taj Mahal could be realized by neural states even when the man never sees the Taj Mahal, which is to say, even when the desire is unfulfilled. By the functional theory the desire is realized, by the ordinary usage it is not. But to judge this as a problem, as Malcolm does, is a mistake. I will assume that Malcolm's ordinary language notion can be formulated in a definition according to which a desire or more generally a mental designator is realized when and only when it is satisfied. That is:

(OR) a property F realizes a type of mental designator D if and only if D has satisfaction conditions, and the instantiation of F meets those conditions.⁵

So by this ordinary notion, *seeing* the Taj Mahal (F) realizes the *desire* to see the Taj Mahal (D) because the desire to see the Taj Mahal (D) has satisfaction conditions, and the perceptual event of seeing the Taj Mahal (F) meets those conditions. The ordinary folk notion of realization thus stands in the representational tradition, since D represents F and hence F realizes D by meeting its conditions for satisfaction. Now the functionalist theory that Malcolm rejects is the popular second-order version of realization *qua* functional roles and occupants:

(FR) a property G realizes property F if and only if some object instantiates F and G , F is the property of having some property that occupies a causal-role R , and G occupies role R .

This is a theory in the metaphysical tradition. It expresses the determination of a functional property by its role-player property. Thus, given a second-order definition of F over a set of first-order properties that lawfully occupy role R , and given that G is a member of that set, it follows that if an object has G it must have F . So my point can be put very simply: the common folk view is an intentional notion that stands in the

⁵ Here and elsewhere I define realization in terms of properties. Others define realization in terms of property instances. Some think this leads to conceptual tangles. See Thomas Polger and Laurence Shapiro (2008). But I think definitions can be formulated either way, with the proper auxiliary assumption that allows the concept of realization to range over multiple categories. See Endicott (2010).

representational tradition, while the functionalist view is a notion of determination that stands in the metaphysical tradition, and the two are perfectly compatible. Certainly the latter should not be faulted for failing to capture the intentionality of the former.

To appreciate the point, notice that the situation Malcolm describes is very much like what is depicted in the previous Fig. 1, save some inconsequential differences that will now be addressed. For one, he discusses the intentional relation in terms of mental states such as a plan, hope, desire, or intention, rather than a broader class that includes linguistic designators. Also, he does not describe these mental states in terms of functional concepts. Even so, the functional case can be depicted below, where ‘*D*’ is a type of mental designator, ‘*F*’ is the represented second-order functional property, ‘*G*’ is the first-order property that plays the causal role associated with *F*, while the representational and metaphysical relations remain the same, as shown in Fig. 2a.

The representational theory of ordinary-language realization (OR) that describes the facts depicted at the top of Fig. 2a does not compete with the metaphysical theory of functional-role realization (FR) that describes the facts depicted on the right side of Fig. 2a. The theories simply explain different things about the object *s* – how it is represented versus how it is determined by its more basic properties. Moreover, the point remains the same if, as Malcolm considered, the mental designator is a desire *D* that is metaphysically realized by neural state *G*. The theories still pick out different but compatible facts, as depicted in Fig. 2b. Furthermore, since the functionalist theory aims to explain a different relation than the ordinary intentional notion, there is no reason to be baffled by Malcolm’s variation on the case according to which the man’s desire to see the Taj Mahal is functional-role realized by his neural states even when the desire is unfulfilled. This just means that the man’s desire is *metaphysically* realized by his neural

Fig. 2 a and b
Ordinary-Language and
Functional-Role Realization

a (mental designator *D*) ← (s has second-order *F*)
 ↑
 (s has first-order *G*)

b (s has second-order mental designator *D*) ← (*s'* has *F*)
 ↑
 (s has first-order neural property *G*)

states (the fact depicted on the left-side of Fig. 2b) even when nothing counts as the *ordinary intentional* realization of that desire (when the fact that s' has F depicted at the top-right of Fig. 2b does not exist).

Of course the ordinary intentional facts and functional-role facts depicted in Fig. 2a and 2b are compatible barring special considerations. And perhaps Malcolm means to indicate a special problem when he goes on to say that the functional theory is unacceptable because it implies dualism under the interpretation according to which a neural state “produces” or is “correlated with” a mental state (1984, p.98). But Malcolm’s interpretation is slightly uncharitable, since functionalists explicitly say more. They say that the neural state must play the causal role of the mental state, which is not implied by a dualist language of mere correlation. Indeed, it is a familiar virtue of functionalist theories that they supply an explanation for mind-body correlations (Levine, 1993, pp.129-132; and Horgan, 1993, p.579). More important for present purposes, Malcolm’s complaint about dualism is different than the preceding argument that the functionalist theory is unacceptable because it conflicts with the ordinary language notion. At least he would need to show that functional role-occupant dualism is inconsistent with the ordinary intentional notion of realization, yet nothing he says would indicate a tension on that score (some argue that nonreductive functionalism is inconsistent with *physicalism*, but the latter is hardly the ordinary folk notion in question). My concern, in any case, is with the alleged conflict and how it can be resolved by minding the different traditions of realization and respecting their different but compatible goals.

