Multiple discoveries in science: a test of the communication theory*

Augustine Brannigan
Richard A. Wanner

Abstract. While social determinists have argued that the appearance of multiple, independent innovations in science and technology demonstrates the inevitability of cultural maturation, contemporary psychological approaches view the patterns as a function of the chance interaction of scientists and the environment. By contrast, it is argued in this paper that multiple discoveries can be accounted for by delays in communication in science, and that these have declined with the growth of scientific institutions. Using a record of multiples covering the last four centuries, we show that the mean number of scientists involved in multiples has been declining, and that the time interval separating independent reports has been approaching zero. Lastly, it is argued that the communication framework refutes the chance model but is consistent in part with the maturational approach

Résumé Pendant que les déterministes discutent que des innovations multiples et indépendantes dans le domaine des sciences et de la technologie démontrent qu'inévitablement il y aura un développement culturel, les approches psychologiques contemporaines, elles, voient le modèle comme une fonction de l'action réciproque des scientistes et de l'environnement. Dans cet article, par contraste, on tente de prouver que de nombreuses découvertes peuvent être expliquées par des délais dans les communications scientifiques et que celles-ci diminuent à mesure que les institutions scientifiques grandissent. En utilisant des dossiers portant sur ces multiples et qui couvrent les quatre (4) derniers siècles, on se rend compte que le nombre moyen de scientistes, impliqués dans les multiples, déclinent et que l'espace de temps séparant les rapports indépendants est presque nul. Enfin on y discute que le domaine des communications nie le facteur hazard mais demeure cependant consistant, en partie, avec la manière d'aborder le développement.

---

* This research was partially supported by an operating grant from the University Research Policy and Grants Committee of the University of Calgary. We wish to thank Dean K. Simonton for providing us with his list of multiple discoveries and their characteristics for this project. The authors' names are in alphabetical order Please address all correspondence and reprint requests to Professor Augustine Brannigan or Richard A. Wanner, Department of Sociology, University of Calgary, Calgary, Alberta T2N 1N4.

Multiple discoveries have come to occupy a curious if not strategic place in the modern sociology of science. Like children born out of wedlock, everyone agrees they should be attended to, but no one knows quite what to do with them. As with lineage, the problem is determining what they mean to whom. Although Merton (1973) must be credited, if not with creating them, with elevating multiples to the position of a "strategic research site" in the sociology of science, the reason they muster interest must be sought in the claims made for them by those who have attempted to explain the phenomenon of scientific discovery. Specifically, for cultural determinists who trace their intellectual ancestry to Kroeber (1917), multiple discoveries provide evidence for the inevitable maturational development of human societies. This view claims that incremental advances that occur in science over time vastly enhance the probability of further advances. Consequently, individuals working in the same area are likely to face duplication of their efforts as a result of the saturation of a culture with common problems, resources and solutions.

In contrast, for advocates of the "chance" model of scientific discovery (Price, 1963; Simonton, 1978; 1979), the existence of multiple discoveries is simply explained: they are a random outcome of the interaction between the natural environment and the scientific community. In this view, the simple interaction of discoveries and scientific discoverers will produce, on one end of the continuum, a situation where most prospective discoveries are not uncovered and, on the other, a situation in which an ever-diminishing group are discovered independently several times over. Consequently, Price and Simonton attach great importance to the apparent Poisson distribution of multiple discoveries according to their grade, i.e., number of scientists making the same discovery. This observed j-shaped distribution suggests that most possible discoveries are never made, the majority of those which are made are discovered by one scientist, a smaller number by two scientists, an even smaller number by three, and so forth. The implied model is that this duplication of effort is produced by the chance interaction between a fixed number of scientists and a fixed number of discoveries.

