Part VI

REALISM

https://doi.org/10.1086/psaprocbienmeetp.1988.1.192982 Published online by Cambridge University Press

Five Theses on Instrumental Realism

Davis Baird

University of South Carolina

1. Introduction

Some may find an oxymoron in my title. But, my use of "instrumental" is to focus attention on the real instruments of science–pumps, dynamos and cyclotrons–and not the view that scientific theories are best understood as instruments. In what follows I characterize and argue for a kind of realism strongly wedded to what we *do* with scientific instruments, and divorced from what our theories may *say* about the entities manipulated by these instruments. My discussion owes much to Ian Hacking's "Experimental Argument for Realism" (Hacking 1983, ch. 16).

The following fantasy might help give some idea of the view I am interested in. Suppose that the diskettes, which many people use in conjunction with a small computer, do not store information magnetically; in fact they store it "radioactively," exactly how, is not relevant. IBM decided that many people in their target personal computer market would be put off by "nuclear diskettes." So they lied; they told people that the diskettes worked magnetically. To date, their secret has been well kept. Everyone outside a few special folks at IBM believes that the diskettes work magnetically. Nonetheless we all use these diskettes, with only an occasional problem, to store and manipulate information.

This situation suggests the standard–too many theories–argument against realism: the behavior of a bit of the world can be adequately predicted by a variety of different theories. Here, the truth is that the diskettes work "radioactively," but, most of us falsely believe that they work magnetically. Nonetheless, our false theory is empirically adequate. Consequently, so the argument goes, we should not have confidence that there really are the entities which the theory we accept says there are; another, equally adequate (in this case, true) theory might describe the world without such entities.

At the same time, no one would doubt that *something* on the diskettes stores information. Our confidence that there is *something* there to store the information comes from our ability to use the diskettes in a variety of ways. The "theory" that information is stored magnetically, is not the source of this confidence. We can be convinced in the reality of things, while at the same time, acknowledge that what we know about these things is both very little and very fallible. One needs only to *act* in the world to be a realist about various parts of the world.

PSA 1988, Volume 1, pp. 165-173

2. The Independence Thesis

Scientific realism has commonly been made into a claim about meaning: the entities, processes or what-not postulated by a realistically interpreted theory must refer to actual entities, processes or what-not; furthermore, and importantly, since the terms for these theoretical entities get their specific meaning from the theories which postulate them, the entities must be largely as they are portrayed in the theories which postulate them. Phlogiston is not real because there is no stuff with the properties ascribed to phlogiston.

The Independence Thesis of Instrumental Realism is that questions about realism–the existence or not of some bit of the world–are independent of questions about meaning. We can be convinced that *something* is there even if we know very little about that something.

Consider artificial intelligence [AI]. The problem of creating computer thought has received attention from philosophers in part because it seems to pose a problem about realism. Terry Winograd published a long paper describing a "computer system for understanding English" (Winograd 1973 p. 1). An Editorial Note preceding the paper assures the reader that:

> Winograd's system is not a "simulation," but it incorporates important ideas about human syntactic, semantic, and problem-solving abilities, and in particular about their interactions in understanding natural language (Winograd p. vii).

And, indeed, Winograd speaks freely about what the computer *knows*, what it *assumes*, and particularly how it *understands* English. Yet John Searle criticizes just this kind of talk in his widely discussed "Minds, Brains, and Programs." Searle criticizes "specifically the claim that the appropriately programmed computer literally has cognitive states and that the program thereby explains human cognition" (Searle p. 351).

The issue seems to be whether or not computers can have *real* cognitive states. Searle contends that they cannot; Winograd and his editor disagree. Yet, no one claims that AI remarks about various semantic categories refer to nothing. *Something* is happening on those silicone chips. The issue is whether these silicone chip events are similar or identical to the biological/mental events which make up human understanding. So we do not have a question of realism after all; we have a question about the relationship of two bits of reality. This is the Independence Thesis at work: We frequently mistake questions about the *nature* of some bit of the world for questions about the *existence* of this bit of the world.

There is a problem. We cannot just point to an electron or a computer belief–as we can to a pliers–and say, "*That*, whatever its properties may be, is a computer belief." At one time it was popular to say that the meaning of a term for a theoretical entity is implicitly given by the specific theory which employs the term. In his "Meaning of Meaning" (1979) Hilary Putnam shows how to avoid this problem by holding the *reference*, but not other aspects of the meaning, of terms constant between different theories. Putnam demonstrates how a theory of meaning could partially disassociate the meaning of theoretical terms from the theories in which they occur. Thus, the words one uses to describe some bit of nature can *refer* to the objectively real world, even when one misunderstands this bit of nature.

