THE COGNITIVE BASIS OF AESTHETIC EXPERIENCE

Margot D. Lasher,* John M. Carroll** and Thomas G. Bever***

Abstract—*This paper outlines a psychological theory of aesthetic experience. The theory is based upon the role of representational conflict and the natural process of resolving such conflict during mental growth in childhood. It proposes that aesthetic experience in adults re-creates the type of mental and emotional activity that is typical of childhood mental growth. Works of art provide the experience of mental growth and the emotional responses accompanying this growth, long after the initial process of understanding the world has been completed. The framework for mental representation and the resolving of representational conflicts comes from the field of cognitive psychology and the developmental theory of Piaget.*

This paper draws upon cognitive psychology to outline a theory of aesthetic experience [1]. One of the central ideas of cognitive psychology is that the mind actively represents the external world through abstract symbols. These symbols, often called mental representations and schemas, have been the subject of a great deal of recent theoretical and experimental work. Questions addressed relate to how symbols are formed in the process of experiencing the world; in what type of code or codes they are stored as memory; how they influence our subsequent experiences of the world; and how they are related within the mind as knowledge of the world. Since art has traditionally been considered a symbolic representation of some aspect of the world, it should not be surprising that the insights of cognitive psychology about the process and structure of such representations should be closely related to questions about the psychological role of art and aesthetic experience.

Experimental methodology has enabled psychologists to observe, indirectly, the formation of a mental representation. In one such experiment, Posner and Keele took a pattern of dots, called the 'archtype', and by successively transforming the pattern through the use of a computer, generated a set of patterns which were variations upon the archtype [2]. People in an experiment were shown a set of these variations but were not shown the original archtype. Later, these same people were shown a second set of variations which included variations actually seen before ('old' patterns) and variations never seen before ('new' patterns). Included in this second set was the archtype, which they had never seen, but which represented the relationships between the dots in the patterns which they had seen.

People were generally quite capable of discriminating between old and new patterns. But in the case of the archtype, they tended to think that it was an old pattern. They thought that they had seen it before. In some sense, they recognized it. The cognitive theory which explains this sense of recognition assumes that people were forming a mental representation of the relationships between the different patterns being experienced in the first set. This mental representation preserved certain features (dots, for example) and relationships between features (contours formed by mentally connecting the dots: a triangular relationship, for example). When shown the archtype, subjects 'recognized' it because it was an instantiation of the acquired mental representation.

Mental representations serve an essential function in our experience of the world: they establish constancy and stability among continual small differences and perceptual change. A brief description of our visual perception of the world illustrates the crucial need for such organizing constancy. As we walk through space the specific section of our surrounding environment which is available to our vision is continuously changing. What was ahead of us and in view at one moment is at our side and then behind us in the next few moments. The same change in the actual visual scene occurs when we sit still and move our head, or when we hold our head still and move only our eyes. But the mind is rarely confused or disoriented by these constant changes. We carry with us a mental representation of our surrounding space, including the space behind us at any particular time, and each separate, slightly different view is fitted into this abstract framework. Many years ago Tolman called this representation of space in animals a "cognitive map" [3]. Gibson has shown that, like Posner and Keele's archtype, the world provides an ordered structure which leads us to connect the variations: in the world, for example, the sky is always above and the earth below, the lamp remains to the right of the plant as long as we are seated at our desk, and so on [4].

It is easy to see that a child needs to develop abstract representations in order to make sense out of the world. Mental representations collect a set of experiences into a unified concept. Piaget's work on the development of knowledge in children, however, has given psychologists insight into a special kind of unifying concept [5]. This kind of concept is created when two already established mental representations are suddenly perceived as conflicting. In the course of the natural process of mental growth the child discovers a third representation which, by its mere existence, dissolves the conflict. This particular kind of representational conflict and resolution occurs in the discovery of object constancy.