3. Polger Against Causal Property Views of Realization

Next I consider a recent argument by Thomas Polger that raises some interesting issues about machine functionalism and a diverse family of theories about realization. Polger says that his argument “applies equally to Gillett’s dimensioned view and to the flat causal views of Kim and Shoemaker” (2007, p.239). The targeted views include any theory wherein the causal powers that individuate the realized property are inherited from the realizing property or properties. The causal theories are then “flat” or “dimensioned” depending upon how they parcel out that causal inheritance.

According to a flat view, (I) the realized and realizing properties are possessed by the same individual, and (II) the causal powers that individuate the realized property are matched by the causal powers of the realizing property. Sydney Shoemaker's (2001) subset view for a single individual counts as a flat theory, since the causal powers of the realized property are a proper subset of and thus are matched by the causal powers of the realizing property (see fn.15 for more detail). Likewise, Jaegwon Kim's (1988) combination of a causal inheritance principle with a functional theory of realization counts as a flat theory, since the causal powers that individuate the realized property are inherited from the realizing property, thus enabling the latter to play the causal role of the former. In contrast, according to a dimensioned view, (I*) the realized and realizing properties are possessed by an individual and its proper parts, respectively, and (II*) the causal powers that individuate the realized property are inherited from the distinct causal powers of the properties of the parts. Carl Gillett (2002), who introduced the flat and dimensioned distinction, promotes a dimensioned causal theory.

What is Polger's argument against these views? Focusing on the paradigm case of a computing machine realizing the mathematical function of addition, he argues that: "the formal or mathematical relations are not themselves causal relations, so their realization cannot be in virtue of the physical system (or its parts) contributing the causal powers that are individuating of them" (2007, p.240). And again he puts the point in terms of "the archetypal case of the realization of a formal program or algorithm," meaning a formal program in the sense associated with an "abstract machine" rather than "physical machine," claiming that the realization of such *abstracta* are "neglected" by causal theories (*loc. cit.*).⁶ Polger then summarizes his complaint as a conflict between causal theories and the abstract machine version of Realization Physicalism (RP) according to which special science properties are realized by but not identical to physical science properties:

⁶ Polger makes it clear that his argument concerns the relationship with machine *abstracta*: "There is a well-known ambiguity in familiar explications of Turing machines, between abstract and physical machines ... Here I am focusing on abstract computational functions, as will be clear from my examples" (*ibid.*, p.240, fn.12).

The crux of my objection to the causal approach of Gillett, Kim, and Shoemaker requires that what is realized be causally individuated. On those views abstract algorithms like addition cannot be realized because they are not causally individuated. But mathematical and computational realization are standard examples for RP, so the causal approach must be mistaken (ibid., p.241).

One way to understand Polger's "Noncausal Realization" argument is as follows:

- (i) If something is realized according to flat and dimensioned causal theories, it is individuated by causal powers inherited from the causal powers of the realizing items.
- (ii) But abstract machines as well as their formal and mathematical properties are not individuated by causal powers inherited from the physical realizing machines, their physical realizing components, or their physical properties.
- (iii) Hence abstract machines and their formal and mathematical properties are not realized according to flat and dimensioned causal theories.

Now I think this part of the argument is sound, as long as one keeps a clear divide between the "abstract machine" and the "physical machine" (more on that later). But even if Polger is correct that causal theories do not account for abstract machine realization, it does not follow that they are mistaken unless one assumes in addition that:

- (iv) To be an acceptable theory of realization, abstract machines and their formal and mathematical properties should be realized according to that theory.
- (v) Hence flat and dimensioned causal theories are not acceptable theories of realization.

Yet I will argue that one should reject premise (iv). Certainly Polger has put his finger on a puzzle. Abstract machine realization has been a central case in the philosophy of mind and science. So it would seem to be a desideratum for any adequate theory of realization to account for it. Nevertheless the key lies in the different conceptual traditions of realization. One can accept that abstract items in the traditional machine-functionalists story are not 'realized' in the sense relevant to causal theories, since machine functionalists used the idiom of realization with a different meaning to account for a different but compatible phenomenon about machines.

Consider again Gillett's dimensioned theory. It is a mereological theory of realization *qua* parts and wholes, where the causal powers of the whole are determined by the causal powers of the proper parts. As a theory of determination it falls squarely within the metaphysical tradition of realization. I will define it more formally thus:

(DR) Properties P_1, \dots, P_n realize property F if and only if there is a system s that instantiates property F , there are objects x_1, \dots, x_n that instantiate P_1, \dots, P_n , objects x_1, \dots, x_n are the proper parts of s , the causal powers that individuate property F are inherited from the causal powers of P_1, \dots, P_n , and, necessarily, for any s and any x_1, \dots, x_n , if x_1, \dots, x_n are the proper parts of s , then if x_1, \dots, x_n have P_1, \dots, P_n then s has F .⁷

In contrast, the traditional theory of machine realization that Polger describes is a relation between a physical machine and either a machine-table *description* or some *abstract machine* that is expressed by that description (Putnam, 1975, pp. 365, 371). In the parlance of the time, the abstract machine supplies the sense of the description, while the physical machine is its referent. For present purposes the difference does not matter, since machine descriptions as well as their *abstracta* stand in the same kind of relationship to the physical machine. Polger addresses that relation in terms of the abstract function of addition: "What must be the case for a physical system to realize addition? The physical system must have states whose causal relations to one another somehow 'correspond' or 'map' onto the mathematical relations characterized by addition" (2007, p.240). Polger then quotes Van Gulick, who puts the point in terms of machine descriptions:

Instantiation [i.e., realization] of such a formal machine description requires roughly that there be some mapping relation from the formal states, inputs, and outputs of the abstract machine table onto the physical

⁷ This is only a basic idea of part-whole realization because, pace Gillett, it contains no reference to a complex aggregate of the parts that is coincident with x and no reference to complex structural property or organizational feature that could be identified as a role-player G for F . Gillett defines dimensioned realization in terms of property instances: "Property/ relation instance(s) P_1-P_n realize an instance of a property F , in an individual s , if and only if s has powers that are individuating of an instance of F in virtue of the powers contributed by P_1-P_n to s or s 's constituent(s), but not vice versa (2002, p. 322, with a change in the variables).

states, inputs, and outputs of the instantiating system, such that under that mapping the relations of temporal sequence among those physical items are isomorphic to the relations of formal succession among the machine table items (Van Gulick, 1988, p.80; Polger, loc. cit.).

Reference to the “instantiation [i.e., realization] of such a formal machine description” and “machine table items” that map onto temporal sequences of the physical machine is just another way of talking about the satisfaction of a machine description by a physical machine – which means that the traditional mapping view is an intentional story that falls under the representational tradition of realization. As Putnam said: “a ‘machine table’ describes a machine if the machine has internal states corresponding to the columns of the table, and if it ‘obeys’ the instructions in the table” (1975, p.365). Indeed, Polger likewise says that the noncausal aspects of abstract machines are “intentionally realized” and that the “relation between physical devices or systems and mathematical functions is usually thought of as an isomorphism, mapping, or representation” (2007, pp. 241, 246).

One slight complication is that the machine-functional concept of realization has a foot in both representational and mathematical traditions by employing a mathematical mapping. But the mapping is an unsurprising component in a traditional intentional story, since it was a once popular way to express the semantic relation of “correspondence” (see Pitcher, 1964, pp.9-10). And the main point is just that machine functionalists used the idiom of realization in the stated intentional way.⁸ So I will define this notion like Van Gulick’s statement, only adding the alternative formulation about abstract machines:

(MR) a machine-functional property F of a system s realizes a machine-table description D (or abstract machine M) if and only if D (M) has satisfaction conditions, the instantiation of F requires a set of input-procedure-output processes, and the instantiation of F by s meets the

⁸ So Arthur Burks and Jesse Wright used the term ‘represents’ as the converse of ‘realizes,’ saying that a logical net N “represents” a circuit C can be expressed alternatively as “ C physically realizes N ” (1953, p.194). I add that, although machine functionalists used the idiom of realization in the stated intentional way, it does not follow that they were unconcerned with metaphysical goals and theories. Rather, they expressed them in a different way. In Fodor’s case, for example, his theory of functional decomposition was a dimensioned theory in the metaphysical tradition, as I explain toward the end of section 4.

satisfaction conditions of D (M) because the elements of D (M) map one-to-one onto members of the set of input-procedure-output processes of s required by F .⁹

Now when this theory of machine-functional realization is placed against the backdrop of the different conceptual traditions of realization, the resolution of Polger's noncausal realization argument seems clear. The machine-functional theory (MR) stands in the representational tradition, while the theory of dimensioned realization (DR) stands in the metaphysical tradition, and the two simply explain different facts about machines. A computational system s can satisfy a machine-functional description D (or an abstract machine M) and have its denoted functional property F determined by its collective engineering properties P_1 - P_n , as shown below:

$$\begin{array}{c} \text{(description } D \text{ or abstract } M) \leftarrow (s \text{ has } F) \\ \uparrow \\ (s\text{'s parts have basic properties } P_1\text{-}P_n) \end{array}$$

Fig. 3. *Machine and Dimensioned Realization*

The fact that the machine description D or the abstract machine M lacks causal powers inherited from the properties of denoted system s or its parts is irrelevant to the account of the relationship between the property F instantiated by s and the properties P_1 - P_n of its parts.¹⁰ As a consequence, the metaphysical theory can be fully adequate to its intended domain (the facts depicted on the right side of Figure 3) without accounting for either the intentional relation or the abstract machines covered by the theory of realization *qua* one-to-one mapping (the facts depicted at the top of Figure 3). So premise (iv) in Polger's noncausal realization argument is false. Specifically, it is false that an adequate metaphysical theory of realization should account for the intentional realization

⁹ I do not believe this is an adequate account of the satisfaction conditions for a machine-table description, only that it approximates the once traditional view. For some added conditions, see Lycan (1987) and Chalmers (1994).

¹⁰ The point remains the same if one replaces the mereological dimensions of part-whole determination with the kind of facts described by a flat causal theory wherein the same object s has a realizing property G that contributes all the causal powers possessed by its realized property F (that is, if one replaces "s's parts have basic properties P_1 - P_n " with "s has G " in Figure 3).

of abstract machines and the formal and mathematical properties they possess. But notice that the premise is not altogether false – a metaphysical theory of realization should account for *something* about machine realization, namely, the metaphysical realization of the denoted machine. That is the goal of the causal theories in question, and nothing said thus far about the noncausal status of machine *abstracta* will frustrate that goal.

