3 • Incongruity in Humor: The Cognitive Dynamics

SHIRLEY WILLIS MAASE • EDWARD L. FINK • STAN A. KAPLOWITZ
University of Maryland • Michigan State University

The study of humor has spanned more than a dozen centuries and has drawn the interest of researchers from a wide variety of fields: anthropology, communication, philosophy, experimental social psychology, education, and sociology. The use of humor and the humor response has been suggested to be related to a variety of social and communication functions (for example, therapy, education, persuasion, and social influence). However, despite the interest in and the suggested social importance of humor, we are no closer to developing a generalized theory of humor than we were in the first century A.D., when Quintilian complained that no one had yet explained what laughter was, though many had tried (Morreall, 1983).

Some approaches to the study of humor have focused on the cognitive activity involved in the response to humor, while others have focused on the emotional aspects of that response. Although both cognitive and affective orientations have laid important groundwork for the development of a more generalized theory of humor, often the most basic assumptions of these theories have not been stated precisely or tested empirically.

The purpose of this chapter is to state formally and precisely some of the most basic assumptions shared by a number of humor theories, and to demonstrate the utility of metric multidimensional scaling (MMDS) models for the explication and test of those assumptions from a cognitive perspective. To accomplish these objectives, the remainder of this chapter will be...
devoted to the presentation of four topics. First, a review of the relevant humor literature will be presented to establish which assumptions and hypotheses warrant further elaboration and research. Second, existing spatial and temporal models will be reviewed. Third, several general MMDS models will be described and their relevance for the study of humor will be explained. The fourth section uses these models to develop empirically testable hypotheses regarding humor. Finally, a summary of this chapter will be provided.

COGNITIVE AND AFFECTIVE THEORIES OF HUMOR

Although theorists seem to agree that the humor response disturbs the balance of both emotion and thought (e.g., see Clark, 1970; Koestler, 1964; Wicker; Thorelli, Barron, & Ponder, 1981), most research focuses on one or the other aspect rather than on both. Eysenck (1942, pp. 303-305) grouped theories of humor into two major categories: (1) cognitive theories, which stress elements such as incongruity, contrast between ideas, or deceived ideational expectancies; and (2) affective theories, which equate laughter with the satisfaction of needs or drives, or which stress the emotional correlates of laughter.

Cognitively oriented theories view humor as a result of cognitions that are perceived to be incongruous. The incongruity creates a cognitive imbalance; the presentation of the incongruity is surprising or unexpected, and the imbalance is rapidly resolved. One of the more celebrated of the early incongruity theorists was Kant (1952), who attempted to link the cognitive and physical aspects of laughter. Laughter, according to Kant, is "an affectation arising from a strained expectation being suddenly reduced to nothing" and "in this presentation, the understanding, missing what it expected, suddenly let go its hold, with the result that the effect of this slackening is felt in the body by the oscillation of the organs" (p. 538).

The basic ideas of most incongruity theorists, however, are similar to the two-stage process outlined by Suls (1972). The first stage occurs when the perceiver encounters an incongruity between the anticipated and the actuality. In the second, or problem-solving, stage, the receiver reconciles the incongruous elements. The humor is the difference between what is expected and the actual state of affairs. Herhardt (1970) explored the hypothesis that the presentation of nonhumorous material would also produce a humorous response if there had been a corresponding violation of expectancies. Herhardt found that subjects lifting weights smiled and laughed when presented with a weight that was divergent from a series of prior
The relationship between incongruity and amusement is not simple, however. While Kenny (1955) found a negative linear relationship between amusement and incongruity, Shurcliff (1968) and Nerhardt (1970) found a positive linear relationship. An additional study by Nerhardt (1975) and evidence provided by Wilson (1979) suggest there is a curvilinear, inverted U-shaped relationship between amusement and incongruity, with moderate levels of incongruity producing the most amusement. Wilson suggests that the earlier studies may have employed a small range of incongruities and thus were reporting only the slope of one side of the inverted U-shaped curve.

Not only is the relation between the amount of incongruity and humor complex, but so is the relation between the type of incongruity and humor. While Kant suggests that the resolution of incongruity is necessary for humor, Clark (1970) suggests that amusement is the enjoyment of perceiving, or indulging in the incongruity. For jokes, moreover, theorists disagree about, or sometimes fail to specify clearly, which parts of a joke are incongruous (Godkewitsch, 1976a). Godkewitsch suggests that incongruity does not lie between joke body and punch line, but rather between occurrences within the joke body; the punch line serves to give the necessary information to resolve the incongruity. In a test of this hypothesis, Godkewitsch found that the relation between incongruity, its placement in a joke (i.e., within the joke body or punch line), and amusement appears to be confounded by the theme of the joke. For example, sex jokes were funnier the more predictable and expected they were and the less surprising their punch lines were. Yet one class of verbal put-ons (nonsense jokes) was rated funnier the more unusual and the more unlikely the occurrences within the joke body were. Funniness of both sexual and nontendentious jokes correlated positively with a measure of punch line appropriateness. The results of studies such as this indicate that incongruity is relevant to the cognitive processing of humor, but the wide variety of interpretations and applications of incongruity has yielded inconclusive and sometimes inconsistent results.

