JOHN STACHEL

SCIENTIFIC DISCOVERIES AS HISTORICAL ARTIFACTS

In his recent book *Wonderful Life* (Gould 1989, pp. 277–291), Steven Jay Gould notes that Harvard now organizes the sciences "according to procedural style rather than conventional discipline [into] the experimental-predictive and the historical" (*ibid.*, p. 279). While the former, such as physics and chemistry, have often been taken as prototypes for all sciences, Gould emphasizes that:

> Historical explanations are distinct from conventional experimental results in many ways. The issue of verification by repetition does not arise because we are trying to account for the uniqueness of detail that cannot, both by the laws of probability and time's arrow of irreversibility, occur together again. We do not attempt to interpret the complex events of narrative by reducing them to simple consequences of natural law; historical events do not, of course, violate any general principles of matter and motion, but their occurrence lies in a realm of contingent detail. (The law of gravity tells us how an apple falls, but not why the apple fell at that moment, and why Newton happened to be sitting there, ripe for inspiration.) And the issue of prediction, a central ingredient in the stereotype, does not enter into a historical narrative. We can explain the event after it occurs, but contingency precludes its repetition, even from an identical starting point (*ibid.*, p. 278).

He then succinctly characterizes the nature of such explanations:

> Historical explanations take the form of narrative: *E*, the phenomenon to be explained, arose because *D* came before, preceded by *C*, *B*, and *A*. If any of these stages had not occurred, or had transpired in a different way, then *E* would not exist (or would be present in substantially altered form, *E'*, requiring a different explanation). Thus, *E* makes sense and can be explained rigorously as the outcome of *A* through *D*. But no law of nature enjoined *E*; any variant *E'* arising from an altered set of antecedents, would have been equally explicable, though massively different in form and effect (*ibid.*, p. 282).

Gould adds, perhaps a bit wistfully:

> When we have established 'just history' as the only complete and acceptable explanation for phenomena that everyone judges important – the evolution of the human intelligence, or of any self-conscious life on earth, for example – then we shall have won (*ibid.*, p. 283).

When applied to historical sciences that involve human consciousness and agency – to which I shall refer as the telic historical sciences for short – Gould's account must be supplemented by an observation that is quite consonant with his viewpoint.[1] Compared with the non-telic historical

Kostas Gavroglu et al. (eds.), Trends in the Historiography of Science, 139-148.

sciences, a new element enters into the telic ones, the element of intention. Conscious human activity (often called "practice") involves a project, the existence of a conscious goal or goals on the part of the participants in that action. Marx emphasized well over a century ago:

A spider conducts operations which resemble those of the weaver, and a bee would put many a human architect to shame by the construction of its honeycomb cells. But what distinguishes the worst of architects from the best of bees is that the architect builds the cell in his mind before he constructs it in wax (Marx 1976, p. 284).

The reference to inferior architects serves to reminds us that the product of any such goal-oriented human activity, which I shall refer to as an artifact,[2] does not always correspond to the initial intent, even when its production involves only a single agent. Apart from competence, the intent of even the best craftsman will often change in the course of a project. When more than one agent participates in an activity, the goals of some participants may initiate the action, while the goals of others may be defined in the course of their reactions. Negotiations between participants with congruent goals and clashes between participants with incompatible goals assure that the artifact produced hardly ever coincides with the original goal of any single participant. While talk of goals in the non-telic historical sciences is quite misplaced, the element of human intention, of project, of goal should never be ignored in discussing the telic ones.

As noted, goals carry the inherent possibility that they will be modified – perhaps drastically – in the course of an activity, or even abandoned. If not abandoned, they may remain forever unfulfilled because unfulfillable. The objective world confronts humanity with a manifold of potentialities for activity. Clearly, if some project is not in accord with any of these potentialities, no amount of effort will produce the desired artifact – nor can it ever arise serendipitously in the course of striving for other goals. For example, if our present understanding of gravitation is at all adequate – even qualitatively – all projects to construct an anti-gravity machine will fail.[3]

You may have become aware by now that one of *my* goals in this paper is to try to develop a vocabulary for talking about human projects that does not imply that the outcome of such a project is *foreordained* by "objective reality," on the one hand; nor on the other, that "anything goes" – that there are *no* objective constraints on such projects. Whether *my* goal will have to be drastically modified or abandoned depends upon the utility of such artifacts as this paper.