The game of peek-a-boo illustrates this special kind of representational conflict and resolution in children. In this game the adult's face is hidden behind a pillow, say, and then emerges again. The infant sees the face, sees the pillow only, and then sees the face again. When the face re-appears, the infant often expresses pleasure, and with experience will want the adult to repeat the disappearance and re-appearance over and over again.

The infant can have two possible representations of the situation when the face goes behind the pillow. In the first representation the face is gone. In this developmentally early representation things which are not immediately present to one of the perceptual systems are simply not there, and the infant would cease to pay any attention to the pillow. This representation is actually backed by a great deal of infant experience: situations tend to remain relatively constant from one moment to the next, and in normal experience when a person walks out of sight beyond the door that person does not re-appear in the same place at the next instant.

*Psychologist, Department of Psychology, Assumption College, Worcester, MA 01609, U.S.A.
**Watson Research Center, IBM.
***Columbia University.

After experience with the game, however, the infant develops a second representation: the memory of the face re-appearing at the next moment from behind the pillow. The excitement of the game is created in part by an expectation of the face re-appearing. The two representations of the face when it is out of sight, not there and yet expected to be there, are resolved through the creation of the unifying representation of object constancy. Objects which cannot be seen or otherwise perceived can still be there. Later, the infant actively tests the validity of this resolution by reaching to push aside the pillow to see if the face is really still there, although hidden.

This unification of conflicting representations through the discovery of a new, resolving representation occurs naturally throughout childhood and is of central importance for mental growth. This kind of resolution occurs also in adult life in an activity which psychologists have broadly called problem solving [6]. In adult experience the discovery of the new representation which allows the conflict to be automatically and effortlessly resolved is often called 'insight'.

In both informal descriptions of the experiences of insight and more formal descriptions by philosophers and psychologists there is general agreement that insight is accompanied by two kinds of emotion. The first emotion can be described as a surge of excitement at the moment of discovery, the 'aha' experience. The second is a calmer feeling of satisfaction following the first emotion.

This second, calmer emotion is the easier of the two to explain. The feeling of satisfaction can be understood as a release from the tension created by the previous representational conflict. In explicit, adult problem solving the period of satisfaction is often purposefully extended by the act of verifying that the solution is correct and can be integrated with established knowledge. The process of verification appears more informally in the behaviour of children who have just solved some representational problem. We have mentioned that the infant who has just discovered object constancy will reach for the pillow to make sure that the face is actually there behind it.

It is more difficult to explain the initial arousal of emotion accompanying insight, especially because the framework of cognitive psychology has not dealt much with the relationship between mental and emotional experience. Two ancient traditions, one in Western and the other in Eastern philosophy, provide a way to approach this emotional arousal. The Western tradition is exemplified by Plato and continues throughout Western culture in both philosophical and religious works. For the Eastern tradition we can refer to Zen Buddhism, a form of philosophy which is historically related to both Indian and Chinese philosophical traditions. These very different Western and Eastern traditions seem to agree on certain things relating to the experience of insight and the role of art.

Both traditions treat art as a potential path toward a fuller understanding of what is true. They both use the word *art* in a broad sense, generally meaning a discipline that can be studied and perfected. In Plato, for example, dialectic and gymnastics are both considered arts [7], as are archery and the tea ceremony in Zen Buddhism [8].

The aspect of the discipline which makes it an art does not seem to reside in any description of the discipline itself, but only in the process by which it is taught and studied. This process involves the creation of harmony between mental, emotional and physical aspects of the student. It involves a giving or moulding of the self to the perfect form of the discipline. The study of the discipline then becomes a path to the attainment of spiritual or philosophical truth.

The important thing for our understanding of the emotional aspect of insight is that in both traditions the truth expresses itself. The truth is always there—it is as though it has object constancy—but its perception is obscured by confused ways of thinking and behaving in the ordinary world. This philosophical approach to truth, constantly there, independent of human mind or will, provides an explanation for certain phenomenological aspects of insight. Hard work and good intentions function only up to a point. Insight often comes spontaneously, at a moment when we are not consciously working on the problem at all. It is not a product of conscious thought or will; it seems just to happen. "If everything depends on the archer's becoming purposeless and effacing himself in the event, then its outward realization must occur automatically with no further need of the controlling or reflecting intelligence." (Eugene Herrigel [9].) In the parable of the cave in *The Republic* the philosopher is *dragged* from the darkness of the cave into the light of the sun. His enlightenment is not only beyond his personal, intellectual control, it is actually against his conscious will [10].