Yet one might think that my compatibilist picture misses an important theme in Polger’s argument. It is *F*’s realization by the system *s* – not just the properties of a machine description *D* or abstract machine *M* – that is supposedly noncausal in a way that conflicts with the flat and dimensioned theories of inherited causal powers. But that claim is unsupported by mere talk about the satisfaction of machine tables, descriptions, and their associated *abstracta*. Even though descriptions that exist outside the system *s* and the *abstracta* that exist in Plato’s heaven are not individuated by causal powers inherited from the denoted system *s* or its components, the instantiation of the functional property *F* by the system *s* would seem to be a different matter. At least one would need another argument about *F*’s realization by *s* that does not trade on the status of items like descriptions and their associated *abstracta*. To that issue I now turn.

4. Responses and Further Reflections on Machine Realization

There is more to Polger’s noncausal realization argument, a point about the individuation of *F* based upon the *abstracta*, which comes out clearly when he considers a criticism that his argument trades on an ambiguity. This issue is worth carefully sorting out, especially given the important place that machine functionalism has occupied in contemporary philosophical thinking. Thus Polger mentions the criticism:

It could be suspected that I am trading on an ambiguity in notions like *computation*, *program*, or *algorithm*. These may pick out abstract relations, but they may also pick out physical operations. We can concede that the mathematical function of *addition* is not causally individuated, and thus cannot be realized according to the causal approaches. But *adding* is a procedure or operation (an algorithm’ in the causal sense) and the property or state of *being an adder* (i.e., an *adding machine*) can be causally individuated. Even if addition cannot be realized, adding and

adders can be realized. And that is all that RP requires (2007, p.245).¹¹

Polger then responds that special considerations about noncausal individuation for symbolic properties create a conflict with the requirements for individuation laid down by causal theories of realization. That is, even though the property is possessed by a physical machine, the status of *F* as a symbolic property requires that it should be individuated by conditions based upon the mapping relations with the *abstracta* of the intentional story rather than the causal powers of a causal theory.¹² So Polger says:

Of course realizing an *adder* involves having some causal powers. It does not, however, involve having the causal powers distinctive of *addition*, for this objector concedes that there addition is not causally individuated. But for the same reason there is no distinctive set of causal powers that individuate adders. What makes something an adder (an adding to machine) is that it has causal powers that stand in some relation to the mathematical *addition* function. The relation between physical devices or systems and mathematical functions is usually thought of as isomorphism, mapping, or representation. So, following Van Gulick, Cummins, and Marr, we may say that when the physical states and state transitions of a device map or represent the formal states and transitions of the addition function, then the device realizes the property or state of being an adding machine. The crucial point is that although an adding machine is realized by something's having particular causal powers, it is not realized—as Gillett, Kim, and Shoemaker would require—by anything's having the causal powers individuating of adding machines (2007, p.246).

¹¹ Polger credits the objection to Chase Wrenn, Jose Luis Bermudez, and several anonymous referees. For the distinction between abstract and concrete programs, or external versus encoded programs, see Newell and Simon (1972, p. 33); Weizenbaum (1976, p. 111). Polger adds that even if the computational case fails to establish the conclusion about noncausal realization, there are other examples that do not trade on any ambiguity between abstract versus physical machines. I will return to that issue in the final section when I address the scope of my proposal.

¹² Interpreting Polger's argument in this way requires that the original premise (ii) be modified to read: "(ii*) But the symbolic properties of a physical machine are not individuated by causal powers inherited from the physical properties of its physical components," with the remaining premises and conclusion revised accordingly. To anticipate, defenders of causal theories should reject (ii*), as long as the symbolic properties are construed in terms of the syntax of a physical machine. The semantics is once again a matter for a theory of intentionality in the representational tradition.

To put Polger's point in a different way, a machine is an adder because of the mathematical *abstracta* it is related to, not because of the powers it inherits from its more basic physical properties. And I agree with much of what he says. I think it is true that an *adder*, understood as a physical machine, does not have "causal powers distinctive of *addition*," since the latter is a relation in *abstracta* that is not causally individuated. I also think it is true that "there is no distinctive set of causal powers that individuate adders," not for the same reason that adding machines are *abstracta* that consequently are not causally individuated, but for the different reason that they are physical objects built in different ways with different parts that operate by different engineering principles. Finally, I think it is also true that adding machines stand in a mapping or representation relation to the *abstracta*. But none of this, I argue, creates a problem for causal theories of realization.