While intuitively appealing, both of these models have received some critical attention. It has been noted (Brannigan, 1980) that the theory of cultural maturation, which is in part based on the existence of multiples, becomes logically misleading where the only evidence for maturation is the occurrence of redundant discoveries.[1] This logical problem is found uniformly in the work of Kroeber (1917), Ogburn and Thomas (1922) and Merton (1973). As for the chance model, Brannigan, Wanner and White (1981a; 1981b) have argued that, as evidence for the theory, the Poisson distribution

---

[1] Additionally, it has been speculated that the social theory of the inevitability of multiples has its basis in a folk theory of social change (Brannigan, 1981: 143ff.).

of grade is insufficient for a number of reasons. The mathematical assumption that the number of scientists (N) varies inversely with the probability of any one discovery being made (p) is not borne out historically. In fact N and p seem to vary directly: the probability of any one discovery being made appears to be greatest in contemporary times when the number of scientists is largest. Secondly, the distribution of multiples, if examined century by century, ceases to approximate the Poisson as the average grade decreases over time from 1500 to 1900. Lastly, the frequency of multiples is not the same in every discipline, as Simonton argues, but appears to be influenced by factors over and above those examined in the Poisson model. In fact, a chief assumption of this paper is that one cannot claim to have explained a phenomenon simply because the shape of its distribution has been identified. In the same way that one has not explained income inequality in a society merely because its distribution is noted to approximate the log normal, to observe that the grade of multiple is approximately distributed as the Poisson is not to explain the phenomenon. Multiple discoveries can be shown to be random phenomena only if it can be demonstrated that no other predetermined variables influence their characteristics.

In this paper we present a new theoretical perspective, based on the development of scientific communication, that explains not only the systematic relationship of grade of multiple discovery to several other characteristics of innovations and scientists, but furthers our understanding of the phenomenon of multiple discovery by highlighting another of their important, but neglected, characteristics, the time interval elapsing between the first and last instances of discoveries comprising a multiple. After elaborating this communication framework and specifying its implications for both grade and interval of multiples, we proceed to test it using a subset of Simonton's (1978; 1979) collection of multiples occurring between the sixteenth and twentieth centuries.

An alternative theory: communication and speed

A mathematician himself, Jacques Hadamard noted that mathematicians frequently monitor the work in which their colleagues are involved to avoid duplication. "After having started a certain set of questions and seeing that several authors had begun to follow the same line, I happened to drop it and to investigate something else" (1949: 132). Had Hadamard not seen that others were conducting similar research, it is not improbable that both he and they would have continued in their research and subsequently reported rival outcomes. The first datum which Merton (1973: 357) cited as evidence for the contention that all discoveries are in principle multiple discoveries was that published singletons often turn out to be multiples. This was illustrated with the examples of Cavendish and Gauss. Their original discoveries had been recorded in unpublished notebooks. Is it any wonder that others subsequently failed to take note of such achievements? Presumably, if such works had been published or communicated widely, the potential rivals would have worked in other directions and consequently the record of multiples would have been truncated significantly.

This was intimated by Ogburn in his work on social change. He stated, "if an invention has become once made and has become widely known, there is no occasion for a second invention" (1950: 86). Of course, the reason for this is that once an invention or discovery has become known, the appearance of any identical announcement will be treated as a replication or duplication whose identity as a discovery will be negated by comparison with the earlier unprecedented discovery or invention. However, the announcements of discoveries do not always meet with unqualified success. Merton (1973: 380) spoke of the possibility that an announcement could get "lost in the great information system" of science. Presumably an innovation could get lost or could fail to become widely known or could be announced well after the achievement had been made elsewhere, because it was not published or circulated (this was the case with Scheele, Cavendish, and Gauss); because it was a matter of political or industrial secrecy (the Manhattan Project); because it was politically or religiously repressed (Galileo in Rome, and Vavilov in Russia); because it was unavailable to others outside the language and/or cultural group in which it was recorded (this was the case until recently with ancient Phoenician stone writing found in North America); or because it was announced in such obscure language that potential readers were disentitled (as in the case of Copernicus, Semmelweis, and Galois). Numerous other examples could be presented to substantiate each point. Nonetheless the common thread uniting each case would be what might be called a failure in the communication of results. We might argue, therefore, that, contrary to Merton (1973), the history of science is not essentially a history of multiple discovery, but a history of poor communication. Consonant with this view is the observation that priority disputes, which are intrinsically associated with multiples, have declined in time with the rise of professional scientific establishments, establishments which are particularly efficient in organizing the communication of research outcomes.