To return to the thread of my argument, which of the densely interwoven web of potentialities will actually be realized as a concrete historical artifact or sequence of artifacts depends upon numerous contingent factors – factors that are not accidental in some absolute sense of that word, but contingent in the sense that the production of these particular artifacts could not be predicted solely on the basis of a knowledge of some initial state and of all the potentialities inherent in it. Indeed, what *some* of the potentialities were often is known only, after the event, and it is doubtful if *all* of them can ever be known; so perhaps it is better to say that the production of artifact, an actualized potentiality, will always appear to depend on factors, the occurrence or non-occurrence of which can be equally well conceived.[4]

What does all this have to do with my topic: "Scientific discoveries as historical artifacts"? If we agree that the history of science is a telic historical science, then all scientific discoveries must be regarded as historical artifacts and studied by the appropriate methods. Perhaps it is not necessary to remind historians of science of this truism, but it is often necessary to remind scientists and philosophers of science that attempts to find a "logic of discovery," a methodology that would fit discoveries into the experimental-predictive mold are fundamentally misguided.

On the other hand – and perhaps here even the historians need a nudge – this observation does not imply that discoveries are not amenable to rational analysis. As applied to historical artifacts, such an analysis must employ the methods of the telic historical sciences. As Gould puts it, patterns of explanation must "take the form of narrative." After the fact, one can make sense of the production of a scientific discovery; its emergence can in principle[5] be rigorously understood as the outcome of some antecedent sequence of events, a sequence that can be causally understood.[6] But if any of these antecedent events had not occurred, or if the events that make up some stage in the process had transpired in a different way, then the event in question would not have occurred, or would have taken a substantially altered form, requiring a different explanation.[7] One must take seriously the lesson of all historical sciences: if we are really constructing a narrative and not a morality play, contingent factors enter in an essential way into the construction of every historical narrative.

How should one go about such a construction? I suggest that we should *not* start from the assumption that there is an ideal world of facts, theories, devices, or what have you, just waiting to be taken off the Platonic shelf ("discovered") by some uniquely creative soul(s). Any attempt to flatten out the rich manifold of intellectual, social, and instutional elements, many of them contingent, that enter into particular discoveries by attributing the entire process to individual creativity uncovering pre-existent truth risks ending in tautology and platitude: A creative individual is one who has made a great discovery, while a great discovery is one made by a truly creative individual.

In thinking about how to organize the manifold elements that enter into the production of a scientific discovery, I have been helped by the definition of creativity given by Mihaly Czikszentmihalyi.[8] He raises the question not *what* but *where* is creativity.

> All of the definitions ... of which I am aware assume that the phenomenon exists ... either inside the person or in the work produced ... After studying creativity for almost a quarter of a century, I have come to the reluctant conclusion that this is not the case. We cannot study creativity by isolating individuals and their works from the social and historical milieu in which their actions are carried out. This is because what we call creative is never the result of individual action alone; it is the product of three main shaping forces: a set of social institutions or *field*, that selects from the variations produced by individuals those that are worth preserving; a stable cultural *domain* that will preserve and transmit the selected new ideas or forms to the following generations; and finally the *individual*, who brings about some change in the domain, a change that the field will consider to be creative (Czikszentmihalyi 1988).

By its very nature, then, creativity involves public activity. We might contrast it with talent, the individual capacity to produce artifacts of some type. The act of production of the artifact does not per se constitute part of a creative process. The artifact has to fall within (or in extreme cases initiate) some cultural domain, and receive a positive appraisal by the appropriate audience for the field in question. Of course, more than one person may participate in the production of an artifact; and it is by no means guaranteed that the appropriate individual(s) will receive recognition for their role in its creation. In a somewhat trivialized nutshell, creativity involves talented individuals producing artifacts that are ultimately integrated into some cultural domain by the socially dominant arbiters of that field.

Analysis of a scientific discovery as a historical artifact, then, must involve a discussion of the *goals* of the participants in the discovery. But these goals should not be taken as given. The analysis should consider such questions as:

How the goals of each participant arose in a definite personal, social, and institutional context that reflects the state of the domain and field as it filters through to the participants.