3. The Intervening Thesis

Putnam shows how questions about meaning can be considered independently from questions about realism. He does not show how to establish reference to a "something-we-know-little-what." What replaces *pointing* in the case of entities like computer beliefs?

The answer is given by the Intervening Thesis of Instrumental Realism: Our ability to intervene and manipulate these whatever-they-may-bes with our instruments, to produce consistent, reliable and reproducible effects, provides one guarantee that we are engaged with a piece of the real world. There may be other guarantees, but the way of instrumental intervention is of paramount importance in science.

The work done by AI researchers is primarily doing. Winograd's paper elaborates on the programming details of *how* the system goes about understanding, remembering and manipulating English language expressions. One small example runs:

> In order to interpret a sentence like "Find a block which is taller than the one I told you to pick up." the system must use a clause ("you to pick up") as the object of a verb ("tell"). It generates a pseudo-object of the type #EVENT, and creates an OSS [Object Semantic Structure] for that object (Winograd, p. 153).

Winograd creates a system which can manipulate formally created semantic objects. The system is programmed to distinguish "real" semantic objects from "pseudo-objects" which are used temporarily to achieve the ultimate "understanding." All of this is implemented. It is his ability to produce desired effects consistently and repeatedly which gives us the confidence that he is manipulating *something*.

It is the Intervening Thesis which Hacking argues for in his "experimental argument for realism" (Hacking, 1983, ch. 16). Hacking describes how a positron gun is built at the Stanford Linear Accelerator Center (SLAC). There is a source of laser light. There is an ordinary polarizer to longitudinally polarize the light. There is a device called a Pockel's cell to circularly polarize the light. There is a radioactive decay device to randomize the direction of the circularly polarized light. Etc. Each researcher on the team brings a particular special expertise to the project; one well understands the operation of the Pockel's cell, another the radioactive randomizer, etc. The expertise the researchers bring has been gained through careful experimentation with the particular devices in question.

It is not the truth of certain propositions that allows a valid inference that electrons are real. Rather, the argument relies on how electrons are treated in the laboratory. Researchers at SLAC treat them in ways which do not differ relevantly from the way ordinary objects are treated. All but the most phenominalist or solipsist of thinkers conclude that rose bushes are real. They are paradigm real things and are treated as such: they are fertilized, hybridized, admired, and so on. The reality of rose bushes is not a consequence of anyone's ability to assert truths about them; it is an aspect of how people interact with them. Once researchers use electrons in the manner of the experiment Hacking describes, they have become part of the run-of-the-mill real world of experimental physicists in much the same way that rose bushes are part of the run-of-the-mill world of gardeners.

4. The Historical Thesis

The Intervening Thesis seems to pose another problem, for scientists and engineers have described themselves as intervening in nature with many entities we are not

https://doi.org/10.1086/psaprocbienmeetp.1988.1.192982 Published online by Cambridge University Press

initially inclined to say are real. James Watt, of steam engine fame, intervened with phlogiston to determine that water was not a simple substance. The Intervening Thesis appears to be too generous, for few now regard phlogiston as real.

This problem can be surmounted by attending to the Historical Thesis of Instrumental Realism: Our attempts to understand the real stuff of the universe have an historical pedigree. As we come to understand more, the detailed descriptions of what we *know* and what we *do* changes. If these descriptions are of something which exists–something real–then it will be possible to trace a history of these uses of language from earlier antiquated names and actions to current usages. If not, such histories simply will stop when we realize that there is nothing of the sort being described.

James Watt is remembered for his improvements to the steam engine. He also engaged in a bitter priority dispute with Henry Cavendish over who discovered that water is not a simple substance. Watt was up on the chemistry of his day. Much of what he says he is doing makes little sense by today's standards; he believed in a modified phlogiston theory, and in a theory of substantial latent and sensible heats. Nonetheless, these outdated ways of *talking* about things did not hinder Watt's *doing* with things.

Newcomen's steam engine works by condensing the steam in a cylinder with its piston withdrawn. This creates a vacuum, so that the weight of the atmosphere forces the piston into the cylinder. Watt was disturbed by the amount of steam it took to fill the cylinder. After a long series of experiments, Watt came to the conclusion that the relative proportion of one of the material components of steam, its "latent heat," was the problem. One of Watt's early ideas was to operate the engine at higher pressures:

> That, in proportion as the sensible heat of steam increases, its latent diminishes, so, in the steam-engine working with pressures above 15 lbs. must be more advantageous than below it; for not only the latent heat is diminished, but the steam is considerably expanded by the sensible heat which is easily added (Letter from Watt to John Roebuck; quoted in Muirhead, p. 161).

