Information as a Measure of Structure in Music
Author(s): Edgar Coons and David Kraehenbuehl
Source: *Journal of Music Theory*, Vol. 2, No. 2 (Nov., 1958), pp. 127-161
Published by: Duke University Press on behalf of the Yale University Department of Music
Stable URL: http://www.jstor.org/stable/843197
Accessed: 26-02-2018 08:19 UTC

INFORMATION AS A MEASURE OF STRUCTURE IN MUSIC

by

Edgar Coons and
David Kraehenbuehl
Yale University

When we wish to communicate by visual means, we draw a linear configuration whose structure is similar to that of the object or idea that we wish to represent. For example, we make a wavy line on a page to represent a snake, the surface of the ocean, the nature of a light ray, the rise and fall of emotional response, a principle of economics, or a cyclic concept of history. In every case, the wavy line serves as a reasonably satisfactory representation of the fact or idea to be conveyed because it demonstrates the most obvious structural property of that which it seeks to represent. The wavy line does not rely upon convention for its communicative effect. It relies solely upon its structural relevance to the fact or idea it is intended to convey.

The possibility of artistic communication by means of music relies also upon the structural relevance of the musical experience to the non-musical experiences that the music is intended to convey. If we are to successfully relate musical experience and general experience to provide a practical aesthetic of music, it must be through a careful definition of the structure of experience. Since musical experience is only a special case of general experience, an effective definition of structure in general experience would also be a definition of musical structure. It is only necessary to formulate this definition in terms which are independent of the elements generating the structure. This has not been accomplished to date in the case of music. Most descriptions of musical structure are couched in terms that make it impossible, except by means of vague and often misleading analogies, to relate musical structure to the structure of general experience. It is our intention to present here a method for measuring the structural properties of any experience, regardless of the specific elements that constitute the experience. We will then demonstrate some applications of this measure to problems of musical structure.

General experience at the lowest level consists of sense-data, those minute neurological responses of the individual human to the external world that are the units of perception. Through the continuing process of grouping these units in a variety of ways and combining them with new sense-data, the mind formulates and stores general concepts regarding experience. Concept formation, the abstraction of ideas from reality, relies upon the capability of the human mind to reduce the minutiae of any sensory experience to a skeleton consisting of only those units of experience which are essential to the recollection and identification of the experience. Our understanding of the natural world around us derives from any indications of structure that that world suggests.

As an example of the above, we may take a single musical pitch sounded by an oboe. We say that we hear an "oboe A." This phrase accurately identifies our extremely general idea of an intricate sensory experience. Anyone who has seen the oscillograph of an oboe tone knows that this tone is experienced by the ear as an intricately varying pattern of successive air pressures. Not one of us could distinguish consciously any single one of these air pressures which are the basic units of our perception of an "oboe A." This is not because our ear is insensitive to them, but rather because our mind immediately generalizes this array of sense-data into an inclusive single experience, the sound of an oboe. We are likewise not specifically conscious of recurrences in this pattern of air pressures. However, because the pattern does demonstrate recurrences at regular intervals, we distinguish its pitch level from those of other patterns where the recurrences are at different time intervals. The natural structural properties of this pitch sounded by an oboe assist us to identify it. Although we are not conscious of the discrete units of this sensory experience, the sense-data themselves, we are nevertheless able to abstract, from the intricate reality of the experience, the general notions of "A" and "oboe." A similar process of concept formation, relying upon the recognition of natural structures in all that we perceive, produces for us such physical concepts as "hot," such emotional concepts as "fear," and such spiritual concepts as "truth." As the total concept becomes more complex, we become able to identify specifically many of the sub-units that compose it. We speak of it as being articulate, that is, possessed of distinct elements or groups of elements which may be considered separately or in various deliberately formed relationships that may give rise to other acceptable concepts.

A musical composition is such a larger articulate structure. If the "oboe A" — an experience of whose sub-units we remain unaware — is placed in a succession of various oboe pitches, a musical phrase will result. This phrase is a larger unit of musical experience in which we are able to distinguish the sub-units, the successive pitch elements. This phrase may in turn become a sub-unit in a musical period, leading to further more extensive expansion of the musical experience to the length of a complete composition. Because this composition partakes of the same structural properties as experience in general, it will serve to represent some part of experience for its listeners. To sum up, experience can be said to exist on two general levels: the level of simple perception and the level of concept formation. The latter, in turn, consists of many levels, each level developing more inclusive concepts regarding the articulate conceptual elements of the less inclusive levels.

Since we cannot identify individual sense-data, they cannot serve as the units for any practical measure of structure in experience. In music, we must begin a study of structure at that already remarkably general level of concept, the individual sounding pitch. In dealing with an extended composition, we might assume some still more inclusive level of concept as the phrase or period to serve as the basis for our considerations. Whatever level of generality we may choose as the base-line for our structural analysis, we will refer to the articulate units of this base-line as "events." A succession of events will be called a "pattern." It is our intention to define the properties of any possible pattern of events in terms of some measure that is independent of the particular sense to which the pattern may appeal. Fortunately, we are provided with such a measure in the universal phenomenon of entropy, expressed in the field of communication as "information."[1]

In the technical sense that the term is used here, information has no necessary connection with information in the every-day sense of general knowledge. "Information" is the technical term for a measure of the degree of randomness exhibited by a pattern of events. A totally random succession of events would produce a state of maximal informedness. Any evidence of order in a pattern of events would result in a reduction of this state of maximal informedness. Since a random situation is one of minimal predictability, it can be said that information is a measure of <u>un</u>predictability. That which is not predicted, should it occur, will provide a maximum of information. That which is readily predicted, should it occur, will provide a minimum of information. In this instance our common-sense meaning for information is not different from its technical meaning. That the sun rises tomorrow is not very informative in the general sense because it is low-informed in the technical sense, i.e. it is readily predicted. Of course, the failure of the sun to rise, being altogether unpredicted in the pattern of astronomical events to date, would be a high-informed new event that would certainly also be an eminently newsworthy bit of general information.

This simple example demonstrates the underlying principle of our measure of structure in pattern. The informedness of each new event in a pattern depends upon the predictions that the pattern of events has led us to formulate to the moment. The new event may confirm these predictions or it may fail to confirm them. We shall use the expression "to nonconfirm" for this latter circumstance. Information will be a measure of the degree to which a single prediction or an array of predictions is "nonconfirmed" by the present event.[2] The general method for determining an exact value for the informedness of any

1. For an account of the application of information in the fields of electronic communication and language, see Shannon and Weaver, <u>The Mathematical Theory of Communication</u> (1949). For an account of the application of the same measures to problems of musical style, see Richard C. Pinkerton, "Information Theory and Melody," <u>Scientific American</u>, Vol. 194, 2:77-87 and Joseph E. Youngblood, "Style as Information," <u>Journal of Music Theory</u>, II:24-35.

2. This definition of information differs from that ordinarily used in the field of communication. The authors' reasons for developing this particular definition of information to measure structure in music are set forth in another article entitled "Information as a Measure of the Experience of Music" which, although prepared for reading over a year before this present article, will appear subsequently in the <u>Journal of Aesthetics and Art Criticism</u>.

given event will involve three steps. We will determine first all the predictions established by the pattern of events to the present. We will then determine whether the new event confirms or nonconfirms each prediction. A summation of the number of predictions nonconfirmed by the new event will constitute an information measure of the new event in the light of the pattern of events which preceded it.

It must be understood that we will be operating here under somewhat aseptic conditions. The patterns under consideration will be isolated from their environment and discussed as if they were each unique and in no wise affected by preceding or accompanying patterns. This will provide us with an analysis of the pattern in the abstract which can then be evaluated in terms of various types of context. This process is not unreasonable since the context of a pattern must also be a pattern whose properties could be similarly ascertained in the abstract. The problem of context does not become significant until empirical investigations are undertaken to substantiate the purely theoretical conclusions developed here. A logical analysis of the properties of pattern would provide the empiricist with a sensible basis for the control of context.

We will begin our demonstration of the information properties of a pattern of events by considering the simplest of all patterns, a single event, designated as "A." If this is the sole experience to date, it is the only possible prediction. Beginning with this experience, only two 2-event patterns are possible. They are "AA" and "AB." In the first of these, the second "A" confirms the array of predictions to date, generating a state of minimal informedness. In the second, the occurrence of "B" after "A" (and here "B" really means any "not-A") constitutes a state of maximal informedness. Maximal informedness is a property of the relationship of dissimilarity, the result of the nonconfirmation of a prediction. Minimal informedness is a property of the relationship of similarity, the result of the confirmation of a prediction.