How these individuals went about trying to realize their goals. This involves a process of navigation among the potentialities proferred by nature, some of them already charted but many of them still uncharted – indeed, the object of the search may involve the creation of new charts – in the domain of research; a process of navigation that utilizes the resources – intellectual, instrumental, financial, etc. – of the field that are available to the researchers, i.e., the social and institutional setting of their research.

How each individual's goals were realized, modified (possibly drastically), or abandoned, as the result of successes achieved or obstacles met in the course of this process of navigation; or in the simultaneous or subsequent process of negotiation with others in the domain having congruent goals, and of conflict with still others having opposing or competing goals.

How, out of this complex process of negotiation, there arose a concensus within the domain (or possibly several complementary or even competing sub-concensuses) that defined the nature (or natures, for there are often several more-or-less accepted variants) of the scientific discovery as it is finally accepted into the field.[9] As emphasized above, this outcome (or these outcomes) will be amenable to causal explanation; but it can hardly be considered unique or inevitable: One can often imagine alternate narratives that would have resulted in a different definition of the discovery that is finally incorporated into the field.

Several comments might be added about the nature of such a program for analysing discoveries in the natural sciences as historical artifacts. First of all, it should be clear from what I said that emphasis on goals does not imply a pure methodological individualism; the origins of individual goals would be explained as a result of an interaction of individual temperament and circumstance with the social milieu. Secondly, it does not imply a purely social-constructivist explanation of discoveries. If a particular goal is not in accord with *any* potentialities inherent in the natural world, *no* amount of goal-oriented behavior – no matter how socially reinforced it may be – will ever produce the desired artifact. (To return to my earlier example, the U.S. Air Force was once willing to put a lot of money into anti-gravity devices, but to no avail.)

Finally, such a program does not imply a purely aleatory *or* a purely deterministic explanation of discovery. If their particular social and institutional setting continually motivates a large number of individuals to strive for a cluster of related goals, one or more of which is in accord with some natural potentialities, then the probability is rather high that before too long, one or more of the goals will be realized as a socially useful artifact, in one form or another. The aura of inevitability surrounding many scientific or technological discoveries, often reinforced by their independent occurence several times, can usually be resolved into such a constonance of social motivation conjoined with the inherent feasibility of the resulting project(s).[10]

II. THE EXAMPLE OF RELATIVITY

I believe that the development of special relativity (SRT) and general relativity (GRT) can be treated with advantage using the approach just indicated. The following brief sketch, drawn primarily from material I have presented in more detail elsewhere, still has to be filled out by considerable further research before making any claim to constitute an account that is adequate when judged by the standards suggested above.

SRT arose from quite a different goal than the one finally attained. Einstein had originally been searching for an electrodynamics of moving bodies that was compatible with all the relevant experimental results, particularly those from optics. This search itself formed part of his wider efforts to give a constructive account of the nature of matter and radiation. The work of Lorentz, Poincaré, Langevin, and others suggests that, if Einstein had not made his contribution, the formalism of special relativity might have been assimilated by the physics community in a context that made no such sharp distinction, as Einstein eventually did, between kinematical and dynamical concepts, a context that probably still would have included the concept of the ether. The shift from the goal of solving problems of electrodynamics to that of a reconstruction of the kinematical foundations of all physics cost Einstein seven years of effort. Even so, in the light postulate, his version of SRT still bears the birthmarks of its origins in electrodynamics, birthmarks that have almost universally been incorporated as an integral element in accounts of the theory. Had Einstein seen clearly in 1905 that he was faced with a purely kinematical problem, he might have realized something that he did not appreciate even after it was pointed out by Ignatowsky in 1909–1910: The light postulate is not needed for a derivation of the Lorentz transformations. The principle of relativity plus suitable assumptions about the homogeneity and isotropy of space and time (assumptions that Einstein also used implicitly), suffice to derive an unique family of kinematic transformations: the relativistic formulae with an arbitrary parameter taking the place of the speed of light.[11] Had Einstein, or the other leaders of the scientific field or community who determined the nature of Einstein's discovery, rejected this final non-kinematical element of the theory, we would have a somewhat different theory than the version today accepted as canonical.