Watt abandoned this idea because high pressure engines were dangerous. Although he continued to look for ways to alter to his advantage the proportion of (a substantial) latent heat in the steam (See Baird 1988 for more details).

Watt first communicated his discovery that water is not a simple substance in a letter to Joseph Preistly. Watt writes, "Water is composed of dephlogisticated air and phlogiston deprived of part of their latent or elementary heat" (Quoted in Muirhead, p. 321). Later in 1783, he wrote to Joseph Banks, the Secretary of the Royal Society, with a recipe for making water:

> To make Water.–
>
> R. Of pure air and of phlogiston Q. S., or if you wish to be very exact, of pure air one part, of phlogiston, in a fluid form, two parts, by measure. Put them into a strong glass vessel, which admits of being shut quite close; mix them, fire them with the electric spark; they will explode, and throw out their elementary heat. Give that time to escape, and you will find the water, (equal in weight to the air), adhering to the sides of the vessel. Keep it in a phial close corked for use (Quoted in Muirhead, p. 322).

Clearly we now would want to say that the "pure air" Watt writes of is what we now call oxygen. Phlogiston is what we now call hydrogen.

Do phlogiston or substantial latent heat exist? Before 1778, phlogiston was supposed to be a substance which, when combined with metallic ore, produced a metal; it was also a substance which humans threw away by respiring. This stuff does not exist. Neither does the stuff with the properties Watt ascribed to latent heat. But, there is a substance, which through suitable re-categorizations of other substances, would play the role of hydrogen–in that, among other things, it combined with pure air (oxygen) to produce water–and which could have many of the other properties phlogiston was supposed to have had. This was the phlogiston of the "modified phlogiston theory" current in 1785 (Conant, p. 110). Similarly, we now know much more about latent heat. The phenomena Watt investigated can be explained in terms of the modern *property*, latent heat.

While Watt held many false beliefs about phlogiston or latent heat, he could still manipulate them to learn about water and the steam engine. In the sense that we may get almost everything wrong about some postulated stuff, but still want to say that our different and successively improved theories are all talking about the same stuff, it seems that phlogiston exists and "hydrogen" is our current name for it. Similarly we can say sensibly that Watt's latent heat–the substance–exists.

Return to artificial intelligence. In his seminal paper, "Computing Machinery and Intelligence", A. M. Turing seems to say that "computer thought" is the same as human thought if a computer's type-written linguistic behavior could not be distinguished from a human's type-written linguistic behavior: the "Turing test". On this view, questions about the similarity of the internal mental states of humans and the (possible) internal states of computers are to be answered on the basis of external behavior of both kinds of being. We know humans have internal mental states; Turing seems to say that computers have internal mental states if their external behavior is sufficiently similar to human external behavior. Turing apparently gives us a criterion for both the reality of internal mental states of computers and for the similarity of these states to human mental states.

On the contrary, Turing actually rejects the question about the possible reality of computer thought: "The original question, 'Can machines think?' I believe to be too meaningless to deserve discussion" (Turing, p. 57). However he further believed that in the near future (for Turing, within 50 years of 1950) it would be possible to construct machines which could *imitate* type-written human language interaction to such a degree that "an average interrogator will not have more than 70 percent chance" (Turing, p. 57) of correctly distinguishing language generated by a computer from that generated by a human. As a consequence "the use of words and general educated opinion will have altered so much that one will be able to speak of machines thinking without expecting to be contradicted" (Turing, p. 57). With this shift in the use of words comes a clear answer to the question of computer thought: computer's do think.

Turing sees the relationship between our talk about computer thought and our interventions with computers historically. With the emergence of a new technology (new ways of doing) comes new ways of speaking; in particular we *will* find (says Turing) that it makes sense to speak of computer thought. Questions about realism involve issues about the change through time of how we use words to talk about what we do. We must recognize that both what we say and what we do changes.

It is possible to sort out the transformation from phlogiston to hydrogen. Since phlogiston is real–as Watt's interventions with phlogiston show–one expects and can find a sensible history from "phlogiston" to "hydrogen." Similarly, it is possible to imagine Turing's prophecy come true: talk of computer thought would become more sensible as we built better computers. Computer thought–internal mental states for computers–would be real. It is also possible to imagine Turing's prophecy failing:

despite repeated attempts, computers never do behave similarly enough to humans, that we feel comfortable speaking of computer thought. Although, no doubt, we would describe the computer somehow, talk of computer thought would disappear. Computer *thought* would not be real. Nothing replaced either; talk of either simply disappeared.