What emerges from these two traditions is the suggestion, translated into psychological terms, that the emotion accompanying insight is one of surprise. If the insight is in some sense unexpected, if we cannot control the moment that insight comes, then we are in a psychological situation in which we are capable of being surprised. Furthermore, since the particular kind of experience with which we are concerned is the discovery of a *new* form of representation, it is clear that the actual form of the insight cannot be anticipated. Again, we are vulnerable to surprise.

We have, then, a representational conflict which is accompanied by some emotional tension. There is next the discovery of a mental resolution which is immediately recognized as being correct in that it dissolves the preceding conflict; this discovery is accompanied initially by emotional surprise, and then a sense of satisfaction and calm created by the dissolution of previous tension.

We propose to define aesthetic experience as the experience of such a representational and emotional sequence. This kind of creative resolution occurs naturally and unconsciously during childhood. Children are kept busy solving representational problems presented to them by ordinary experience. These mental activities are enjoyable, both because they solve problems and because in doing so they release emotional energy. But, by adolescence, problems of this perceptually based, unconscious kind are in some sense solved.

Aesthetic experience fills the function in adults that normal cognitive growth serves in childhood. The experience of art is a special case of aesthetic experience. Works of art provide adults with new problems to solve, or perhaps more accurately, with more complex, difficult versions of the old problems. Art re-creates the mental and emotional processes typical of childhood mental growth. Just as the world provides the structure and coherence for the discovery of a unifying resolution in childhood, the successful work of art provides this structure and coherence for adults. The experience of art is thus of central importance to the continuation of mental and emotional growth throughout life.

We would now provide two examples of aesthetic experiences evoked by works of art (which contrast with peek-a-boo, for example). 'Mary, don't you weep' is an American folk song surviving through oral tradition. It illustrates the typical folk song structure.

> If I could I surely would
> Stand on the rock where Moses stood.
> Pharaoh's army got drowned,
> Oh, Mary, don't you weep.
> *Oh, Mary, don't you weep don't you moan*
> *Oh, Mary, don't you weep don't you moan*
> *Pharaoh's army got drowned,*
> *Oh, Mary, don't you weep.*
> Mary wore three links of chain,

Every link was Jesus' name.
Pharaoh's army got drowned,
Oh, Mary, don't you weep.
(*Chorus*)
Mary wore three links of chain,
Every link was Freedom's name.
Pharaoh's army got drowned,
Oh, Mary, don't you weep.
(*Chorus*)
One of these nights about twelve o'clock
This whole world's goin' reel and rock
Pharaoh's army got drowned,
Oh, Mary, don't you weep.
(*Chorus*)
Moses stood on the Red Sea shore
Smotin' the water with a two-by-four.
Pharaoh's army got drowned
Oh, Mary, don't you weep.
(*Chorus*)

The typical folk song consists of a set of verses. The lyrics of each verse change, but the musical pattern remains constant. Each verse is typically followed by a chorus. In the simplest form the chorus remains constant in both musical pattern and words throughout the song. One structural representation for mentally organizing the folk song assigns the perceived song elements (the actual sounds in the form of notes and words) to a reduplicating pattern: verse, chorus, verse, chorus, ...

At the same time, however, the typical folk song tells a story that develops sequentially from the beginning to the end of the song. This second representation assigns the song elements to a sequential structure which moves forward over time. The repetitive representation conflicts with the sequential representation.

There is also a typical approach to the resolution of this representational conflict: each verse ends with a common line or lines which are capable of both advancing the sequential narrative structure and strengthening the repetitive verse–chorus structure. In many songs this common line is then repeated, either exactly or with variation, to form the chorus.