Instead, the transition from Einstein's designation of his work on the relativity *principle* to the ~~domain~~'s characterization of his "discovery" as field's the *theory* of relativity, and the acceptance of this "theory" by the scientific community, involved a process of negotiation among such augurs of the German-speaking theoretical physics community as Planck, Lorentz, Minkowski, and Ehrenfest, a process that rejected Ignatowski's insight and accepted a certain reading of Einstein's accomplishment. In many cases, no clear distinction was made between Lorentz's "constructive approach" and Einstein's "principle approach" (see the studies of the reception of SR by various national physics communities by Stanley Goldberg and others). What Einstein set out to do and what the community decided he had accomplished were in many respects rather different things.

In contrast to the story of SRT, as early as 1907 Einstein rather clearly formulated the intertwined goals of generalizing the relativity principle beyond the Lorentz transformation group, and of inventing (his favourite term) a relativistic theory of gravitation. While almost all other physicists rejected the first problem, a number of prominent ones worked on the second, but they did so within the confines of SRT. Work on the problem of formulating a special-relativistic theory of gravitation did not stop even after the augers of physics agreed that Einstein had found the "correct" relativistic theory of gravitation by combining the two problems, with the resultant shift to the search for a geometrical, non-flat space-time theory of gravitation. Even so, it was only a series of "lucky accidents" and "fruitful errors" that diverted him from exploring special-relativistic scalar and tensor theories of gravitation that would have been quite compatible with his outlook on the gravitational problem in 1907.

Indeed, after the formulation of special-relativistic quantum field theories, there was an upsurge of interest in such flatspace theories of gravitation because Einstein's non-linear gravitational theory proved so resistant to quantization (it remains so to this day). Suppose Einstein had not been around to give the crucial 'geometrical turn' to the gravitational problem long before the advent of quantum field theory. Could some version of GRT have been developed out of a special-relativistic quantum field-theoretical approach? Feyman actually raised this question in 1957: "Suppose Einstein never existed, and the theory [GRT] was not available" (p. 151). He proceeded to show how a non-linear gravitational theory *formally* identical to GRT could have been developed as the theory of a self-interacting spin-two field in Minkowski space-time – which is just the way most elementary particle theorists insist on viewing Einstein's theory to this day. Of course, as concerns its *conceptual* structure, the resultant theory is *not* equivalent to Einstein's GRT, and this is just my point. Had history taken a different course, we might have had a different version of a non-linear gravitation theory that would not have been a "general relativity theory."

I believe that construction of alternative historical scenarios, such as the ones I have tried to sketch out here for SRT and GRT, serves a valuable purpose in reminding us of the contingent and constructive nature of many features of a scientific "discovery" that, with hindsight, we have become accustomed to regard as inevitable and natural. I emphasize again that the approach sketched does *not* suggest that such contingencies are without their causes, but only that the occurrence of these factors should not be regarded as inevitable. In short, such an approach turns our attention away from attempts to find some sort of *prospective* method of creativity or logic of discovery towards *retrospective* attempts to understand scientific discoveries as historical artifacts.

I have found no discussion by Einstein of the role of social or institutional factors in the development of new scientific ideas. But it is interesting that, rather than "discovery" [*Entdeckung*] (which implies previous existence), or "creation" [*Schaffung*] (which implies complete human control) as the best metaphor for the process that results in a new idea, he seems to have preferred "invention" [*Erfindung*], which implies a process in which human agency acts in accord with something outside of full human control.[12] This preference for the term "invention" rather than "discovery" suggests he was well aware of the role of contingent personal factors in the emergence of such theories. I shall end with another quotation that also suggests he might have favored a retrospective approach:

> A new idea comes suddenly and in a rather intuitive way. That means that it is not reached by conscious logical conclusions. But thinking it through afterwards you can always discover the reasons which have led you unconsciously to your guess and you will find a logical way to justify it.[13]

Boston University

NOTES

[1] I am *not* asserting that *all* sciences involving human agency must be historical. However, the history of science certainly is, so the question of whether there are human experimental-predictive sciences need not be considered here.