We may take the symbol $\underline{d}$ to represent both a relationship of dissimilarity and a condition of maximal informedness. We may take the symbol $\underline{s}$ to represent both a relationship of similarity and a condition of minimal informedness. The symbols $\underline{d}$ and $\underline{s}$ will have a double significance wherever they appear in our analyses. As substantive indicators they will denote the relationships of dissimilarity and similarity respectively; as evaluative indicators, they will represent maximal and minimal states of informedness respectively. The relationships they denote will be the building blocks of our system; the values they represent will be the units of measurement in our system. The equation $A/A = \underline{s}$ will be an abbreviation for the expression "if event 'A' follows event 'A,' a relationship of similarity will exist at the second 'A,' and hence the second 'A' will be minimally informed." The equation $A/B = \underline{d}$ will be an abbreviation for "if event 'B' follows event 'A,' a relationship of dissimilarity will exist at event 'B,' and hence event 'B' will be maximally informed." Since the relationships $\underline{d}$ and $\underline{s}$ are substantive, they may themselves be the events in a pattern of events. This allows us to formulate the four basic equations, similar to those given above, that will provide the operational foundation for our measurement of information in a pattern of events: $\underline{d}/\underline{d} = \underline{s}$, $\underline{s}/\underline{s} = \underline{s}$, $\underline{d}/\underline{s} = \underline{d}$, and $\underline{s}/\underline{d} = \underline{d}$.

When these equations are translated into words, their special relevance to problems of musical structure is evident. For example, the first equation ($\underline{d}/\underline{d} = \underline{s}$) tells us that the listener's past is one of maximal informedness ($\underline{d}$). He therefore predicts or anticipates a continuation of this state of maximal informedness. Our knowledge of the future in any case is nothing more than a presumed extension of all that we know to the present. If this condition of maximal informedness which the listener predicts on the basis of his past experience actually comes to be ($\underline{d}/\underline{d}$), then he is not astonished ($\underline{d}/\underline{d} = \underline{s}$). He is minimally informed. This is why compositions which doggedly persist in seeking out the unexpected become increasingly less interesting rather than increasingly more exciting. On the other hand, if the listener's prediction is nonconfirmed, if something minimally informed occurs ($\underline{d}/\underline{s}$), the overall experience will be one of surprise, maximal informedness ($\underline{d}/\underline{s} = \underline{d}$). An example of this would be the occurrence of so simple a fact as a major triad in a long series of extremely intense and complicated chords. The triad, being low-informed by nature, comes as a decided shock in a context which has led the listener to expect nothing but the most unpredictable stream of dissimilar chords. These basic equations of our system, then, are not dry abstractions from reality. They are actually succinct statements in two values of the simplest patterns of actual experience. Out of them we will develop a measure that will indicate with subtlety the infinite degrees of informedness that may develop between the crude extremes of maximal and minimal informedness.

In order to demonstrate the application of the basic equations given above to a specific problem in pattern, we will examine some selected 3-event patterns. Having presented two events in the pattern "AB," we wish to determine the difference in structural effect if we select "A," "B," or some heretofore unknown event, "C," as the next event in the pattern. To do so, we must first identify all the predictions established by the pattern "AB." There are four, one of which cannot always be tested in the light of the new event. They are "A," "B," a relationship of $\underline{d}$ which developed when "B" followed "A," and "AB" considered as a unit, that is, as if "B" were an extension of "A" rather than distinct from "A." This last-named prediction can only be considered under certain conditions, as indicated in Table I (see p. 133). In all the tables, where an $\underline{x}$ appears as a test value for a prediction, rather than a $\underline{d}$ or an $\underline{s}$, this indicates that the prediction under consideration either cannot or need not be tested. For example, in the pattern "ABA," it is possible that the structural unit "AB" is about to be repeated as the third and fourth events of the pattern. However, since this cannot be determined for certain, the prediction of "AB" as a unit cannot be tested at the third event in the pattern "ABA." On the other hand, in the pattern "ABB," it is clear that the predicted structural unit "AB" is already nonconfirmed by the occurrence of "B" as the third event. Therefore we may test the prediction "AB" with confidence at the third event and declare that it is nonconfirmed. Of course, this nonconfirmation at the third event makes a further test of the prediction at the fourth event meaningless (as shown in Table III, p. 135). Its nonconfirmation was already firmly established at the third event.[3]

Table I (p. 133) is a summary of the predictions and the tests of the predictions for the patterns "ABA," "ABB," and "ABC." The table shows that there are three predictions that may be tested summarily if they can be tested at all. They are either fully confirmed or fully nonconfirmed by the third event in the pattern. As mentioned before, the prediction "AB" cannot be tested in the pattern "ABA." The fourth prediction, however, presents a more elaborate situation. Each test is a partial test of the fourth prediction, and the findings of the three tests must be averaged in order to determine exactly how much the fourth prediction is nonconfirmed. An averaging process requires numerical values. We may use "1" to indicate maximal informedness, "zero" to indicate minimal informedness. All information values must lie between these two extremes. Since d indicates a state of maximal informedness, we may substitute "1" for each d in Table I; likewise, since s indicates minimal informedness, we may substitute "zero" for each s in Table I. The results are shown in Table II (p. 133). Note that the values of d and s as partial tests of the fourth prediction have been averaged resulting in an overall nonconfirmation of this prediction of 0.50 for third event "A," 0.33 for third event "B," and 0.00 for third event "C."

As can be read from Table II, three predictions are tested at the third event in the pattern "ABA" and these predictions are 1.50 nonconfirmed. In the case of the pattern "ABB," four predictions are tested and 2.33 are nonconfirmed. In the case of the pattern "ABC," four predictions are tested and 3.00 are nonconfirmed. Our final informedness values should lie between "zero" (minimal informedness) and "1" (maximal informedness). This will be accomplished if the nonconfirmations are expressed as a percentage of the total predictions tested, that is, if the upper total in Table II is divided by the lower. Event "A" following the pattern "AB" is 50.0% informed, event "B" is 58.3% informed, and event "C" is 75.0% informed. From this, we know that the pattern "AB" predicts "A" most effectively since, when "A" occurs, it is the least informed of all the possible consequents to the pattern "AB"; it predicts "B" less effectively and "C" least effectively.

At first glance, it might seem that the pattern "AB" would in no wise predict the occurrence of event "C" and that, therefore, the occurrence of event "C" should be 100% informed. It is true that the events "A" and "B" do not, in a literal sense, suggest the possibility of anything other than "A" or "B" occurring in the future. However, taken together, they predict the possible recurrence of a relationship of dissimilarity. When event "C" occurs, this prediction is confirmed in that both "A/C" and "B/C" are structurally equivalent to "A/B."

---

3. Some not serious differences in information values developed in this article and in the article cited in footnote 2 resulted from a revision of our method for calculating the tests of this sort of prediction. The method described here produces more refined results than the method developed previously.

Table I

Tests of the Predictions for "AB."

Event 1		A		
Event 2		B		
Proposed events 3		A or	B or	C

Predictions	Tests	Test values		
Event 1 (A)	(1/3)	s	d	d
Event 2 (B)	(2/3)	d	s	d
Events 1-2 as a unit (AB)	(12/3?)	x	d	d
Events 1/2 (d)	(1/2:1/3)	d	s	s
	(1/2:2/3)	s	d	s
	(1/2:12/3?)	x	s	s

Table II

A Numerical Equivalent for Table I.

Event 1		A		
Event 2		B		
Proposed events 3		A or	B or	C

Predictions	Tests		Test values		
Event 1	(1/3)		0	1	1
Event 2	(2/3)		1	0	1
Events 1-2 as a unit	(12/3?)		-	1	1
Events 1/2	(1/2:1/3)				
	(1/2:2/3)	Avg.	0.50	0.33	0.00
	(1/2:12/3?)				
Total nonconfirmations			1.50	2.33	3.00
Total predictions tested			3.00	4.00	4.00

Predictions in the form "A" or "ABC," for example, will be called literal predictions and would be confirmed only by the recurrence of the identical event or events that established the prediction. Predictions in the form "A/B" will be called first-order analogical predictions and would be confirmed by any other pair of events having a similar relationship. Higher-order analogical predictions (the form "A/B:A/C" would be an example) also figure in our computations (see Table III, p. 135). These analogical predictions have a considerable influence on the informedness of a given event when it occurs. This can be most effectively demonstrated by the pattern "ABCDEF." In the pattern "ABCDE" no event has occurred that would constitute a literal prediction of the event "F." However, because the pattern "ABCDE" predicts the relationship of dissimilarity extensively, the least informed event that can follow this pattern is the heretofore unknown event "F" (informedness 20.833%), while the recurrence of any one of the events which has already occurred is more surprising, hence, more informed (event "E," 26.664%). The reverse would be true if we considered only the literal predictions in the pattern "ABCDE."