I will state my position in terms of Robert Cummins' account of how a physical system computes the function of addition, since Polger says that Cummins' account supports his noncausal realization argument (2007, p.240), and since Cummins' account clearly distinguishes the numerals in a machine from the numbers they represent – a distinction between the syntax and the semantics that will loom large in my resolution of Polger's worry. Cummins says:

To add is to be described by the plus function, $+\{m, n\} = s$. Hence, to explain what makes a system an adder is to explain the fact that the system is described by $+$. But $+$ is a function whose arguments and values are numbers; and whatever numbers are, they are not states or processes or events in any physical system. How, then, can a physical system be described by $+$? How can a physical system traffic in numbers and hence add? The answer, of course, is that numerals – i.e., representations of numbers – can be states of a physical system, even if numbers themselves cannot (1989, p.89).¹³

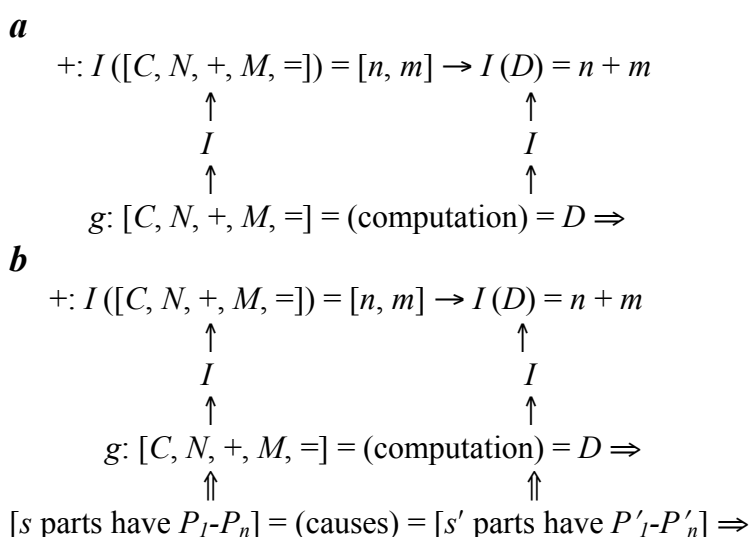
¹³ Polger quotes the more detailed passage below, but the moral is the same:

We may think of the button-pressing sequences as arguments to a function g that gives display states as values. An adding machine *satisfies* g ; that is, the arguments and values of g are literally states of the physical system. Addition, as was remarked above, relates numbers, not physical states of some machine, so a physical system cannot literally satisfy the plus function. What an adding machine does is *instantiate* the plus function. It *instantiates* addition by *satisfying* the function g whose arguments and

So the plus function takes numbers in its argument place and yields numbers as values. But numbers are abstract objects that cannot be physical states of a machine. Consequently, a physical machine can only be said to add because it has physical symbols or numerals that *represent* numbers, or can be *interpreted* as numbers (hence Cummins calls this framework “interpretational semantics”). Cummins presents the relationship between the abstract numbers and the computation over symbols by a “Tower Bridge” structure, where the input to the adding machine is a sequence $[C, M, +, N, =]$ for *clear, first addend, plus, second addend, equals*, and where the output is a display D which represents the sum of the two addends, as shown in Fig. 4a.

Fig. 4a Cummins’ Tower Bridge (1989, p.90)

b The Tower Bridge with Metaphysical Foundations



Cummins says that “the vertical arrows correspond to interpretation” (ibid., p.91), which is an intentional relation like ones depicted in Fig. 1 through 3 (only they were depicted horizontally rather than vertically). Hence, interpreting a function g computed

values represent arguments and values of the addition function, or in other words, have those arguments and values as interpretations” (Cummins, loc. cit.).

Call this “instantiation by interpretation”: addition is *instantiated* in this intentional sense when the physical machine *satisfies* the function g in the Fregean sense that its physical symbols are arguments and values of g and when those arguments and values are *interpreted* as numbers. Note that this does not make the computed function g an abstract object, like the numbers and the plus function. Physical things can be treated as functions too. E.g., the biological process of generation can be treated as a function that takes parents as arguments and yields offspring as values.

within a physical machine system s as the function of addition is like interpreting a functional property F of a physical system s in accordance with a machine-table description. And this picture fits my proposal perfectly. Symbols within a physical machine *represent* abstract objects. This is ‘realization’ in the sense specified by the representational tradition. Flat and dimensioned theories may then account for how the symbols are determined within a physical machine by its underlying engineering. That is ‘realization’ in the sense specified by the metaphysical tradition. To complete the larger picture, consider how a dimensioned theory may enter at the foundations of the Tower Bridge as a vertical relation of determination between steps in the process of s computing g and stages in a causal process involving s ’s more basic parts and properties, as shown in Fig. 4b.

The important point is that the metaphysical theory is only intended to explain the realization of the symbolic representations within the machine, *not* the abstract numbers they represent. Consequently, to answer Polger’s worry about individuation, the machine-functional property F that is causally individuated and subject to the inheritance of causal powers is *being a symbol* (e.g., being an N), which is a property of physical states in the machine, *not being a number* (e.g., being an n), which is a property of abstract objects, and *not even being a symbol for or representation of a number*, which is an intentional relation accounted for by the interpretational semantics (e.g., the interpretation I of N as n). That is, the machine is a physical symbol system whose inherited causal powers attach to the *syntax*, *not the semantics*. Indeed, familiar issues of wide content already yield the conclusion that meaning is not determined solely by facts about the system, including facts about the system’s parts. So one should not expect the kind of metaphysical theory in question to explain the semantics of a machine system.