The common line serves as the structural link between conflicting representations. In the resolving representations, sequential motion and static repetition are unified through a structure in which sequential events and developing meanings are seen as repeating in a cyclical pattern. The common line, by belonging to both of the conflicting representations, enables the discovery of the cyclical structure which incorporates both sequence and repetition.

The relationships between representations range from very simple to extremely subtle, and of course depend upon the lyric content. 'Mary, don't you weep' has a highly repetitive structure and appears deceptively simple, especially when the lyrics are divorced from the music. If we look at the narrative structure, we see that the first and fifth verses refer to the miracle of Moses' parting of the Red Sea, and the second and third verses refer to the miracle of Jesus, linking Moses and Jesus with the concepts of bondage and freedom. Obliquely in the first verse, and directly in the fourth verse, there is a reference to a miracle of the future. Thus the narrative itself does not follow any natural order in real or historical Biblical time: time weaves back and forth as (1) future and past (Old Testament), (2) past (New Testament), (3) past (New Testament), (4) future, (5) past (Old Testament). The one temporal element that is missing, thus far, is the present.

Common lines completing each verse are:

Pharaoh's army got drowned
Oh, Mary, don't you weep.

These lines are followed by the chorus, which consists of variations upon and an exact repetition of these common lines. The first of the common lines completes the narrative of the past: in both the Old and New Testament miracles, the enemy was defeated and the oppressed people of the world became free. In the second of the two common lines we are brought directly into the present, and symbolically into the future: Mary should not despair because the song, which encapsulates all of Biblical history, is assuring her that another miracle is going to occur. The common lines and the chorus seem unrelentingly repetitious because they carry the burden of pulling together both the past and the future into a meaning for the present. The repetition of the musical pattern increases the strength of this experience.

Thus the common lines complete the sequential structure, bring a unifying order to the temporal sequence by pulling together past and future into a present context, and at the same time, just by being a repeating element, strengthen the repetitive verse–chorus structure. The unifying representation is a cyclical structure which sees events, in this case miracles, as repeating throughout time.

Our example from dance will be more general because of the difficulty of reproducing a text for the reader. When we watch dance we are watching the human body creating patterns of motion over time. We recognize many of these patterns through their conformity to an 'archetypical' mental representation: a run and then a leap is a familiar pattern whether observed in the context of a child leaping over a puddle, a basketball player making a jumpshot, or a dancer dancing. The mental representation of a coherent pattern of motions can be called an event unit. Hochberg [11], Carroll [12] and Lasher [13] have all worked on descriptions of the structure of such event units.

In the perception of any sequence of stimuli, such as the sounds of language and music or the visual sequences of film and dance, we tend to divide the ongoing flow of perception into distinct units. The division of the sequence, and thus the structure of the units, is made according to meaningful principles. In the case of human motion, one of these principles is the function of the motion. The motion is assumed to have some function, and a unit is perceived as completed only when that function is perceived as completed.

In ordinary motions the function of the event is often related to some visible object. We lean to reach a pen on the other side of the table; we leap to avoid stepping into a puddle. The presence of these objects acts as a powerful determinant of function when we are watching someone else's motion. Dance abstracts bodily motion from actual objects, and thereby abstracts the motion from any obvious function. The dancer does not lean forward for an object or for distance; it is for form.

Yet dance motions are derived from ordinary motions and are interpretable in terms of our understanding of our own and others' ordinary motions. The mental representation of ordinary human motion acts as a base for the development of representations of dance, just as the representation of ordinary language acts as a base for representations of poetry and song. The ongoing flow of dance is divided into phrase units which are related in formal ways to event units.

The a-functional nature of dance creates manifold possibilities for conflicting mental representations. Potentially conflicting representations which are normally resolved by subsuming both under the perceived function of the motion become experienced in dance as in conflict. An interesting example is the potentially powerful conflict between the two perceptual systems primarily involved in the understanding of dance, the kinaesthetic and the visual systems.