Literal predictions never demonstrate the property of partial testing. Each literal prediction is tested immediately and uniquely, if it is tested at all, as each new event occurs. Analogical predictions of any order always exhibit the property of partial testing and must therefore be dealt with separately from the literal predictions.

In Table III (p. 135), which lists all the predictions and all the tests related to each prediction for the fourth event in every possible 4-event pattern, the tests indicated between the double lines are tests of literal predictions. All those listed below the lower double line are partial tests of the indicated first-order analogical predictions. The tests of each prediction are separated by a broken line. The tests involving higher-order relationships are given below the single solid line and, at the fourth event, all such tests constitute partial tests of a single prediction as indicated. It is important to distinguish between tests involving different-order relationships, even if they be tests of the same prediction. Note that the prediction "1/2" is tested under 6 conditions by various first-order relationships but under 72 conditions by various second-order relationships. The effect of any given test is extensively reduced as its order increases.

Table IV (p. 138) is a numerical summary of the tests listed in Table III. Note that the partial tests of any single prediction have been averaged in the same fashion as demonstrated in Table II. The total nonconfirmations, the total predictions tested, and finally the informedness of each possible fourth event are given at the bottom of Table IV.

Such a tabulation as shown in Table III becomes virtually impossible for 5-event patterns because a total of 197,210 tests of predictions might be necessary to determine the informedness of any selected fifth event. Fortunately, such a tabulation is not necessary because the entire operation can be reduced to a computation which, although cumbersome because of the size of the numbers, makes it possible to calculate the informedness of an event at any position in a pattern of any

Table III

Tests of the Predictive Array at Every Possible Fourth Event.

Event 1	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
Event 2	A	A	A	A	A	B	B	B	B	B	B	B	B	B	B
Event 3	A	A	B	B	B	A	A	A	B	B	B	C	C	C	C
Event 4	A	B	A	B	C	A	B	C	A	B	C	A	B	C	D
Predictions and tests*	**Test values**														
1/4	s	d	s	d	d	s	d	d	s	d	d	s	d	d	d
2/4	s	d	s	d	d	d	s	d	d	s	d	d	s	d	d
3/4	s	d	d	s	d	s	d	d	d	s	d	d	d	s	d
12/34	s	d	x	x	x	d	s	d	x	x	x	x	x	x	x
23/4?	x	d	x	d	d	d	x	d	d	x	d	d	x	d	d
123/4??	x	d	x	d	d	x	d	d	x	d	d	x	d	d	d
1/2:1/4	s	d	s	d	d	d	s	s	d	s	s	d	s	s	s
:2/4	s	d	s	d	d	s	d	s	s	d	s	s	d	s	s
:3/4	s	d	d	s	d	d	s	s	s	d	s	s	s	d	s
:12/34	s	d	x	x	x	s	d	s	x	x	x	x	x	x	x
:23/4?	x	d	x	d	d	s	x	s	s	x	s	s	x	s	s
:123/4? ?	x	d	x	d	d	x	s	s	x	s	s	x	s	s	s
1/3:1/4	s	d	d	s	s	s	d	d	d	s	s	d	s	s	s
:2/4	s	d	s	d	d	s	d	s	s	d	s	s	d	s	s
:3/4	s	d	s	d	s	s	d	d	s	d	s	s	s	d	s
:12/34	s	d	x	x	x	d	s	d	x	x	x	x	x	x	x
:23/4?	x	d	x	s	s	d	x	d	s	x	s	s	x	s	s
:123/4? ?	x	d	x	s	s	x	d	d	x	s	s	x	s	s	s
2/3:1/4	s	d	d	s	s	d	s	s	s	d	d	d	s	s	s
:2/4	s	d	d	s	s	s	d	s	d	s	d	s	d	s	s
:3/4	s	d	s	d	s	d	s	s	d	s	d	s	s	d	s
:12/34	s	d	x	x	x	s	d	s	x	x	x	x	x	x	x
:23/4?	x	d	x	s	s	s	x	s	d	x	d	s	x	s	s
:123/4? ?	x	d	x	s	s	x	s	s	x	d	d	x	s	s	s
12/3?:1/4	x	x	d	s	s	x	x	x	d	s	s	d	s	s	s
:2/4	x	x	d	s	s	x	x	x	s	d	s	s	d	s	s
:3/4	x	x	s	d	s	x	x	x	s	d	s	s	s	d	s
:12/34	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
:23/4?	x	x	x	s	s	x	x	x	s	x	s	s	x	s	s
:123/4? ?	x	x	x	s	s	x	x	x	x	s	s	x	s	s	s

Table III (cont.)

Event 1	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
Event 2	A	A	A	A	A	B	B	B	B	B	B	B	B	B	B
Event 3	A	A	B	B	B	A	A	A	B	B	B	C	C	C	C
Event 4	A	B	A	B	C	A	B	C	A	B	C	A	B	C	D
Predictions and tests*	Test values														
1/2:1/3#1/2:1/4	s	d	d	s	s	s	d	d	d	s	s	d	s	s	s
:2/4	s	d	d	s	s	d	s	d	s	d	s	s	d	s	s
:3/4	s	d	s	d	s	s	d	d	s	d	s	s	s	d	s
:12/34	s	d	x	x	x	d	x	d	x	x	x	x	x	x	x
:23/4?	x	d	x	s	s	d	x	d	s	x	s	s	x	s	s
:123/4? ?	x	d	x	s	s	x	d	d	x	s	s	x	s	s	s
#1/3:1/4	s	d	s	d	d	d	s	s	d	s	s	d	s	s	s
:2/4	s	d	s	d	d	s	d	s	s	d	s	s	d	s	s
:3/4	s	d	d	s	d	d	s	s	s	d	s	s	s	d	s
:12/34	s	d	x	x	x	s	d	s	x	x	x	x	x	x	x
:23/4?	x	d	x	d	d	s	x	s	s	x	s	s	x	s	s
:123/4? ?	x	d	x	d	d	x	s	s	x	s	s	x	s	s	s
#2/3:1/4	s	d	s	d	d	s	d	d	s	d	d	d	s	s	s
:2/4	s	d	s	d	d	d	s	d	d	s	d	s	d	s	s
:3/4	s	d	d	s	d	s	d	d	d	s	d	s	s	d	s
:12/34	s	d	x	x	x	d	s	d	x	x	x	x	x	x	x
:23/4?	x	d	x	d	d	d	x	d	d	x	d	s	x	s	s
:123/4? ?	x	d	x	d	d	x	d	d	x	d	d	x	s	s	s
#12/3?:1/4	x	x	s	d	d	x	x	x	d	s	s	d	s	s	s
:2/4	x	x	s	d	d	x	x	x	s	d	s	s	d	s	s
:3/4	x	x	d	s	d	x	x	x	s	d	s	s	s	d	s
:12/34	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
:23/4?	x	x	x	d	d	x	x	x	s	x	s	s	x	s	s
:123/4? ?	x	x	x	d	d	x	x	x	x	s	s	x	s	s	s
:2/3#1/2:1/4	s	d	d	s	s	d	s	s	s	d	d	d	s	s	s
:2/4	s	d	d	s	s	s	d	s	d	s	d	s	d	s	s
:3/4	s	d	s	d	s	d	s	s	d	s	d	s	s	d	s
:12/34	s	d	x	x	x	s	d	s	x	x	x	x	x	x	x
:23/4?	x	d	x	s	s	s	x	s	d	x	d	s	x	s	s
:123/4? ?	x	d	x	s	s	x	s	s	x	d	d	x	s	s	s
#1/3:1/4	s	d	s	d	d	s	d	d	s	d	d	d	s	s	s
:2/4	s	d	s	d	d	d	s	d	d	s	d	s	d	s	s
:3/4	s	d	d	s	d	s	d	d	d	s	d	s	s	d	s
:12/34	s	d	x	x	x	d	s	d	x	x	x	x	x	x	x
:23/4?	x	d	x	d	d	d	x	d	d	x	d	s	x	s	s
:123/4? ?	x	d	x	d	d	x	d	d	x	d	d	x	s	s	s
#2/3:1/4	s	d	s	d	d	d	s	s	d	s	s	d	s	s	s
:2/4	s	d	s	d	d	s	d	s	s	d	s	s	d	s	s
:3/4	s	d	d	s	d	d	s	s	s	d	s	s	s	d	s
:12/34	s	d	x	x	x	s	d	s	x	x	x	x	x	x	x
:23/4?	x	d	x	d	d	s	x	s	s	x	s	s	x	s	s
:123/4? ?	x	d	x	d	d	x	s	s	x	s	s	x	s	s	s

Table III (cont.)