Thus, to summarize, there is a division of labor such that the representational theory of realization explains the semantic properties of a physical machine, while the metaphysical theory of realization explains the syntactic (and other) properties determined by the basic facts of the machine. So what Polger worries about – conditions for individuation based on the relation to the mathematical *abstracta* – apply to the semantic properties of machine symbols, *not* their syntactic properties. Let me also emphasize that the larger picture of machine functionalism with the intentional

superstructure and the added metaphysical foundations was firmly grounded in traditional functionalist philosophy. In other words, I reject the idea that there was a *single* relation under consideration, “*the realization relation*,” introduced by Putnam and central to the developing machine-functionalist views (Polger, 2007, p.247). To the contrary, the machine-functional notion of intentional realization *qua* one-to-one mapping was clearly in play, as already established. A flat theory of realization was also in play, in the form of the already discussed second-order functional theory.¹⁴ And a type of dimensioned theory of realization was also in play, though not described under Gillett’s rubric of a ‘dimensioned’ theory.

Thus Cummins subsumes realization under a general theory that explains the instantiation of a property in a kind of system *S* by “the properties of *S*’s components and their mode of organization” (1983, p.15). That is the mereology of a dimensioned perspective. In fact, this mereology is central to the more publicized species of a property instantiation theory introduced to the philosophical community by Fodor (1968) and developed by Cummins (1975), namely, a functional analysis wherein higher-level (realized) functional properties or associated capacities decompose into lower-level sub-capacities that ultimately decompose into simple mechanistic-level processes (the basic realizing parts). Functional decomposition is just composition is reverse. Therefore as long as the properties in question are associated with *causal capacities*, the resulting functional analysis counts as a dimensioned theory of realization – the causal capacities, with their attendant causal powers, decompose into more basic physical parts, which, put the other way around, means that the causal capacities are determined by those compositional parts. So the traditional intentional theory of machine satisfaction and the traditional dimensioned theory of part-whole (de)composition were common parts of a larger machine-functional view.

Finally, I should add that my point is not simply that the different theories of realization in question are formally compatible. They *complement* one another. On the

¹⁴ Anthony and Levine report: “Now back in the 60s, we thought we had all this straightened out. The trick was going to be to treat psychological properties as ‘higher-order’ properties ‘realized’ by physical properties. A higher-order property is a property you have to have in virtue of having some *other* property that meets certain specifications” (1997, p.85).

one hand, the deeper mereological theory serves to explain how a physical machine instantiates the functional property of computing certain symbols (and thus by interpretation certain numbers) by virtue of its more basic computational capacities and ultimately its most basic engineering parts. On the other hand, the intentional story plays an important role in the epistemics of machine realization. Satisfaction of a mapping relation is part of the justification for applying the machine description to the object in question. In other words, mapping elements of a machine table onto the component activities of a physical system is like going through a checklist to verify that the object is indeed what someone claims it to be. And ditto for other theories of interpretation that posit additional facts to justify the intentional ascription. This kind of epistemic practice can be crucial to understanding a machine, since one typically does not have epistemic access to the metaphysical details about how a machine is constituted by its basic physical mechanisms, save the exceptional engineer. So intentional and metaphysical notions of realization form a larger picture of machine-functional systems that includes not only their metaphysical status *vis-à-vis* underlying implementing mechanisms but also how they are described and how they are known.

5. Concluding Observations

I want to finish by discussing the scope and limits of my proposal. Regarding the limits, as I stated at the outset, some arguments over different theories of realization have a point. For example, if the choice is between two theories in the same conceptual tradition with the same explanatory goals, then they are genuine competitors and my compatibilist proposal does not apply. Thus I count Yablo's (1992) theory of realization *qua* determinables and determinates and Shoemaker's (2001) theory of realization *qua* subsets of causal powers for single objects as genuine competitors. Both theories involve property determination over the same individual and thus both theories fall under the metaphysical tradition.¹⁵ In contrast, my proposal requires two theories that belong to different conceptual traditions with different explanatory goals.

¹⁵ To substantiate this claim, consider: (DR) a property *F* is realized by a property *G* if and only if there is an object *x* that instantiates *F* and *G*, *F* is a determinable, and *G* is one of its determinates. Yablo explains the kind of determination in question: "*G* determines

Regarding the scope of my proposal, aside from the cases already examined, other types of cases can be understood in the same way by minding the different traditions of realization. Consider Fodor's (1981) economic example of multiply realized monetary units (cf. Polger, 2007, pp.241, 245-246). A dollar bill has a symbolic or representative function, meaning that the symbols printed on its front and back represent the economic backing of the United States government and its Federal Reserve. Hence the same division of labor applies. The intentional or semantic properties of the symbols are realized according to theories in the representational tradition, while the syntactic (and other) properties of a dollar bill are realized according to theories in the metaphysical tradition. Consequently, the symbolic property *F* that is causally individuated and subject to the inheritance of causal powers is once again the syntactic property of *being a symbol* (rather than *representing the specific monetary value of a dollar*), in this case an arrangement of ink that is determined by the material composition of the dollar bill – the ink, linen, cotton, and synthetic threads.