5. The Instruments-Do-Not-Always-Work-Right Thesis

Instruments establish the phenomenal world of the scientist. This much is what is right about the logical positivist's interest in meter readings. However, instruments do not simply create their output from whole cloth; with them we interact with the real world. There are two kinds of reasons for this.

The first is the Intervening Thesis at work: When we have an instrument working properly our ability to do things with the instrument establishes a relation between us and the real that is not different in kind from the relation we have to any of the less controversial parts of the real world. A reliable instrument allows us to produce and manipulate phenomena. We similarly interact with rose bushes.

The second follows from the Instruments-Do-Not-Always-Work-Right Thesis: In the construction of an instrument, we do *not* always get just what we expect; this fact insures that when we *do* get an instrument to work properly our success is not simply a social construction. We think the objects we study with our instruments are real because we are not "the only ones doing the talking."

The development of the cyclotron during the 1930s provides a good example. The basic idea behind a cyclotron is to use a negative charge in electric potential to accelerate a positive ion. By forcing the ion to follow a spiral path with an electromagnet, the same potential difference is used repeatedly to accelerate the ion to higher and higher energies. The simple physical trick which convinced E. O. Lawrence that a cyclotron would be feasible, followed from a simple calculation he performed in 1929: ions in a uniform magnetic field revolve with a fixed frequency; the faster they go, the wider the orbit (Lawrence 1951, p. 431). Thus a fixed frequency power oscillator could be used repeatedly to accelerate an ion spirally out from a central ion source.

For this reason, Lawrence initially conceived of the cyclotron with a perfectly uniform magnetic field. Yet when M. S. Livingston began building cyclotrons he found that magnetic shims substantially improved their operation. Initially, Livingston and Lawrence thought the shims corrected irregularities in what supposed to be a uniform magnetic field. In a literal process of cutting and trying, Livingston added shims "to compensate for these irregularities." Fortunately he checked his progress by looking directly at the operation of the cyclotron. As they subsequently found out, the point of the shims is to create a magnetic field which decreases in intensity at larger radii; this helps to focus the ion beam (See Baird and Faust 1988, and Livingston 1969 for more details.).

There is a predominantly empirical constraint that guides construction. A phenomenon is sought. The instrument must be *reliable*; it must produce appropriately similar effects with each use. This output should be as free from interpretive ambiguity as possible. Ideally we should be able to control nature to a certain degree with the instrument; we should be able to anticipate outputs for a relatively wide variety of inputs to the instrument.

These constraints are, in part, social constraints, for it is the group of practicing scientists which collectively judges similarity of inputs and outputs and which determines whether the results are ambiguous. But importantly these constraints are empirical. Lawrence and Livingston would have been happier had their instrument not required magnetic shimming; *nature did not comply*. Thus, while people judge the similarity of results, it is nature which produces these (similar) results. We tinker with our nascent instruments to get what we want from nature. But, the fact that we need to tinker, shows us that we are tussling with a bit of the real world.

6. The Tinkering Thesis

Much of the work that goes into getting a new instrument running properly is tinkering. A good idea might suggest a design, but getting all the bugs out of the device frequently is a cut and try process. This is the Tinkering Thesis, and it argues against a more generous *theoretical* realism.

In broad terms instrument creation is a process of emulation, adaptation and tinkering. Lawrence adapted an idea of R. Wideroe's (1928) for a linear accelerator. The high frequency power source was adapted from amateur radio transmitters (McMillan 1959, pp. 668-70). New techniques were developed by tinkering; for example, R. R. Wilson devised a simple, but useful, way to communicate linear and rotary motion across the vacuum barrier (Wilson 1941).

These techniques are collected and preserved in books such as John H. Moore *et al.*'s *Building Scientific Apparatus: A Practical Guide to Design and Construction* (1983). For example, Moore *et. al.* discuss a problem that arises with metal bearings in vacuum systems; it seems that the bearings become rough very quickly:

> The tendency for a bearing to gall is reduced if the two mating bearing surfaces are made of different metals. For example, a steel shaft rotating without lubrication in a brass or bronze journal will hold up better than in a steel bushing. A solid lubricant may be applied to one of the bearing surfaces. Silver, lead-indium, and molybdenum disulfide have been used for this purpose. Graphite does not lubricate in a vacuum. MoS_2 is probably best. The lubricant should be burnished into the bearing surface. The part to be lubricated is placed in a lathe. As the part turns, the lubricant is applied and rubbed into the surface with the rounded end of a hardwood stick. By this means, the lubricant is forced into the pores. After burnishing, the surface should be wiped free of loose lubricant (Moore et al. 1983, p. 90).