Event 1	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
Event 2	A	A	A	A	A	B	B	B	B	B	B	B	B	B	B
Event 3	A	A	B	B	B	A	A	A	B	B	B	C	C	C	C
Event 4	A	B	A	B	C	A	B	C	A	B	C	A	B	C	D
Predictions and tests*	**Test values**														
#12/3?:1/4	x	x	s	d	d	x	x	x	s	d	d	d	s	s	s
:2/4	x	x	s	d	d	x	x	x	d	s	d	s	d	s	s
:3/4	x	x	d	s	d	x	x	x	d	s	d	s	s	d	s
:12/34	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
:23/4?	x	x	x	d	d	x	x	x	d	x	d	s	x	s	s
:123/4? ?	x	x	x	d	d	x	x	x	x	d	d	x	s	s	s
:12/3?#1/2:1/4	x	x	d	s	s	x	x	x	d	s	s	d	s	s	s
:2/4	x	x	d	s	s	x	x	x	s	d	s	s	d	s	s
:3/4	x	x	s	d	s	x	x	x	s	d	s	s	s	d	s
:12/34	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
:23/4?	x	x	x	s	s	x	x	x	s	x	s	s	x	s	s
:123/4? ?	x	x	x	s	s	x	x	x	x	s	s	x	s	s	s
#1/3:1/4	x	x	s	d	d	x	x	x	d	s	s	d	s	s	s
:2/4	x	x	s	d	d	x	x	x	s	d	s	s	d	s	s
:3/4	x	x	d	s	d	x	x	x	s	d	s	s	s	d	s
:12/34	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
:23/4?	x	x	x	d	d	x	x	x	s	x	s	s	x	s	s
:123/4? ?	x	x	x	d	d	x	x	x	x	s	s	x	s	s	s
#2/3:1/4	x	x	s	d	d	x	x	x	s	d	d	d	s	s	s
:2/4	x	x	s	d	d	x	x	x	d	s	d	s	d	s	s
:3/4	x	x	d	s	d	x	x	x	d	s	d	s	s	d	s
:12/34	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
:23/4?	x	x	x	d	d	x	x	x	d	x	d	s	x	s	s
:123/4? ?	x	x	x	d	d	x	x	x	x	d	d	x	s	s	s
#12/3?:1/4	x	x	s	d	d	x	x	x	d	s	s	d	s	s	s
:2/4	x	x	s	d	d	x	x	x	s	d	s	s	d	s	s
:3/4	x	x	d	s	d	x	x	x	s	d	s	s	s	d	s
:12/34	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
:23/4?	x	x	x	d	d	x	x	x	s	x	s	s	x	s	s
:123/4? ?	x	x	x	d	d	x	x	x	x	s	s	x	s	s	s

*The prediction being tested is that value which appears only once at the far left of each column of test relationships.

"/" indicates a literal test
":" indicates a first-order analogical test
"#" indicates a second-order analogical test

Table IV

A Numerical Equivalent for Table III.

Event 1	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
Event 2	A	A	A	A	A	B	B	B	B	B	B	B	B	B	B
Event 3	A	A	B	B	B	A	A	A	B	B	B	C	C	C	C
Event 4	A	B	A	B	C	A	B	C	A	B	C	A	B	C	D
Predictions								**Test values**							
1	0.00	1.00	0.00	1.00	1.00	0.00	1.00	1.00	0.00	1.00	1.00	0.00	1.00	1.00	1.00
2	0.00	1.00	0.00	1.00	1.00	1.00	0.00	1.00	1.00	0.00	1.00	1.00	0.00	1.00	1.00
3	0.00	1.00	1.00	0.00	1.00	0.00	1.00	1.00	1.00	0.00	1.00	1.00	1.00	0.00	1.00
12	0.00	1.00	----	----	----	1.00	0.00	1.00	----	----	----	----	----	----	----
23	----	1.00	----	1.00	1.00	1.00	----	1.00	1.00	----	1.00	1.00	----	1.00	1.00
123	----	1.00	----	1.00	1.00	----	1.00	1.00	----	1.00	1.00	----	1.00	1.00	1.00
1/2	0.00	1.00	0.33	0.80	1.00	0.40	0.40	0.00	0.25	0.50	0.00	0.25	0.25	0.20	0.00
1/3	0.00	1.00	0.67	0.20	0.00	0.60	0.60	1.00	0.25	0.50	0.00	0.25	0.25	0.20	0.00
2/3	0.00	1.00	0.67	0.20	0.00	0.40	0.40	0.00	0.75	0.50	1.00	0.25	0.25	0.20	0.00
12/3?	----	----	0.67	0.20	0.00	----	----	----	0.25	0.50	0.00	0.25	0.25	0.20	0.00
1/2	0.00	1.00	0.42	0.65	0.75	0.50	0.50	0.50	0.46	0.50	0.42	0.25	0.25	0.20	0.00
Tests	0.00	10.00	3.76	6.05	6.75	4.90	4.90	7.50	4.96	4.50	6.42	4.25	4.25	5.00	5.00
Preds	8.00	10.00	8.00	10.00	10.00	9.00	9.00	10.00	9.00	9.00	10.00	9.00	9.00	10.00	10.00
Info (%)	00.0	100.0	46.9	60.5	67.5	54.4	54.4	75.0	55.1	50.0	64.2	47.2	47.2	50.0	50.0

length. A rigorous proof of this computation would be appropriate only to a mathematical journal and will be omitted here. The method will be outlined and proved by demonstration. Interested readers may then apply the method to ascertain the correctness of the authors' information values, given in Table XI (p. 152), which will be used later to develop certain points regarding musical structure. Readers who are not mathematically inclined may turn directly to page 148 where the musical applications of the calculated results are discussed.

We will begin with a general formula that will represent information in terms of predictions and tests of predictions. Information is the degree to which an array of predictions is nonconfirmed. Expressed as a formula:

$$1)\quad \text{information} = \frac{\text{nonconfirming tests of predictions}}{\text{predictions tested}}\,.$$

The correctness of this general formula can be tested by considering again the patterns "AA" and "AB." In each case, one prediction (A) is tested. In the case of "AA," no prediction is nonconfirmed; in the case of "AB," the single prediction "A" is nonconfirmed. Using formula 1), we get familiar results:

$$\text{information "AA"} = \frac{0}{1} = 0,\ \text{and}$$

$$\text{information "AB"} = \frac{1}{1} = 1.$$

Formula 1) may be rewritten in a more useful form which will provide the same answers as above:

$$2)\quad \text{information} = \frac{1}{\text{predictions tested}} \times \left[\frac{\text{predictions tested}}{\text{total tests}} \times \text{nonconfirming tests.}\right]$$

Formula 2) does not alter the values of formula 1) because the fraction inside the brackets is equal to unity and therefore does not change the value of the product. This fraction is inserted here to provide a form in which a general formula for the informedness of event $\underline{n}$ (E_n) in a pattern of any length can be expressed as a single summation (cf. formulas 2) and 8)).