Implications for the grade and interval of multiples

What the communication theory would suggest is that, while the number of discoveries may be increasing absolutely and relatively over time with the growth of science, the size of the average grade of multiple should decline as scientific institutions and their journals forestall possible duplication. In other words, over time each multiple event will involve increasingly fewer independent contributions. Secondly, because individual disciplines have varied in their patterns of cognitive growth and institutional development, they will vary in the degree to which they experience duplications, and this will change over time. In other words, the grade of multiple will be influenced by both time frame and discipline. Third, following the communication framework, we hypothesize that the interval or time elapsed between the first and last date of discoveries belonging to a multiple will be reduced over time with the growth of science. Whereas in Renaissance science independent multiple discoveries may have been separated by decades, this interval will contract to no more than a few months in contemporary science. And lastly, the interval will further differ across individual disciplines which, as they form over time, will further close the interval between the first and last date of a discovery. As with grade, the interval will vary by both time frame and discipline.

The various theories of multiples which we have discussed have different implications for change in the distribution of grade and interval over time. Specifically, for the chance model, since the only relevant parameters are N (number of discoverers) and p (probability of any one discovery), and since the time factor for the Poisson model is assumed to be a constant, this model implies that the mean of neither grade nor interval will vary over time, as shown in Figure 1a.

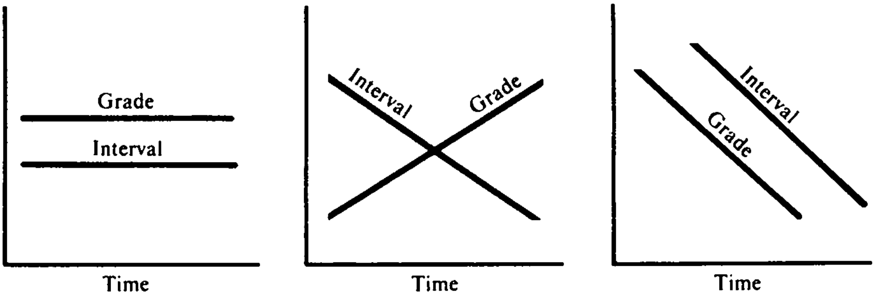
For its adherents, the culturological model would imply that with the growth of science, multiples should become increasingly likely. Consequently, the average grade of multiples increases with time. However, the selfsame growth of scientific institutions should contribute to a closing of the interval between the various breakthroughs. In other words, scientific growth fosters an increasing cross-fertilization of knowledge that is both more and more widely shared among members of the scientific community and of an increasingly competitive nature over time. Consequently, this model would predict that, over time, the average grade of multiples would increase while the average interval would decrease, as illustrated in Figure 1b.

Lastly, the communication framework implies that since duplication is least likely where the communication of results is least problematic, i.e., during periods of institutional and journal development, both the average grade of multiple, as well as the average interval, ought to decrease over time. This situation is shown in Figure 1c. The remainder of this paper is devoted to a test of these separate predictions regarding the patterns of grade and interval distributions over time, as well as to an exploration of the influence of other factors on grade and interval.

Data and measures

There have been various lists of multiple discoveries compiled to date. Kroeber (1917), Ogburn and Thomas (1922), and Merton (1973) have all developed lists.[2] The most exhaustive data set currently available is a collec-

Figure 1. Predictions for trend in mean grade of multiple discoveries and mean interval between instances from alternative theories.



a. Chance model b. Maturational model c. Communication model

tion of 579 multiples, including 1546 separate discoveries, gathered by Simonton. In the analysis to be performed here, we will treat the individual discovery as the unit of analysis and membership in a multiple of a particular grade or interval as a characteristic of the discovery. In this way it is possible to examine the effects of characteristics of both discoveries, such as the year of their occurrence, and scientists, such as their nationality, on grade and interval. In addition the analysis is facilitated by the large increase in the number of cases available.

The specific subsample of discoveries in the Simonton list to be used here excludes those that have been classified as false or disconfirmed; any multiples containing ancient discoveries as one of their components; all geographic and geological discoveries, the former because their scientific status is contentious, and the latter because there are too few of them; and any discoveries for which data is missing on any of the variables in the models developed below.[3] After excluding discoveries on these grounds, 1,235 remained

---

2. In 1961 Robert Merton reported that he and Elinor Barber were undertaking a detailed investigation of their sample of 264 multiples (see Merton, 1973: 364ff.). Merton reported coding discipline, historical period, the interval between the discoveries, the number of co-discoverers (grade), and nationality, among other features of the multiples in their list. The results of this analysis have never been fully published. We suspect that the technical limitations of statistical methods available to sociologists at the time, as well as the small sample size probably undermined the project.