When the visual system is freed from its customary task of picking up clues to the motion's function, it can experience the moving body more fully as a pattern of geometrical shapes changing over time. This pattern has its own visual structure.

When a dancer leaps, for example, the visual pattern presents a frozen image of the dancer's body at the height of the leap. In the purely visual world of representation, the change from stillness to motion signifies the beginning of an event, and the change from motion to stillness signifies its end. Thus the freezing of the image at the height of the leap would signify the end of that event: the leap would end at its peak.

But the leap is also being interpreted kinaesthetically, and in kinaesthetic representations the perceived three-dimensional, weighted body is always moving in relation to the force of gravity and the ground. In the kinaesthetic representation the dancer must return to the stability of the ground before the event unit can coherently end. There is, then, a conflict between the visual and kinaesthetic representations of this event: they end at different times and in different places.

The aesthetic resolution occurs through the creation of a unifying representation of static, bodily form which is capable of symbolizing both the visually and kinaesthetically represented motion. The unifying representation symbolizes the entire event unit as a single configuration. The dancer's body is abstracted momentarily from time and place; that is, from motion. The entire event unit of a run–leap, for example, is represented as the configuration of the body at its most unstable moment, the height of the leap.

Some people say that they watch athletic events for the beauty of the bodily motion as well as for the competitive or recreational aspects. For these people the athletic event is an aesthetic experience because the structure of the bodily motion (leaping to catch a ball or turning and reaching in a highly precise repetitive form in pitching) allows them to abstract a static, unifying configuration from the visual and kinaesthetic structures. The stop–action of television cameras can provide something analogous to Posner and Keele's archtype for television viewers: in an aesthetically satisfying picture of frozen motion we can sometimes recognize an instantiation of our own unifying mental representation of that event. Then we have that

'aha' experience, the sense of recognition and, in some sense, surprise at the rightness of the image.

These two art forms, the folk song and dance, illustrate what we mean by aesthetic conflict-resolution. It seems that the creative resolutions established by works of art are so complex that, unlike object constancy, we never establish them as permanent mental assumptions: thus a work of art continues to create conflict-resolution over multiple experiences.

REFERENCES

1. The paper is based upon a forthcoming book by M. D. Lasher, J. M. Carroll and T. G. Bever, *The Cognitive Basis of Aesthetic Experience: a Generative Theory.*
2. M. Posner and S. Keele, On the Genesis of Abstract Ideas, *J. Exp. Psychol.* **77**, 353 (1968).
3. E. Tolman, Cognitive Maps in Rats and Men, *Psychol. Review* **55**, 189 (1948).
4. J. J. Gibson, *The Senses Considered as Perceptual Systems* (New York: Houghton Mifflin, 1966).
   J. J. Gibson, *The Ecological Approach to Visual Perception* (New York: Houghton Mifflin, 1979).
5. J. Piaget, *The Psychology of Intelligence* (New Jersey: Littlefield, Adams, 1966), see especially pp. 107–116.
6. A. Newell and H. Simon, *Human Problem Solving* (New Jersey: Prentice Hall, 1972).
7. *The Dialogues of Plato*, trans. by B. Jowett (New York: Random House, 1937).
8. D. Suzuki, *Zen and Japanese Culture* (Princeton: Princeton University, 1970).
9. E. Herrigel, *Zen* (New York: McGraw-Hill, 1964) p. 61.
10. *The Dialogues of Plato, ibid.,* Vol. I, p. 775.
11. J. Hochberg, In the Mind's Eye, *Contemporary Theory and Research in Visual Perception,* R. Haber, ed. (New York: Holt, Rinehart & Winston, 1968).
12. J. Carroll, *Toward a Structural Psychology of Cinema* (The Hague: Mouton, 1980).
13. M. Lasher, The Cognitive Representation of an Event Involving Human Motion, *Cognitive Psychology* **13**, 391 (1981).