Let us now examine the more intricate situation of the 3-event patterns considered in Tables I and II. There are three or four predictions tested, depending upon the pattern under consideration. There are from four to six tests of these predictions. Furthermore, two or three of the tests deal summarily with two or three of the literal predictions ($pred_0$) while the other two or three are partial tests of a single first-order analogical prediction ($pred_1$). To represent this situation we must distinguish carefully between literal tests ($test_0$) of literal predictions and first-order tests ($test_1$) of first-order analogical predictions. We will later have to consider predictions ($pred_2$) as they are tested by second-order analogical relationships ($test_2$). In the case of the 3-event patterns under consideration, the formula for determining the informedness of the third event (E_3) is merely an expansion of formula 2):

$$3)\ \text{information } E_3 = \frac{1}{pred_0 + pred_1} \times$$

$$\left[\left(\frac{pred_0}{test_0} \times test_0\text{-nonconfirming}\right) + \left(\frac{pred_1}{test_1} \times test_1\text{-nonconfirming}\right)\right]$$

If we substitute the indicated values in formula 3), we get precisely the same percentile results as indicated on page 132:

$$\text{info ABA } (E_3) = \frac{1}{3}\left[\left(\frac{2}{2} \times 1\right) + \left(\frac{1}{2} \times 1\right)\right] = \frac{1}{3}(1.50) = .500,$$

$$\text{info ABB } (E_3) = \frac{1}{4}\left[\left(\frac{3}{3} \times 2\right) + \left(\frac{1}{3} \times 1\right)\right] = \frac{1}{4}(2.33) = .583,$$

$$\text{info ABC } (E_3) = \frac{1}{4}\left[\left(\frac{3}{3} \times 3\right) + \left(\frac{1}{3} \times 0\right)\right] = \frac{1}{4}(3.00) = .750.$$

Formula 3) may be further expanded to become a general statement for the computation of the informedness of the *n*th event in any pattern of events:

$$4)\ \text{info } (E_n) = \frac{1}{\sum_{i=n}^{i=0} pred_i} \sum_{i=n}^{i=0} \frac{pred_i}{test_i} (test_i\text{-nonconfirming}).$$

The application of this formula requires only that we have numerical values for each term of the formula. In our illustrations to date, we have derived these values from tabular computations for 3-event and 4-event patterns. We have already indicated that such tabular computations are very nearly impossible for patterns of any greater length. We will now develop a method for computing in a summary fashion the values needed to operate formula 4).

It is first necessary to determine values for $pred_0$, $test_0$, and to represent numerically the value of $test_0$-nonconfirming. These values will provide us with the material from which we may compute by methods to be demonstrated the values for $pred_i$, $test_i$, and $test_i$-nonconfirming. To date the authors have discovered no more efficient procedure than the tabular method for deriving the values for the literal level of calculation. All other values may be operationally derived from these.

As an example of the computation procedures, we will use the pattern "ABBA." We may then compare our computed results with the tabular results given in Tables I to IV. We will begin by determining $pred_0$ and $test_0$ for the pattern "ABBA." Table VI (p. 142) represents all the literal predictions and the tests of these predictions. As we already knew from formulas 2) and 3), the values for $pred_0$ and $test_0$ are the same.

We must now compute some summary value for $test_0$-nonconfirming. In order to do so, we must provide numerical equivalents for the d and s values in Table VI. Furthermore, these numerical equivalents must be so chosen that the four basic operational equations for our system (see p. 131) will all be satisfied. In Table V (p. 142) this is demonstrated to be the case if the value +1 is substituted for s; the value -1, for d. It must be understood that these values stand only for the substantive meanings of s and d; +1 represents a relationship of similarity while -1 represents a relationship of dissimilarity. If this is kept in mind, these numerical equivalents for letter-symbols will not be confused with the numbers "0" and "1" that were earlier selected to represent states of minimal and maximal informedness.

We may now compute a summary value for the tests of predictions for each event in the pattern "ABBA" by substituting the numerical equivalents for the letters in Table VI and adding. The results are given in Table VII (p. 142). The summary of the test values (sum_0) given in Table VII represents the difference between the s and d test values. For example, the -1 under the second "B" in the pattern does not mean that there is a single d value for the tests at that event, but rather that the d values exceed the s by one. The -2 under the final "A" indicates that the d values exceed the s by 2. This can be confirmed by observing the columns of letters given for the same events under "Test values" in Table VI. If this summary of the test values of any event were positive rather than negative, this would indicate that the s values exceed the d.

We are concerned only with those tests of a prediction that prove to be nonconfirming. We wish to know the value of $test_0$-nonconfirming at each event. This may be computed from sum_0 by applying the following formula:

$$5)\quad test_0\text{-nonconfirming} = \frac{test_0 - sum_0}{2} .$$

Applying this formula to the values given in Table VII, we get the values for $test_0$-nonconfirming given in Table VIII (p. 142).

Formula 5) may be rewritten in a more general form to read:

$$6)\quad test_i\text{-nonconfirming} = \frac{test_i - sum_i}{2} .$$

By substituting the right-hand side of formula 6) for the left-hand side of formula 6) when the latter appears in formula 4), formula 4) becomes:

Table V

Numerical Equivalents for Operational Equations.

$\underline{s}/\underline{s} = \underline{s}$	$+1/+1 = +1$
$\underline{d}/\underline{d} = \underline{s}$	$-1/-1 = +1$
$\underline{s}/\underline{d} = \underline{d}$	$+1/-1 = -1$
$\underline{d}/\underline{s} = \underline{d}$	$-1/+1 = -1$

Table VI

Test Values for Pattern "ABBA."

Predictions	Test values			
	A	B	B	A
Event 1		d	d	s
Event 2			s	d
Event 3				d
Events 1-2 as a unit			d	x
Events 2-3 as a unit				d
Events 1-2-3 as a unit				x
$Pred_0$		1	3	4
$Test_0$		1	3	4

Table VII

A Numerical Equivalent for Table VI.

Predictions	Test values			
	A	B	B	A
Event 1		-1	-1	+1
Event 2			+1	-1
Event 3				-1
Events 1-2 as a unit			-1	x
Events 2-3 as a unit				-1
Events 1-2-3 as a unit				x
Summary of test values (sum_0)		-1	-1	-2

Table VIII

Computation of $Test_0$-nonconfirming.

	A	B	B	A
$Test_0$		1	3	4
Sum_0		-1	-1	-2
$Test_0$-nonconfirming		1	2	3

$$7)\quad \text{info } E_n = \frac{1}{\sum_{i=n}^{i=0} \text{pred}_i} \sum_{i=n}^{i=0} \frac{\text{pred}_i}{\text{test}_i} \left(\frac{\text{test}_i - \text{sum}_i}{2} \right) .$$

By actually performing the operations indicated in formula 7) and re-arranging the terms we come to the most useful form of the formula for purposes of actual computation:

$$8)\quad \text{info } E_n = .50 - \left[\frac{1}{2\sum_{i=n}^{i=0} \text{pred}_i} \sum_{i=n}^{i=0} \frac{\text{pred}_i \times \text{sum}_i}{\text{test}_i} \right] .$$

It remains to demonstrate the general validity of formula 6). To do so, we must actually compute the values for sum_i and test_i and demonstrate that they may be operated with in the same way as formula 5) operates with sum_0 and test_0. We must find a method for computing sum_i at the nth event so that its value will always represent the difference between s and d test values at any level of calculation. The general formula for this computation is:

$$9)\quad \text{sum}_i(E_n) = \left[\text{sum}_{i-1}(E_n)\right] \sum_{j=n-i-2}^{j=n-1} \left[\text{sum}_{i-1}(E_j)\right] .$$

In words, $\text{sum}_i(E_n)$ is the product of the value of the present event in the next lower order of relationships and the sum of the test values for all the past events in that level. When the computation is laid out in the diagrammatic form indicated below, it can be easily remembered as "the sum of the past multiplied by the present," all at the preceding level. For the pattern "ABBA," this computation would be:

	A	B		B		A
sum_0		[(-1)	+	(-1)]	x	(-2)
sum_1				1		4
sum_2						4

These results may be checked against the actual values in Tables I and III. For example, the tests of pred_1 given for the form "ABB" in Table I show 2 s values and one d value, an excess of one s value as our computed result +1 for sum_1 above would indicate. The values for sum_1 at the fourth event in the pattern "ABBA" may be found in the column for that pattern between the lower double line and the solid single line in Table III. Actual counting will show that there are 10 s values and 6 d values, the excess of 4 s values indicated by the above calculation. Finally, values for sum_2 for the fourth event may be found below the single solid line in the "ABBA" column of Table III. Again, an actual count shows 26 s values and 22 d values, a preponderance of 4 s values as computed above. It is clear that this method of calculation does produce values for sum_i at any level which are of the same type as those already determined for sum_0.

Table IX

Computation of Informedness for Pattern "ABBA."