3. Although we are largely ignoring them here, we recognize the enormous problems involved in assuring reliability and validity in the measurement of multiple discoveries. These difficulties include the role of nationalism in attributing a discovery to a scientist, the aggrandizement of disciplinary founders, diminishing recollection of original discoveries as time passes, whether or not discoveries deemed to be multiples are truly equivalent, and the quality and importance of individual entries. See Brannigan, Wanner and White (1981b) for a detailed discussion of these issues.

for analysis.

As indicated above, the two dependent variables in this study are the *grade* of multiple, measured as the number of recorded instances of the discovery, and the *time interval* between the first and last instance, measured as the number of years elapsing between the first and last recorded instance. It is likely that these two characteristics of multiples are not independent of one another, since higher grade multiples will likely have taken a longer period of time to occur. The major independent variables include time period, which will be effectively measured by the century in which the multiple occurred, discipline, nationality of the scientist making the discovery, whether it is a discovery or an invention, and whether it is theoretical or empirical in nature. The general disciplines to be included are mathematics, astronomy, physics, chemistry, biology, medicine, and technology. The coding of scientific innovations as either discoveries or inventions was considered important, since the latter, being recombinations of existing ideas or elements, would seem to be more likely to eventuate in multiples.[4] Basic discoveries, which are essentially new scientific ideas or elements, would in all probability be less obvious to workers in any specific field. A similar logic lies behind distinguishing between theoretical and empirical innovations. Theoretical work is thought to be less likely to be reproduced, since it is often less obvious to workers in a field than empirical verification research.[5]

A major problem to be faced in developing models to describe the effects of our independent variables on grade and interval stems from the fact that neither of these variables is normally distributed, thus precluding the use of ordinary least-squares (OLS) regression analysis. Since both of these variables are instead distributed as the Poisson, or at least approximately so, we can expect more variation around the mean for discoveries with a higher grade or interval than for discoveries with a low mean grade or interval. What this implies is that the conditional variances of these two variables will be unequal, violating the assumption of homoscedasticity underlying OLS regression. In addition, OLS regression also assumes that the errors of estimate are normally distributed, certainly not the case in our models predicting grade and interval. Instead, the error structures for these models are Poisson distributions.

Fortunately, a recently developed statistical analysis system, called Generalized Linear Interactive Modelling, or GLIM, is available to estimate

---

[4] We propose to maintain the important distinction between discoveries and inventions by referring to them collectively as "innovations."

[5] The designation of an innovation as a discovery or an invention or as theoretical or empirical was largely a subjective decision on our part. While we feel that we have been quite consistent in applying our operational definitions to the innovations in the sample, a number were so ambiguous that they were not coded and therefore deleted from the analysis.

such models with non-standard assumptions such as Poisson distributions of error and heteroscedasticity (Baker and Nelder, 1978). By making use of a maximum-likelihood estimation procedure along with the assumption that errors in prediction are distributed as the Poisson, GLIM allows us to estimate models of the form

$$\log G = \Sigma\beta \log C + \xi_1 \qquad (1)$$

in which G is grade — or interval — of multiple and C is characteristics of discoveries and scientists. Later in the analysis we shall also estimate some models that contain multiplicative terms capturing interactions among the independent variables, but they take the same general form as equation 1.

Findings

Grade of multiple discoveries

Table 1 presents five separate models, each expressing the zero-order effects of each of our independent variables on grade of multiple. Here, the coefficients express deviations from the mean of a baseline category, e.g. mathematics in the case of discipline. While in no case is more than 5 percent of the variance explained in grade, for all the independent variables but nationality of scientist the effects are statistically significant at the .01 level, indicating that the log of mean grade does differ across centuries, disciplines, and types of innovation. As the communication theory predicts, there is a fairly systematic decline in log grade from the sixteenth to the twentieth centuries, with only a slight reversal in the nineteenth century. The differences across disciplines show mathematics and biology with the highest means, while chemistry and astronomy have the lowest. Overall, inventions have a higher mean log grade than do discoveries, while the very slight difference between theoretical and empirical work shows the latter to have a higher mean. In the case of nationality of scientist, there is essentially no difference in mean grade.