[2] I have borrowed this term from Marx Wartosky. (See Wartofsky 1979.) An artifact need not be a material object: a dance, a song, the Greek state, a historical event, are artifacts. On the other hand, a naturally produced object may become an artifact without any physical modification by virtue of its cultural role, e.g., sacred rocks or pools.

[3] An article appeared in the *New York Times* of December 28, 1989, reporting a claim by a two Japanese scientists to have created a weight loss by spinning an object. While this "astounding claim" is "almost certainly wrong," it serves as a salutary reminder of how tentative must be all claims to have exhaustive knowledge of natural potentialities.

[4] This does not imply that they are equally probable, just that the probability for neither appears to be zero.

[5] I say "in principle" because lack of all the relevant historical information (which *could* have been known if adequate records had been kept) usually prevents anything approaching such a complete reconstruction.

[6] Here, the distinction between causality and determinism is important. As I have discussed elsewhere (Stachel 1969), the concept of determinism refers to closed systems, for which a knowledge of the state of the system at any time would permit the prediction of the future (and retrodiction of the past) states of the system at any other time. The concept of causality refers to open systems, for which the effects of external intervention on the system can be lawfully explained. Hence, causal explanation never implies inevitability, since by definition the external intervention is unpredictable on the basis of a complete knowledge of the state of an open system. Universal determinism is the dogma that any open system can be closed by sufficiently enlarging it.

[7] Note that I am here just paraphrasing Gould's account of historical explanation, quoted above, and attempting to apply it to the process of scientific discovery.

[8] The relevance of Cszikszentmihalyi's work was brought to my attention by Howard Gardner (unpublished talk at the Osgood Hill meeting on "Einstein: The Early Years," October 1990).

[9] Here the work of Augustine Brannigan on *The Social Basis of Scientific Discoveries* (Cambridge University Press, 1981) should be incorporated into the program – and will be when this work is elaborated.

[10] For an excellent analysis of the invention of television from this point of view, see Williams 1974. I am much indebted to the ideas of Raymond Williams for my approach to these problems.

[11] Galilean kinematics is just the limiting case of this family when the parameter becomes infinite.

[12] According to Alexander Moszkowski, who reported in Moszkowski 1921 on his extensive conversations with Einstein in 1919–1920: "At first I was almost dumbfounded to hear Einstein say: The expression 'discovery' is itself to be deprecated. For discovery is equivalent to becoming aware of a thing which is already completely formed ... Discovery is not really a creative act!' ... Einstein supplemented this by emphasizing the concept of 'invention,' and ascribed a considerable role to it" (Moszkowski 1921, pp. 100–101).

[13] Einstein to Dr. H. L. Gordon, 3 May 1949 (Item 58–217 in the Control Index to the Einstein Archive).

REFERENCES

Brannigan, Augustine, 1981, *The Social Basis of Scientific Discoveries* (Cambridge: Cambridge University Press).

Czikszentmihalyi, Mihalyi, 1988, "Society, culture and person: a systems view of creativity," in Robert J. Sternberg, ed., *The Nature of Creativity* (Cambridge: Cambridge University Press).

Richard P. Feynman, 1957, "Critical Comments," in *Conference on the Role of Gravitation in Physics.* Wright-Patterson Air Force Base: Wright Air Development Center Technical Report 57-216, pp. 149-153.

Gould, Stephen Jay, 1989, *Wonderful Life: The Burgess Shale and the Nature of History* (New York/London: W.W. Norton).

Marx, Karl, 1976, *Capital: Volume One*, transl. by Ben Fowkes (Hammondsworth: Penguin).

Moszkowski, Alexander, 1921, *Einstein the Searcher: His Work Explained From Dialogues With Einstein,* transl. by Henry L. Brose (New York: E.P. Dutton).

Stachel, John, 1969, "Comments on 'Causality Requirements and the Theory of Relativity,' " in R.S. Cohen and M.W. Wartofsky, eds. *Boston Studies in the Philosophy of Science* vol. 5 (Boston: Reidel) pp. 179-197.

Wartofsky, Marx W., 1979, *Models/Representation and Scientific Understanding* (Dordrecht/Boston: D. Reidel).

Williams, Raymond, 1974, *Television: Technology and Cultural Form* (London: Fontana/Collins).