$Pred_i$

A	B	B	A
	1	3	4
		1	4
			1
Σ	1	4	9

Sum_i

A	B	B	A
	-1	-1	-2
		1	4
			4

$Test_i$

A	B	B	A
	1	3	4
		3	16
			48

$\frac{Pred_i \times sum_i}{Test_i}$

A	B	B	A
	-1.00	-1.00	-2.00
		0.33	1.00
			0.08
Σ	-1.00	-0.67	-0.92

$$\text{info } E_2 = .50 - \frac{1}{2}(-1.00) = 1.00$$

$$\text{info } E_3 = .50 - \frac{1}{8}(-0.67) = 0.58$$

$$\text{info } E_4 = .50 - \frac{1}{18}(-0.92) = 0.55$$

It remains to determine the values for test_i. If all the values for test_0 were assumed to be the same, s, the computation for test_i would be identical to the computation for sum_i. This is because the computation of sum_i is merely a summary procedure for operating repeatedly the basic equations of our system given on page 131. If we assume that all values of test_0 are s, then all our results at any level of relationship would be s since s/s can lead to no value other than s itself. We may then determine test_i from test_0 by exactly the same calculation that we used to determine sum_i from sum_0. The results are:

	A	B	B	A
test_0		(1 +	3) x	4
test_1			3	16
test_2				48

These values may be checked in the appropriate places in Tables I and III, and they will be found to be correct. For example, below the single solid line in Table III in the "ABBA" column, there are indeed 48 tests made. We have now demonstrated the generality of formula 6), and its use to develop formula 8) is justified.

We are still without a procedure for determining readily the values of pred_i which are essential to the complete operation of formula 8). We do have values for pred_0. The general formula for the computation of pred_i is:

10) $\text{pred}_i(E_n) = \text{pred}_i(E_{n-1}) + \text{pred}_{i-1}(E_{n-1})$.

The accuracy of the results obtained by applying this formula can be tested for the pattern "ABBA":

	A	B	B	A
pred_0		1	3	4
pred_1			1	4
pred_2				1

A complete computation for the informedness of each event in the pattern "ABBA" may now be made by applying formula 8). Table IX (p. 144) demonstrates the complete computation. Table X is a final complete demonstration of the application of formula 8) to compute the informedness of each event in the 6-event pattern "ABAACA." It is this type of computation which produced the values given in Table XI (p. 152) where the informedness for each event in every possible 6-event pattern of events may be found.

It is the flux of information created by progression from event to event in a pattern of events that constitutes the reality of experience, and it is this information flux that is the mensurable aspect of experience. If methods for interpreting this measure of experience can be developed, they will provide a useful tool for the analysis of musical structure. The authors will propose a few interpretations, none of

Table X

Computation of Informedness for Pattern "ABAACA."

Predictions	Test Values					
Event	A	B	A	A	C	A
1		d	s	s	d	s
2			d	d	d	d
3				s	d	s
4					d	s
5						d
1-2			x	d	d	x
2-3				d	x	d
3-4					d	x
4-5						x
1-2-3				x	d	x
2-3-4					d	x
3-4-5						x
1-2-3-4					d	x
2-3-4-5						d
1-2-3-4-5						x
$Pred_0$		1	2	5	9	7
$Test_0$		1	2	5	9	7
Sum_0		-1	0	-1	-9	-1

$Pred_i$

	A	B	A	A	C	A
		1	2	5	9	7
			1	3	8	17
				1	4	12
					1	5
						1
Σ		1	3	9	22	42

Sum_i

A	B	A	A	C	A
	-1	0	-1	-9	-1
		0	1	18	11
			0	18	209
				0	3762
					0

$Test_i$

A	B	A	A	C	A
	1	2	5	9	7
		2	15	72	119
			30	1224	10591
				36720	13281114
					487682506080

Table X (cont.)

$\frac{Pred_i \times sum_i}{Test_i}$

A	B	A	A	C	A
	-1.000	0.000	-1.000	-9.000	-1.000
		0.000	0.200	2.000	1.571
			0.000	0.588	0.239
				0.000	0.001
					0.000
Σ	-1.000	0.000	-0.800	-6.412	0.811

$$\text{info } E_2 = .50 - \frac{1}{2}(-1.000) = 1.000^*$$

$$\text{info } E_3 = .50 - \frac{1}{6}(0.000) = 0.500$$

$$\text{info } E_4 = .50 - \frac{1}{18}(-0.800) = 0.544$$

$$\text{info } E_5 = .50 - \frac{1}{44}(-6.412) = 0.641$$

$$\text{info } E_6 = .50 - \frac{1}{84}(0.010) = 0.490$$

* This and the following results may be checked in Table XI.

which have been sufficiently tested empirically to confirm their validity.[4] It is, however, such rationally developed methods of interpretation, formulated on the basis of an objective measure, that could serve as the basis for countless carefully planned empirical tests.

The listener responds to music in two ways: he attends to it, and he is satisfied by it. If a composition is to be effective, its pattern must be one that, first of all, attracts and holds the attention of the listener and, secondly, rewards the listener for his attention. It is evident that only something which is informative will attract the listener's attention. This means that the pattern must be as high-informed as possible. On the other hand, the listener is rewarded by confirmations of his predictions. He would become frustrated and ultimately inattentive if a composition consistently nonconfirmed his predictions. This means that the pattern must be as low-informed as possible. The paradox of artistic creation could be no more bluntly stated. It is necessary to make a pattern which is simultaneously both as high-informed and as low-informed as possible in order to accomplish the contradictory requirements for keeping a listener. The composer is saved from despair in this dilemma by the fact of temporal succession which permits him to satisfy the conflicting desires of his listener by dealing with them alternately. He arranges two events in succession to achieve a maximum information gain in order to capture his listener's attention and rewards him with a third event which achieves a maximum information reduction. He continues in somewhat this fashion on every level of structural organization throughout the course of a composition. Although the concept of level in structural organization would require a second lengthy article to develop, it can be said here that the existence of various levels of organization makes it possible to arrange a pattern in such a way that both the requirements for maintaining the interest of a listener may be consistently achieved simultaneously, one at one level, the other at another. This successful presentation of an apparent paradox is characteristic of the most significant musical creations.

When a composition achieves this goal, we speak of it as being "highly articulate." We mean this compliment in two senses: it is a forceful communication which is extremely clearly organized. If we wish to assess the articulateness (on a single level only) of the many 6-event patterns for which we have given exact information values in Table XI, we must find some numerical expression for the paradoxical process described above. This is not difficult. Articulateness can be computed as the average information flux in the course of a pattern. The pattern exhibiting both the largest gains and the largest reductions of information would have the maximum index of articulateness. Patterns which demonstrate lesser gains or reductions of information would have smaller indices of articulateness. Obviously, the least articulate pattern would be "AAAAAA" with an index of "00.000." The pattern "AAAAAB" is somewhat more articulate with an index of

---

4. Other interpretations of information values and other applications of the methods described here can be found in the article cited in footnote 2.

"20.000." Such patterns as "ABACAD" (index 41.375) and "ABACDE" (index 42.917) are obviously very articulate.

Such a general measure of communicative effectiveness is somewhat misleading however. A listener wants not only to be attracted by a composition while it is in progress; he wishes also to carry away with him some concise impression of the composition. He hopes to reduce the large number of aural sense-data that constitutes his experience of the composition to a few brief musical patterns that will characterize the composition for him. In terms of the layman, he wishes his musical experience to "have a point."

By making another application of the principle of reward, we may demonstrate how a pattern of events establishes one or more of the events in the pattern as its "point." It is a principle of general learning theory that reward is one of the basic means by which learning is reinforced. We learn most readily those advices that are associated with satisfaction, with rewarding results. We have already indicated that the listener is rewarded by a confirmation of his predictions. The greater the confirmation, the greater his satisfaction. Any reduction in information represents an increased state of confirmation; therefore, the information reduction occasioned by the occurrence of any particular event constitutes a reward for the listener. The listener learns this rewarding event more readily than other less rewarding events that occasion gains in information, and he regards the better-learned event as more significant than the other events in the pattern. The greater the information reduction associated with a particular event, the more effectively that event becomes the "point," "subject," or "theme" of the pattern of events.

This line of reasoning allows us to make another useful numerical interpretation of the information values given in Table XI. We wish to know how effectively a given pattern establishes one or more of its elements as significant with respect to other elements in the pattern. We will call this an index of hierarchy. This index will be computed as the average of the information reductions created by a pattern. To demonstrate the refinement that this second index provides, let us reconsider the patterns "ABACAD" and "ABACDE." With regard to articulateness, it was found that pattern "ABACDE" was superior. With regard to hierarchy, however, pattern "ABACAD" (index 38.272) far excels pattern "ABACDE" (index 29.862). Although the maximum information reduction in both patterns is associated with occurrences of event "A," thus making event "A" the point of both patterns, it is evident that the pattern "ABACAD" makes this point far more effectively than the pattern "ABACDE." A common sense examination of the information values for each pattern provides an explanation for this.