Table 2 addresses the question of whether or not these effects of discipline and type of innovation vary over time. That is, are there interactions between these variables and century? This question is important for it adds weight to the relevance of advances in scientific communication as this is evidenced through time. In all cases we find that the model including interaction terms accounts for a significantly larger proportion of the variance in log grade than does its corresponding main-effects model. Surprisingly, this is even the case for the century-nationality model. Despite the fact that nationality had no zero-order effects on grade, such effects do appear across centuries. The strongest interaction effects appear in the century-discipline model where nearly 10 percent of additional variance in log grade is accounted for by the addition of the interaction terms. This is certainly

Table 1. Bivariate regressions of grade of multiple discovery (natural log) on characteristics of discoveries and scientists.

Independent variable	*Coefficient*[1]	*R^2***	*F-ratio*[2]	*d.f.*
Century		.041	13.15*	4/1230
16th century (constant)	1.328			
17th century	-.211			
18th century	-.346			
19th century	-.304			
20th century	-.469			
Discipline		.021	4.39*	6/1228
Mathematics (constant)	1.172			
Astronomy	-.259			
Physics	-.142			
Chemistry	-.243			
Biology	-.083			
Medicine	-.152			
Technology	-.153			
Discovery or invention		.049	63.52*	1/1233
Discovery (constant)	.952			
Invention	.200			
Theoretical or empirical		.010	12.45*	1/1233
Theoretical (constant)	1.095			
Empirical	-.096			
Nationality of scientist		.002	0.62	4/1230
Great Britain (constant)	1.037			
Germany	-.037			
France	-.020			
United States	-.019			
Other	-.009			

Note: All coefficients are maximum likelihood estimates calculated under the assumption that the error structure takes a poisson distribution. GLIM was used to compute these estimates. An approximate R^2 was estimated from the fitted model and total scaled deviance.

1 All coefficients are interpreted as deviations from the constant, i.e., as deviations from the mean of the reference category

2. Ratio of scaled deviances divided by their degrees of freedom. The tests shown here are of the null hypothesis that the mean grade across categories of the independent variable equal the grand mean

* $p < .01$

** Approximate R^2 estimated from the fitted model and total scaled deviance

consistent with the communication-modernization theory which allows that while all sciences will improve their communications over time, there will be significant variation across disciplines in the rate of development.

While the presence of these interaction effects is interesting, we must examine the patterns of means to determine their interpretability. The means

Table 2. Coefficients of determination, F-ratios, and degrees of freedom for models of grade of multiple discovery (natural logs).

Model	R^2**	*F-ratio*[1] *for increase in* R^2	*d.f.*
C and D main effects only	.055		
C and D main effects and interactions	.149	5.52*	24/1200
C and I main effects only	.083		
C and I main effects and interactions	.144	12.82*	4/1225
C and T main effects only	.047		
C and T main effects and interactions	.069	7.24*	4/1225
C and N main effects only	.042		
C and N main effects and interactions	.080	3.12*	16/1210

Note. Symbols are used as follows C, century, D, discipline, I, discovery or invention, T, theoretical or empirical work, N, nationality of scientist

[1] Based on the increase in scaled deviance of the interaction model over the main effects model

* $p < .01$

** Approximate R^2 estimated from the scaled deviance for the fitted model and total scaled deviance

Table 3. Mean grade of multiple discovery (natural log) expected from interaction models: century and discipline, century and discovery or invention, century and theoretical or empirical, and century and nationality.

			Century		
	16th	*17th*	*18th*	*19th*	*20th*
Discipline					
Mathematics	1.815	.938	.972	.978	*
Astronomy	*	.973	.780	.889	*
Physics	1.195	1.165	.913	.917	.808
Chemistry	*	*	.925	.838	.947
Biology	*	.828	1.020	1.018	.990
Medicine	.727	.943	.951	.978	.795
Technology	*	*	.905	.964	.828
Discovery or invention					
Discovery	.771	.914	.932	.911	.845
Invention	1.606	1.179	.910	1.010	.766
Theoretical or empirical					
Theoretical	1.449	1.009	.975	.996	.748
Empirical	.923	1.005	.908	.932	.843
Nationality					
Great Britain	1.069	.943	.988	.963	.805
Germany	1.362	1.078	.829	.926	.867
France	*	.932	.932	.925	.828
United States	*	*	.751	1.035	.804
Other	.967	1.061	.945	.934	.832

* Too few cases for stable estimates.

of log grade predicted by the interaction models are shown in Table 3.[6] In the case of the century-discipline panel, it is clear that the overall decline over time is maintained, though the rate of decline varies by discipline. In general, in earlier centuries there are large differences across disciplines which decrease up to the twentieth century. While a careful study of the development of communications systems in the various disciplines would be necessary to conclusively attribute these patterns to this factor, there is a definite difference between astronomy and medicine, on the one hand, and biology, on the other, that requires such investigation.