I could carry on with other examples. The point is that instruments are created by emulating and adapting proven techniques. There is no general theory of instrument construction; there is only a grab-bag of tried and true techniques.

Each successful instrument presents a phenomenon; each provides us with one connection with the real world. But, the very means that promotes our success at creating new instruments constrains what we see of the real world. In the first place, this is true because we demand a phenomenon of an instrument; this is what gives us confidence that the instrument is working properly–making contact with the real world. But much of nature is complex; it may not be reducible to a series of individual, instrumentally controlled, phenomena. Secondly, instruments are created by emulation, adaptation and tinkering. Thus the sample of phenomena we access with our instruments is a *convenience* sample; there is no reason to suppose it is representative of all the phenomena nature has to offer. Convenience samples are biased, and inductions based on biased samples are not to be trusted. We should be shy to take any claim a theory makes that substantially generalizes away from our instrumental practice too realistically.

Thus Instrumental Realism emphasizes our ability to manipulate and control aspects of nature and not our ability to say true things about nature. When we reach the stage where we can intervene in some bit of the world to produce reasonably consistent and reliable results then we know we have "made contact." Our actions touch on structures or processes which exist independently of us; we are engaged in commerce with the public, objective–real–world.

But if Instrumental Realism does not emphasize the meaning of electron (or computer thought, or latent heat, etc.) then *what* is real when we successfully build an instrument which can intervene with "electrons?" The answer, that it is *electrons* that are real, can mislead. If we make contact with our instruments, and we describe what we are doing with a word like "electron," then "electron" refers–trivially so. But we may get everything wrong about electrons–what their properties are, even whether they are an entity or a property of entities. Still our ability to produce a publicly accessible phenomenon shows that we have made contact. This is Instrumental Realism. We then have the charge of learning more about this bit of nature we are intervening with. As we do learn more, the way we talk about the part of the world will change, but not so radically that we cannot follow the changes and keep sense in our talk.

References

Baird, D. (1988). "Instruments on the Cusp of Science and Technology: The Indicator Diagram." *Knowledge and Society*. Forthcoming.

Baird, D. and Faust, T. (1988). "Scientific Instruments, Scientific Progress and the Cyclotron." Unpublished manuscript.

Conant, J.B. and Nash, L.K. eds. (1957). *Harvard Case Studies in Experimental Science*, Volume 1. Cambridge, Ma: Harvard University Press.

Dennett, D. and Hofstadter, D. eds. (1981). *The Mind's I*. New York: Bantam Books.

Hacking, Ian. (1983). *Representing and Intervening*. Cambridge, England: Cambridge University Press.

Lawrence, E.O. (1951). "The Evolution of the Cyclotron." Nobel Lecture, December 11, 1951. In Nobel Lectures. 1965. *Nobel Lectures for Physics: 1922-1941*. Amsterdam: Elsevier Publishing Company.

Livingston, M.S. (1969). *Particle Accelerators: A Brief History*. Cambridge, Ma: Harvard University Press.

McMillan, E.M. (1959). "Particle Accelerators." *Experimental Nuclear Physics*, Volume III. Edited by E. Segre. Pages 639-785. New York: John Wiley & Sons.

Moore, J.H., Davis, C.C. and Coplan, M.A. (1983). *Building Scientific Apparatus: A Practical Guide to Design and Construction*. London: Addison-Wesley Publishing Company, Advanced Book Program/World Science Division.

Muirhead, J.P. (1859). *The Life of James Watt with Selections from his Correspondence*. London: John Murray.

Putnam, H. (1979). "The Meaning of 'Meaning'." Reprinted in *Mind, Language and Reality: Philosophical Papers*, Volume 2. Cambridge, England: Cambridge University Press.

Searle, J. (1980). "Minds, Brains and Progress." *The Behavioral and Brain Sciences*. Reprinted in Dennett and Hofstader, 1981. Page references are to that volume.

Turing, A. (1950). "Computing Machinery and Intelligence." *Mind*. Reprinted in Dennett and Hofstadter, 1981. Page references are to that volume.

Wideroe, R. (1928). "*Ueber ein Neues Princip zur Herstellung hohen Spannugen* [New Principle in Generating High Voltages]." *Archive Elektrotechische*. Volume 21, pages 387-406.

Wilson, R.R. (1941). "A Vacuum-Tight Sliding Seal." *Review of Scientific Instruments*. Volume 12, pages 91-3.

Winograd, T. (1973). *Understanding Natural Language*. New York: Academic Press.

https://doi.org/10.1086/psaprocbienmeetp.1988.1.192982 Published online by Cambridge University Press