As can be seen in Table XII, pattern "ABACAD" develops two information reductions and both are associated with the same event, event "A." In the pattern "ABACDE," there are three information reductions and, although the largest one is still associated with event "A," two lesser ones are associated with events "D" and "E." These latter two reductions tend to confuse our initial impression that event

"A" is of primary significance since they suggest events other than "A" as weaker but possible points in the pattern. In the pattern "ABACAD," the second information reduction confirms our initial impression that event "A" is the point of the pattern.

Table XII
Comparison of Patterns "ABACAD" and "ABACDE."

	A	B	A	C	A	D
Information Reductions	00.000	100.000	50.000 50.000	75.000	48.456 27.544	50.762
	A	**B**	**A**	**C**	**D**	**E**
Information Reductions	00.000	100.000	50.000 50.000	75.000	50.152 24.848	35.414 14.738

A pattern with a single large information reduction will make a point more effectively than a pattern with several lesser information reductions, even though the overall information reduction achieved by the latter pattern may be greater. Because our index of hierarchy involves an averaging process, it automatically advises us whether the information reduction achieved by the entire pattern is associated effectively with a minimum number of events or is distributed ineffectually over a large number of events in the pattern. The former case, as illustrated by the two patterns selected above, produces, as it should, a larger index of hierarchy. Furthermore, since it is always true that larger information reductions are associated with recurrences of an event which has predominated the pattern already, the largest indices of hierarchy will result for those patterns where the information reductions are consistently associated with recurrences of the same event, that is, those patterns that make a single point with great force.

It has probably become apparent to our readers that what we have been discussing throughout this article is well known to them as that admixture of "unity" and "variety" that is so often encountered in critical discussions of effective artistic communication. It is generally conceded that a significant work of art achieves "unity" through "variety" or, put differently, "variety" within "unity." Most discussions proceed from this acceptable but very general maxim to various hardly helpful accounts of the deep mystery which only unfettered genius can penetrate to achieve this desirable artistic goal. In contrast, our measures of communicative effectiveness are forthright statements of this artistic principle expressed in concrete terms. Indeed, the basic values that we have used for computing information are nothing but "unity" and "variety" in their primordial states, the relationships of similarity and dissimilarity existing between individual events in a pattern. Our index of articulateness is a measure of how neatly the conditions of "unity" and "variety" have been arranged so that the force of neither is dulled. Our index of hierarchy is a measure of how successfully a "variety" of events has been arranged to leave an impression of "unity." Taken together, these two measures constitute an objective analysis of a pattern of events to demonstrate how effectively it achieves simultaneously the desirable but apparently contradictory aims of maximal "unity" within maximal "variety." A careful investigation of the actual values given in Table XI will demonstrate to the reader that "unity" enhances "variety" and, conversely, "variety" enhances "unity." A noticeable predominance of either results in a decrease in the effectiveness of both.

Information, as computed here, provides us with a way of measuring more exactly those properties of pattern that have always been recognized as essential. Once the facts regarding the nature of pattern can be stated rigorously, it will become possible to develop a general theory of formal process in music. Since information is a measure of formal effectiveness that is independent of the specific nature of the elements composing the pattern, exact structural comparisons between musical and non-musical experiences will become feasible, paving the way to a sound theory of the symbolic processes in music. It seems reasonable to assume that the musical symbol and the reality it symbolizes have in common nothing more than their structural properties. It is these which our measure defines explicitly. There would seem to be no end to the practical and experimental avenues that such an exact account of structure would open up to the imaginative music theorist.

Table XI

An Analysis of all 6-event Patterns.

Informedness of Each Event						Index of Articulateness	Index of Hierarchy
A 00.000	A 00.000	A 00.000	A 00.000	A 00.000	A 00.000	00.000	00.000
	(00.000)*	(00.000)	(00.000)	(00.000)	B 100.000	20.000	00.000
	[00.000]	[00.000]	[00.000]	[00.000]			
				B 100.000	A 47.873	30.425	52.127
				(25.000)	B 60.945	27.811	39.055
				[00.000]	C 63.788	27.242	36.212
			B 100.000	A 47.500	A 73.759	35.752	52.500
			(33.333)	(38.125)	B 59.535	32.907	52.500
			[00.000]	[52.500]	C 64.561	33.912	52.500
				B 59.886	A 50.000	30.000	25.000
				(35.028)	B 39.741	32.052	30.130
				[40.114]	C 50.923	29.815	24.538
				C 59.545	A 50.000	30.000	25.000
				(35.114)	B 47.570	30.486	26.215
				[40.455]	C 48.425	30.315	25.788
					D 48.751	30.249	25.624
		B 100.000	A 46.875	A 50.000	A 48.911	31.468	27.107
		(50.000)	(50.911)	(39.062)	B 58.395	32.929	53.425
		[00.000]	[53.125]	[53.125]	C 64.219	34.094	53.525

* The indices shown in the last two columns are for the full 6-event patterns. The indices for patterns of shorter length are found under the final events of the patterns in the main section of the table. The index of articulateness is in parentheses; the index of hierarchy, in brackets.

Table XI (cont.)

						Index of Articulateness	Index of Hierarchy
A 00.000	A 00.000	B 100.000	A 46.875	B 60.412	A 49.829	35.449	31.854
	(00.000)	(50.000)	(50.911)	(41.666)	B 52.666	34.882	30.436
	[00.000]	[00.000]	[53.125]	[53.125]	C 55.427	34.329	29.055
				C 70.100	A 50.000	39.290	36.112
				(44.088)	B 47.872	39.716	37.176
				[53.125]	C 49.509	39.388	36.358
					D 48.421	39.606	36.902
			B 60.500	A 49.830	A 44.823	31.035	18.374
			(46.454)	(37.542)	B 52.190	30.506	25.085
			[39.500]	[25.085]	C 56.547	31.377	25.085
				B 51.616	A 49.176	30.165	16.924
				(37.096)	B 49.941	30.012	16.670
				[24.192]	C 51.571	29.686	16.127
				C 55.781	A 44.760	31.048	18.395
				(36.055)	B 44.567	31.087	18.459
				[22.110]	C 43.345	31.331	18.866
					D 41.685	31.663	19.419
			C 67.500	A 48.467	A 49.911	30.595	25.766
			(44.122)	(37.883)	B 51.189	30.851	25.766
			[32.500]	[25.766]	C 52.067	31.027	25.766
					D 53.469	31.307	25.766

Table XI (cont.)

Informedness of Each Event						Index of Articulateness	Index of Hierarchy
A 00.000	A 00.000	B 100.000	C 67.500	B 48.271	A 44.753	31.049	18.397
	(00.000)	(50.000)	(44.122)	(37.932)	B 44.904	31.019	18.347
	[00.000]	[00.000]	[32.500]	[25.864]	C 45.232	30.954	18.238
					D 43.280	31.344	18.888
				C 49.893	A 44.326	31.135	18.539
				(37.527)	B 41.812	31.638	19.377
				[25.054]	C 44.146	31.171	18.599
					D 40.736	31.853	19.735
				D 49.858	A 41.928	31.614	19.338
				(37.536)	B 37.643	32.471	20.765
				[25.071]	C 37.643	32.471	20.765
					D 38.389	32.322	20.516
					E 35.489	32.902	21.482
	B 100.000	A 50.000	A 54.444	A 50.000	A 50.000	31.778	27.222
	(100.000)	(75.000)	(51.430)	(39.722)	B 55.695	32.917	27.222
	[00.000]	[50.000]	[50.000]	[27.222]	C 62.193	34.216	27.222
				B 58.824	A 50.000	33.715	29.412
				(39.822)	B 51.153	33.485	28.836
				[50.000]	C 54.614	32.792	27.105
				C 64.145	A 49.036	35.951	32.554
				(41.444)	B 46.758	36.306	33.694
				[50.000]	C 48.517	35.955	32.814
					D 48.187	36.021	32.979

Table XI (cont.)