Differences in the types of innovation by century also show a pattern of steady decline and convergence over time. In the case of the discovery-invention contrast, however, there is a distinct reversal, with inventions having a much higher mean grade in earlier centuries, but discoveries acquiring a slight advantage by the twentieth century, though a t-test shows the difference to be non-significant. It is likely that the system of scientific communication embraced largely scientists working in basic research in earlier centuries, accounting for the higher rate of duplication among inventors. The small interaction effect of century with the theoretical-empirical contrast seen in Table 2 appears to be largely the result of a large difference in means in the sixteenth century. Afterward there remains no significant difference between the two types of work. Consequently, our initial supposition that empirical work is more likely to be independently reproduced than theoretical work is unfounded. The same is largely true of the century-nationality interaction. By the eighteenth century only small differences remain across countries.

Interval within multiples

The second characteristic of multiple discoveries or inventions that the communication theory predicts should vary systematically is the time interval between the first and last instance of a multiple. Table 4 lists a series of bivariate models for *interval* that is an exact parallel to those presented in Table 1 for *grade*. Here, the association between century and interval is considerably stronger than the association between century and grade, with over 23 percent of the variance in interval accounted for by century. In addition, the pattern of means is identical to that predicted by the communication theory: the mean grade falls steadily with the passage of time from about 86 years in the sixteenth century (the antilog of 4.457) to just over 2 years by the twentieth century.

---

[6] One might convert these predicted means of log grade into mean grade itself by taking the antilog of each value in the table. When this is done, the contrasts are certainly more striking, but we chose not to do so since the pattern of results remains exactly the same.

Table 4. Bivariate regressions of interval of multiple discovery (natural log) on characteristics of discoveries and scientists.

Independent variable	*Coefficient*	*R'*	*F-ratio*	*d.f*
Century		.235	94.46*	4/1230
16th century (constant)	4.457			
17th century	-0.546			
18th century	-1 445			
19th century	-2.240			
20th century	-3.624			
Discipline		.074	16.36*	6/1228
Mathematics (constant)	3.797			
Astronomy	-0.860			
Physics	-0.986			
Chemistry	-1.977			
Biology	-1.020			
Medicine	-0.976			
Technology	-1.634			
Discovery or invention		.009	11 19*	1/1233
Discovery (constant)	2.947			
Invention	-0.387			
Theoretical or empirical		.037	47.38*	1/1233
Theoretical (constant)	3.334			
Empirical	-0 754			
Nationality of scientist		057	18 59*	4/1230
Great Britain (constant)	2.889			
Germany	-0 451			
France	-0.347			
United States	-1.116			
Other	0.474			

Note All coefficients are maximum likelihood, estimates calculated under the assumption that the error structure takes a poisson distribution GLIM was used to compute these estimates See footnotes to Table 1 for descriptions of the measures reported here

*p< 01

In the case of discipline, mathematics exhibits by far the highest interval overall, while chemistry has the lowest, followed closely by technology. Astronomy, physics, and biology share approximately equal mean intervals. The typologies of innovation bear only tenuous relationships to interval, though the difference between their means is statistically significant. Nationality of scientist accounts for almost 6 percent of the variance in interval, with discoveries made by scientists in "other" countries having the highest mean, and discoveries in the United States the lowest. As we shall see, this is in part a spurious association, since the low American mean is largely the result of the lack of discoveries in that country during the sixteenth and seventeenth centuries.

Table 5 contains coefficients of determination and F-tests for the interval models which contain both main and interaction effects for century and the other characteristics. Here, by far the largest interaction effects take place between century and discipline, i.e., the disciplines clearly differ in their mean intervals over time. While the interactions between century and types of innovation are statistically significant, their magnitude is likely small, since they explain only a small proportion of additional variance in interval. In this case, the nationality-century interactions are effectively zero; the trend toward declining intervals does not vary across nationalities.