Of course there might be other properties that cannot be easily assimilated into this division of labor between representational and metaphysical traditions. Thus, many philosophers argue that such things as minds and dollar bills also have historically individuated functions. On such views, a mental program is individuated in part by the historical-evolutionary fact that it arose in ancestors to fulfill a particular cognitive niche,

F iff: for a thing to be *G* is for it to be *F*, not *simpliciter*, but in a specific way" (ibid., p.252, with a slight change in the variables). He thus offers the following principle as a partial analysis: "*G* determines *F* only if: (i) necessarily, for all *x*, if *x* has *G* then *x* has *F*; and (ii) possibly, for some *x*, *x* has *F* but lacks *G*" (loc. cit.). Clause (i) is a straightforward one-way conditional law of the kind found in the metaphysical literature on realization. Now compare subset realization, which can be defined as follows: (SR) a property *F* is realized by a property *G* if and only if there is an object *x* that instantiates *F* and *G*, and the causal powers of *F* are a proper subset of the causal powers of *G*. This also implies the same kind of property determination, at least with the appropriate background assumptions about the connection between causal powers and laws. Specifically, if properties are necessarily connected to their associated powers, the powers of *G* contain the powers of *F* across all metaphysically possible worlds and the one-way conditional law expressed by Yablo's clause (i) is true with metaphysical necessity. If properties are contingently connected to their powers, the powers of *G* contain the powers *F* across the pertinent nomologically possible worlds and the same one-way conditional law expressed by Yablo's clause (i) is true with nomological necessity.

and a dollar bill is individuated at least in part by the historical-causal fact that it was printed for the United States at the U.S. Bureau of Engraving and Printing. If so, such historical properties are not realized according to flat and dimensioned causal theories, since the properties are not wholly determined by the causal powers inherited from the system's material composition.

But I see no reason why causal theorists could not accept a similar division of labor in their explanation of this larger history of the world. Specifically, they could account for the realization of syntactic and other properties within the chosen set of minded, economic, and other-natured systems, and leave the realization of the semantic, historical, and other-natured properties to the appropriate semantic, historical, and other-natured theories. Any call to make the idiom of realization *univocal* would already run afoul of the fact that there are established conceptual traditions of realization with importantly different theoretical and explanatory goals.

Finally, I believe my proposal has an even larger scope than what has been previously indicated. Recall that Malcolm and Polger argue against certain metaphysical theories of realization on grounds that they do not capture certain intentional facts explained by a representational theory of realization. But I also believe philosophers have taken the converse attitude by avoiding the use of representational theories of realization on grounds that they do not capture certain facts of determination explained by a metaphysical theory of realization. Specifically, I think there is a widespread attitude that the once popular theory of machine realization *qua* one-to-one mapping was rightly replaced by a preferred metaphysical theory of determination such as realization *qua* functional roles and occupants, realization *qua* subsets of causal powers, or realization *qua* parts and wholes.

The history certainly fits. The mapping view of realization dominated discussions in the philosophy of mind from the 1960s through the 1970s. Then Ned Block (1978) and William Lycan (1979) published papers at the end of that era vividly illustrating how an unadorned mapping relation could be satisfied by nonminded or otherwise noncomputational systems (the Chinese Nation or a group of Lilliputians, all taken collectively). Call this the problem of "wayward correspondence." Soon afterward philosophers almost universally dismissed the mapping view of realization. Indeed, aside

from historical references to the mapping view, or citations of the Block-Lycan-style problem of wayward correspondence, few philosophers even bothered to modify the mapping view with additional conditions for machine-table or machine-description satisfaction (but see Lycan, 1987, pp. 37-41; Chalmers, 1994). Otherwise metaphysical theories of determination dominated the discussion from the 1980s until the present day.

Perhaps philosophers simply became more interested in metaphysical theories of realization. But it would seem that, unless philosophers viewed theories of mathematical mapping and the subsequent metaphysical theories as genuine *competitors* such that the latter should be chosen over the former, the natural reaction to the problem of wayward correspondence would have been to continue with the representational tradition and focus upon efforts to strengthen the mapping view into an adequate account of the conditions under which machine descriptions and their expressed functions are satisfied.

But whether one argues against certain metaphysical theories of realization on grounds that they do not capture certain intentional facts explained by a representational theory of realization, or whether one avoids the use of representational theories of realization because they do not capture certain facts of determination explained by a metaphysical theory of realization, the reasoning fits a general pattern. A philosopher argues that one theory of realization is better than another because it provides a better explanation for a particular range of phenomena, say, accounting for common sense cases, or cases within the sciences, when in fact the theories in question are not genuine competitors because of their different but compatible goals. The moral is that the division between different conceptual traditions carries consequences that have been underappreciated. While philosophers have been aware of the many different concepts of realization, some have failed to situate the theories within their proper categories, leading to apparent tensions where in fact there is none.

Ronald Endicott

North Carolina State University

Acknowledgements: I have benefited from several discussions, some that trace back several years. I owe a special debt of gratitude to Carl Gillett, Jaegwon Kim, William Lycan, Brian McLaughlin, Thomas Polger, John Post, and an anonymous reviewer.