Informedness of Each Event						Index of Articulateness	Index of Hierarchy
A 00.000	B 100.000	A 50.000	B 54.444	A 50.000	A 54.735	32.725	27.222
	(100.000)	(75.000)	(51.430)	(39.722)	B 50.945	31.967	27.222
	[00.000]	[50.000]	[50.000]	[27.222]	C 62.193	34.216	27.222
				B 57.073	A 50.892	32.651	28.091
				(39.268)	B 50.446	32.740	28.314
				[50.000]	C 54.993	31.831	26.040
				C 64.145	A 47.286	36.201	33.430
				(41.444)	B 47.108	36.236	33.518
				[50.000]	C 47.629	36.132	33.258
					D 47.036	36.251	33.554
			C 75.000	A 48.456	A 50.762	40.770	38.272
			(58.275)	(50.386)	B 50.342	40.686	38.272
			[50.000]	[38.272]	C 51.841	40.986	38.272
					D 53.785	41.775	38.272
				B 48.413	A 44.698	41.060	26.741
				(50.397)	B 44.862	41.028	26.686
				[38.294]	C 43.665	41.267	27.085
					D 43.350	41.330	27.189
				C 50.114	A 44.004	41.199	26.972
				(49.972)	B 44.848	41.030	26.691
				[37.443]	C 44.104	41.179	26.938
					D 40.717	41.857	28.066

Table XI (cont.)

Informedness of Each Event						Index of Articulateness	Index of Hierarchy
A 00.000	B 100.000	A 50.000	C 75.000	D 50.152	A 41.232	41.754	27.895
	(100.000)	(75.000)	(58.275)	(49.962)	B 36.855	42.629	29.352
	[00.000]	[50.000]	[50.000]	[37.424]	C 37.536	42.493	29.126
					D 38.238	42.352	28.892
					E 35.414	42.917	29.832
		B 58.333	A 55.093	A 52.379	A 50.697	29.861	12.326
		(70.834)	(48.254)	(36.905)	B 49.874	30.025	12.532
		[41.667]	[22.454]	[15.858]	C 53.807	29.810	15.858
				B 51.109	A 51.960	29.948	16.281
				(37.223)	B 50.215	29.957	12.446
				[16.281]	C 56.266	30.810	16.281
				C 59.416	A 44.577	32.814	19.929
				(37.308)	B 44.760	32.777	19.868
				[22.454]	C 43.386	33.052	20.325
					D 40.931	33.543	21.143
			B 50.000	A 58.391	A 52.149	32.927	18.729
			(50.000)	(39.598)	B 49.316	33.493	19.672
			[25.000]	[25.000]	C 54.358	32.485	17.993
				B 48.322	A 58.903	32.452	17.209
				(37.920)	B 46.931	30.614	13.267
				[17.209]	C 61.704	33.012	17.209

Table XI (cont.)

Informedness of Each Event						Index of Articulateness	Index of Hierarchy
A 00.000	B 100.000	B 58.333	B 50.000	C 65.074	A 46.808	36.468	22.733
	(100.000)	(70.834)	(50.000)	(41.268)	B 50.000	36.029	21.670
	[00.000]	[41.667]	[25.000]	[25.000]	C 47.624	36.505	22.461
					D 47.036	36.622	22.657
			C 64.167	A 45.362	A 46.096	33.408	30.236
			(49.118)	(41.576)	B 47.505	33.648	30.236
			[41.667]	[30.236]	C 42.858	33.762	20.971
					D 45.046	33.324	20.242
				B 47.681	A 49.234	33.108	29.076
				(40.997)	B 48.893	33.040	29.076
				[29.076]	C 49.401	33.141	29.076
					D 44.528	33.428	20.415
				C 48.497	A 41.148	34.104	21.540
				(40.793)	B 43.945	33.545	20.609
				[28.668]	C 43.673	33.599	20.700
					D 39.901	34.353	21.956
				D 47.996	A 36.676	34.998	23.030
				(40.918)	B 41.371	34.059	21.466
				[28.919]	C 36.676	34.998	23.030
					D 37.415	34.851	22.784
					E 34.272	35.479	23.830

Table XI (cont.)

Informedness of Each Event						Index of Articulateness	Index of Hierarchy
A 00.000	B 100.000	C 75.000	A 47.222	A 48.583	A 48.970	30.905	26.389
	(100.000)	(62.500)	(50.875)	(38.535)	B 47.906	30.963	17.801
	[00.000]	[25.000]	[26.389]	[26.389]	C 48.845	30.900	26.389
					D 49.672	31.046	26.389
				B 47.639	A 45.327	31.101	18.345
				(38.299)	B 45.327	31.101	18.345
				[26.389]	C 45.327	31.101	18.345
					D 44.674	31.232	18.562
				C 49.338	A 46.519	31.543	18.514
				(38.724)	B 41.882	32.470	20.058
				[26.389]	C 44.520	31.943	19.179
					D 42.121	32.422	19.978
				D 51.330	A 40.602	33.523	21.147
				(39.222)	B 35.908	34.462	22.711
				[26.389]	C 36.549	34.333	22.497
					D 37.300	34.183	22.247
					E 34.130	34.817	23.303
			B 47.222	A 47.639	A 45.852	30.996	18.170
			(50.875)	(38.299)	B 47.378	30.691	17.662
			[26.389]	[26.389]	C 44.666	31.234	18.565
					D 44.674	31.232	18.562

Table XI (cont.)

Informedness of Each Event						Index of Articulateness	Index of Hierarchy
A 00.000	B 100.000	C 75.000	B 47.222	B 49.559	A 46.694	31.596	18.529
	(100.000)	(62.500)	(50.875)	(38.779)	B 48.342	31.266	17.980
	[00.000]	[25.000]	[26.389]	[26.389]	C 45.855	31.764	18.809
					D 47.011	31.533	18.424
				C 49.559	A 46.694	31.596	18.529
				(38.779)	B 48.342	31.266	17.980
				[26.389]	C 47.541	31.427	18.247
					D 47.011	31.533	18.424
				D 51.330	A 36.549	34.333	22.497
				(39.222)	B 41.276	33.388	20.923
				[26.389]	C 36.549	34.333	22.497
					D 37.300	34.183	22.247
					E 34.130	34.817	23.303
			C 50.000	A 43.904	A 45.505	31.539	18.680
			(50.000)	(39.024)	B 44.673	31.373	18.680
			[25.000]	[18.680]	C 45.020	31.442	18.680
					D 44.135	31.265	18.680
				B 45.389	A 42.352	31.530	14.412
				(38.554)	B 43.892	31.222	14.027
				[18.185]	C 43.823	31.235	14.044
					D 41.536	31.693	14.616

Table XI (cont.)

Informedness of Each Event						Index of Articulateness	Index of Hierarchy
A 00.000	B 100.000	C 75.000	C 50.000	C 47.234	A 46.545	30.691	13.364
	(100.000)	(62.500)	(50.000)	(38.192)	B 45.741	30.852	13.565
	[00.000]	[25.000]	[25.000]	[17.571]	C 49.390	30.484	17.571
					D 46.822	30.636	13.294
				D 45.999	A 34.630	32.874	16.092
				(38.500)	B 33.891	33.222	16.527
				[17.982]	C 40.012	31.998	14.997
					D 36.361	32.728	15.910
					E 32.954	33.409	16.762
			D 50.000	A 26.129	A 40.299	37.608	24.599
			(50.000)	(43.468)	B 39.924	37.533	24.599
			[25.000]	[24.599]	C 37.531	37.055	24.599
					D 37.531	37.055	24.599
					E 36.124	36.773	24.599
				B 37.013	A 35.471	32.906	16.132
				(40.747)	B 39.095	33.014	20.975
				[20.975]	C 34.734	33.053	16.316
					D 35.471	32.906	16.132
					E 33.266	33.347	16.684

Table XI (cont.)

Informedness of Each Event						Index of Articulateness	Index of Hierarchy
A 00.000	B 100.000	C 75.000	D 50.000	C 37.013	A 35.471	32.906	16.132
	(100.000)	(62.500)	(50.000)	(40.747)	B 34.734	33.053	16.316
	[00.000]	[25.000]	[25.000]	[20.975]	C 39.098	33.014	20.975
					D 35.471	32.906	16.132
					E 33.266	33.347	16.684
				D 38.587	A 33.712	33.257	16.572
				(40.353)	B 33.136	33.373	16.716
				[20.451]	C 33.712	33.257	16.572
					D 38.361	32.328	15.410
					E 30.542	33.892	17.364
				E 34.783	A 26.001	34.800	18.500
				(41.304)	B 25.480	34.904	18.630
				[21.717]	C 26.001	34.800	18.500
					D 26.001	34.800	18.500
					E 26.664	34.667	18.334
					F 20.833	35.833	19.792