To interpret the patterns of these interaction effects, we must examine the means themselves, as shown in Table 6. Of special interest here is the maintenance of a large average interval in mathematics up to the twentieth century, though there were too few cases in the twentieth century itself to estimate a mean for this discipline. Secondly, the dramatic drop in mean interval in astronomy in the nineteenth century is consistent with the highly developed system of scholarly communication that had developed in this discipline up to that time. A similar dramatic drop takes place in physics in the twentieth century, such that there are no multiples with an interval of more than a year, represented by a mean interval of zero in Table 6. The inconsistently high mean for biology is likely the result of the re-discovery of Mende-

Table 5. Coefficients of determination, F-ratios, and degrees of freedom for models of interval of multiple discovery (natural logs).

Model	*R^2***	*F-ratio*[1] *for increase in R^2*	*d.f*
C and D main effects only	.262		
C and D main effects and interactions	.316	3.95*	24/1200
C and I main effects only	.246		
C and I main effects and interactions	.258	4.95*	4/1225
C and T main effects only	.255		
C and T main effects and interactions	.270	6.29*	4/1225
C and N main effects only	.246		
C and N main effects and interactions	.253	0.71	16/1210

Note: Symbols are used as follows: C, century, D, discipline; I, discovery or invention, T, theoretical or empirical work, N, nationality of scientist

[1] Based on the increase in scaled deviance of the interaction model over the main effects model

* $p < .01$

** Approximate R^2 estimated from the scaled deviance for the fitted model and total scaled deviance

Table 6. Mean interval of multiple discovery (natural log) expected from interaction models: century and discipline, century and discovery or invention, century and theoretical or empirical, and century and nationality.[1]

	Century				
	16th	*17th*	*18th*	*19th*	*20th*
Discipline					
Mathematics	3.701	1.918	3.935	1.853	*
Astronomy	*	1.498	1.591	.204	*
Physics	3.016	2.233	1.976	1.189	**
Chemistry	*	*	.851	.675	.484
Biology	*	1.815	1.582	.834	1.821
Medicine	3.376	2.645	1.923	1.491	.568
Technology	*	*	1.763	.873	.216
Discovery or invention					
Discovery	2.833	2.156	1.679	1.129	.484
Invention	3.413	2.009	2.007	1.267	.169
Theoretical or empirical					
Theoretical	3.598	2.029	2.430	1.330	.131
Empirical	2.769	2.141	1.513	1.132	.474
Nationality					
Great Britain	3.035	2.438	2.111	1.232	.445
Germany	2.943	2.202	2.035	.953	.489
France	*	2.136	1.233	1.162	.515
United States	*	*	1.584	1.342	.322
Other	3.042	1.761	1.633	1.360	.400

1 F-tests in Table 5 are for differences in these means

* Too few cases for stable estimates

** Interval less than one year for all physics discoveries in the 20th century

lian genetics in the early years of that century, though this was in some respects a dubious simultaneous discovery (see Brannigan, Wanner and White, 1981a).

Even though the interactions for century by type of innovation are statistically significant, the main pattern is one of convergence over time. The fairly large differences in the sixteenth century, favoring inventions and theoretical work, are considerably reduced by the seventeenth century, with some tendency toward divergence in the eighteenth. As indicated by the test in Table 5, we can safely assume that no important differences among nationalities exist over time.

Discussion and implications

We have observed a systematic decline over time in both the average grade of multiple discoveries and inventions and the average interval between the first and last instances of a multiple. These findings are clearly more consistent with the communication theory proposed here than they are with either the chance theory of Price and Simonton or the culturological model of Kroeber, Ogburn and Thomas, and Merton. The chance theory, based as it is on the assumption of an unvarying distribution of multiples, is particularly called into question, since both grade and interval seem to be influenced by disciplinary context and type of innovation, though only interval is affected by nationality of scientist.

While the temporal patterns in grade and interval observed here appear to support the communication theory, it has been difficult to interpret the effects of the other independent variables and their interactions with century. What is required is a detailed study of the development of communications media —journals and scholarly associations, as well as improvements in communications technology generally — to link them to the incidence of multiples and the time interval separating their instances. What this suggests is that the study of multiples cannot be separated from the study of the incidence of innovation more broadly conceived, though, as Merton (1973) intimated, the careful examination of multiple discoveries can provide important clues to the understanding of the distribution of scientific knowledge.