References

- Anthony, L., and Levine, J. (1997). "Reduction with Autonomy," *Philosophical Perspectives* 11 (1997): 83-105.
- Arbib, M. (1969). *Theories of Abstract Automata*, Englewood Cliffs, NJ: Prentice-Hall.
- Block, N. (1978). "Troubles with Functionalism," in *Minnesota Studies in the Philosophy of Science* 9: 261-325.
- Burks, A., and Wright, J. (1953). "Theory of Logical Nets," *Proceedings of the Institute of Radio Engineers* 41: 1357-1365.
- Chalmers, D. (1994). "On Implementing a Computation," *Minds and Machines* 4: 391-402.
- Cummins, R. (1975). "Functional Analysis," *Journal of Philosophy* 72: 741-765.
- Cummins, R. (1983). *The Nature of Psychological Explanation*, Cambridge MA: MIT Press.
- Cummins, R. (1989). *Meaning and Mental Representation*, Cambridge MA: MIT Press.
- Endicott, R. (2005). "Multiple Realizability," in D. Borchert, ed., *The Encyclopedia of Philosophy*, vol. 6, 2nd Edition, US: Thomson Gale, Macmillan Reference: 427-432.
- Endicott, R. (2010). "Realization, Reductios, and Category Inclusion," *Journal of Philosophy*, vol. 107, no. 4: 213-219.
- Endicott, R. (forthcoming). "Flat versus Dimensioned: The What and How of Functional Realization" *Journal of Philosophical Research*.
- Fodor, J. (1968). "The Appeal to Tacit Knowledge in Psychological Explanation," *Journal of Philosophy* 65: 627-40.
- Fodor, J. (1981). "Special Sciences." Reprinted in *Representations: Philosophical Essays on the Foundations of Cognitive Science*. Cambridge, MA: MIT Press, 127-145.
- Gillett, C. (2002). "The Dimensions of Realization: A Critique of the Standard View," *Analysis* 62: 316-322.

- Hakimi, S. (1962). "On Realizability of a Set of Integers as Degrees of the Vertices of a Linear Graph," *Journal of the Society for Industrial and Applied Mathematics* 10: 496-506.
- Horgan, T. (1993). "From Supervenience to Superdupervenience: Meeting the Demands of a Material World," *Mind* 102: 555-586.
- Kim, J. (1993). "Psychophysical Supervenience," rpt. in *Supervenience and Mind*, Cambridge: Cambridge University Press, 175-193.
- Kim, J. (1996). *Philosophy of Mind*, Boulder, CO: Westview Press.
- Kim, J. (1998). *Mind in a Physical World*, Cambridge MA: MIT Press.
- Leibniz, G. (1697), "On the Ultimate Origination of Things," in Philosophical Essays, trans. by D. Garber and R. Ariew. Indianapolis, IN: Hackett Publishing (1989), 41-48.
- Lepore, E., and Loewer, B. (1989). "More on Making Mind Matter," *Philosophical Topics* 17: 175-191.
- Levine, J. 1993: "On Leaving Out What It's Like," in Davies, M. and Humphreys, G., eds., *Consciousness: Psychological and Philosophical Essays*, Oxford: Blackwell, 121 - 136.
- Lewis, D. (1980). "Psychophysical and Theoretical Identifications," rpt. in *Readings in Philosophy of Psychology* 1, Cambridge MA: Harvard University Press, 207-215.
- Lycan, W. (1979). "The New Lilliputian Argument Against Machine Functionalism," *Philosophical Studies* 35: 279-287.
- Lycan, W. (1987). *Consciousness*, Cambridge MA: MIT Press.
- Malcolm, N. (1984). "The Causal Theory of Mind," in *Consciousness and Causality: A Debate on the Nature of Mind*, Oxford: Basil Blackwell.
- Newell, A., and Simon, H. (1972). *Human Problem Solving*, Englewood Cliffs, NJ: Prentice-Hall.
- Pitcher, G. (1964). "Introduction," in *Truth*, ed. G. Pitcher, Englewood Cliffs, NJ: Prentice-Hall, Inc.
- Polger, T. (2007). "Realization and the Metaphysics of Mind," *Australasian Journal of Philosophy* 85: 233-259

Polger, T., and Shapiro, L. (2008). "Comments and Criticism: Understanding the Dimensions of Realization," *Journal of Philosophy* 105: 213-222.

Putnam, H. (1975). "Minds and Machines," rpt. in *Mind, Language and Reality: Philosophical Papers*, vol.2, London: Cambridge University Press, 362-385.

Shoemaker, S. (2001). "Realization and Mental Causation," in C. Gillett and B. Loewer, eds., *Physicalism and Its Discontents*. Cambridge: Cambridge University Press, 74-98.

Tarski, A. (1956). "On the Concept of Logical Consequence," rpt. in *Logic, Semantics, Metamathematics: Papers from 1923 to 1938*, Oxford: Clarendon Press, 409-420.

Tye, M. (1995). *Ten Problems of Consciousness*, Cambridge MA: MIT Press.

Van Gulick, R. (1988). "Consciousness, Intrinsic Intentionality, and Self-Understanding Machines," in *Consciousness in Contemporary Science*, ed. A. J. Marcel and E. Bisiach, Oxford: Clarendon Press, 78-100.

Weizenbaum, J. (1976). *Computer Power and Human Reason: From Judgment to Calculation*, NY: W.H. Freeman and Company.

Yablo, S. (1992). "Mental Causation," *Philosophical Review* 101: 245-280.