Another question raised by the data concerns the differences between changes in interval and grade. While both reveal patterns consistent with the explanation, the changes in interval are clearly more consistent with the predictions. With modernization, interval contractions are more uniform than the contraction for grade. Why is this? If the operative factor is the effective communication of results and if we presume that such developments are crudely a function of time, implying the evolution of knowledge, research opportunities, formal organization, etc., we should expect the interval contraction to be *more* a function of modern communication than is the decline in grade. While the increase in scientific careers and professions contributes to interval contraction, it tends to have the opposite effect on grade. Whereas the steady contraction in interval is influenced by the publication of results, the growth of the scientific population tends to increase the number of competitors, which contributes in turn to a slower contraction in grade. Consistent with this interpretation, we would expect to find that the variance of interval declines more rapidly over time than that of grade. That this is the case is clearly shown in Table 7. While the variance decreases for both over time, it decreases considerably more rapidly for interval.

These results tend to support the supposition that interval contraction is a better indicator of the process of scientific change than is grade contraction, because the population growth and development of scientific infrastructures, which reduce the delay in communication results, simultaneously create a situation in which more and more scientists compete for the same achievements and discoveries. Consequently, not only does mean grade de-

Table 7. Variances of grade and interval by century.

	Century				
Variable	*16th*	*17th*	*18th*	*19th*	*20th*
Grade	1.457	1.201	1.108	1.137	1.058
Interval	38.48	18.86	15 10	6.440	2.062

Note Reported variances were computed by taking antilogs of the variances of natural logarithms of grade and interval

cline more slowly than mean interval, but the variance of grade shrinks more slowly over time than does the variance of interval.

There is a final point which must be clarified regarding the communication theory and our assumptions about the development of the scientific community on which it is premised. Where we earlier argued that maturational or culturological theory was one we disputed, plainly we adopt the same assumption about the progress of scientific culture associated with this theory. However, where Kroeber argues that multiples become more inevitable with the growth of science, we caution that such growth ultimately precludes multiples. But we must concur that innovations in general — discoveries, inventions and conceptual advances — become more likely with the growth of scientific institutions and the proliferation of scientific careers and specialties. Like Kroeber's, our position is premised on a developmental model, though the empirical significance of multiples is substantially revised. Nonetheless, the orientations of the maturational and communication models are fundamentally complementary. Further research is needed to explain differences in the rates of growth of separate disciplines and the importance of national and organizational factors underlying these patterns.

References

Baker, R.J. and J.A. Nelder

1978 *The GLIM System: Generalized Linear Interactive Modelling*. Oxford: Numerical Algorithms Group.

Brannigan, A.

1979 "The reification of Mendel." *Social Studies of Science* 9: 423-54.

1980 "Naturalistic and sociological models of the problem of scientific discoveries." *British Journal of Sociology* 31: 559-73.

1981 *The Social Basis of Scientific Discoveries*. Cambridge: Cambridge University Press.

Brannigan, A., R.A. Wanner, and J.M. White

1981a "The phenomenon of multiple discoveries and the re-publication of Mendel's work in 1900." *Philosophy of the Social Sciences* 11: 263-76.

1981b "Explanations of scientific innovation and the frequency distribution of multiple discoveries." A paper presented at the annual meeting of the American Sociological Association, Toronto.

Hadamard, J.

1949 *The Psychology of Invention in the Mathematical Field.* Princeton. Princeton University Press.

Kroeber, A.L.

1917 "The superorganic." *American Anthropologist* 19: 163-214.

Merton, Robert K.

1973 *The Sociology of Science.* Chicago: University of Chicago Press.

Ogburn, William F.

1950 *Social Change.* New York: Viking Press.

Ogburn, W.F. and D.S. Thomas

1922 "Are inventions inevitable? A note on social evolution." *Political Science Quarterly* 37: 83-93.

Price, Derek de Solla

1963 *Little Science, Big Science.* New York: Columbia University Press.

Simonton, D.K.

1978 "Independent discovery in science and technology: a closer look at the Poisson distribution." *Social Studies of Science* 8: 521-32.

1979 "Multiple discovery and invention: Zeitgeist, genius or chance?" *Journal of Personality and Social Psychology* 37: 1603-1616.

Sorokin, Pitirim

1973 *Social and Cultural Dynamics.* Vol. 2. New York· Bedminster Press