Oretic or affective theories assume that the incongruous elements of a joke cause increased tension, drive, or arousal, which is released with the resolution of the joke. Berlyne (1960) distinguishes two ways in which arousal may be linked to humor. The arousal-boost hypothesis suggests that the joke creates arousal to a pleasant level, and amusement stems from the short-term maintenance of this rewarding level of arousal. The arousal-jag hypothesis proposes that the joke body induces a massive and aversive increase in arousal, which is then rapidly reduced as the punch line is understood. The reduction of arousal is pleasant and is thus the source of humor.
Incongruity in Humor

smiling and not simple, hip between 1970) found a (1975) and evidence, inverted U-shaped suggests that incongruities and U-shaped

appreciation. The arousal-jag hypothesis is very similar to the view expressed by Freud (1960):

We should say that laughter arises if a quota of physical energy which has earlier been used for the cathexis of particular physical paths has become unusable, so that it can find free discharge. (p. 147)

The arousal-boost and arousal-jag hypotheses make different predictions regarding the relation of amusement to degree and timing of arousal. The arousal-boost hypothesis predicts an inverted U-shape relationship, with a modest level of arousal evoking the maximum amount of amusement. Moreover, the amusement should start prior to the punch line. The arousal-jag hypothesis predicts a positive, linear relation between amusement and arousal, and the amusement should start after the punch line reduces the arousal.

Some researchers have attempted to examine empirically the relation between arousal and humor. Goldstein, Harman, McGhee, and Karasik (1975), for example, monitored the heart rate and skin conductance of individuals presented with riddles and with nonhumorous problems. Riddles are jokes that begin with a question that cannot be answered correctly to which, following the respondent's incorrect response, the actual answer is provided. For the nonhumorous problems utilized in this study the answer was also provided after the respondents failed to answer the question. Although heart rate accelerated with the question and decelerated with the answer for both stimuli, there were significant differences for skin conductance in the conditions. This research suggests that autonomic responses may be different for jokes than for ordinary problem-solving situations. Part of the difference may be due to the affective loading that occurs with the presentation of humorous material—that is, people expect to enjoy the riddle. According to Goldstein et al. (1975), an additional possibility for the difference in physiological response to riddles and problems may be that the temporal dimension in the two situations may be fundamentally different. Intuitively, problem solving takes longer, and there is some evidence to suggest that the resolution of a joke must occur rather rapidly in order for the joke to be perceived as humorous.

Goldstein et al., in a test of Berlyne's arousal hypothesis, correlated arousal changes and absolute arousal with funniness ratings of riddles. Humor appreciation was found to be the greatest for subjects who showed a modest amount of arousal. This was consistent with the arousal-boost hypothesis. Wilson (1979) suggests that the best way to distinguish the arousal-boost and arousal-jag hypotheses would be to examine the temporal relation of arousal and amusement. We know of no study, however, that
has done this. Hence, while a number of studies (Borlyne, 1960; Godekowski, 1976b; Goldstein et al., 1975) suggest that arousal may be necessary for humor, they do not tell us how much arousal maximizes humor or whether the increase in arousal or the subsequent reduction in it causes humor.

Thus far we have distinguished theories of humor that emphasize cognitive mechanisms from those that emphasize affective mechanisms. Eysenck (1942) warned, however, that the distinction between the two categories must not be carried too far. In practice, cognitive and affective responses are necessarily bound together, and a synthesis of cognitive and affective mechanisms may be necessary to advance our understanding of humor. Clearly, the tension or arousal experienced in humor has a cognitive basis. Hence, an examination of the cognitive correlates of arousal also needs to be made. The magnitude of a physiological response tells us little about the conscious reasoning or unconscious reactions that may have contributed to that response.

The cognitive and affective theories depict a process in which there is a disturbance to an existing state, some action to counter or reduce that disturbance, and some effect that occurs as a result of that action. Some theorists (e.g., Batson, 1952; Fry, 1963) suggest there may even be a rapid oscillation between alternative meanings as the individual seeks to resolve or reduce incongruity. Each of the theories suggests, either implicitly or explicitly, that joke resolution is accompanied by a change in the cognitive representation (or meaning) of the concepts involved in the joke. None of the theories, however, has demonstrated empirically the change in the cognitive configuration following joke resolution.

**HUMOR IN TIME AND SPACE**

The consideration of time as a key factor in the cognitive response to humor has been explored minimally. On the interactional or interpersonal level, comedians have long recognized the importance of timing for their success on stage. Eastman (1936, p. 318) suggests that the superior comic takes the audience the “optimal distance down the mental path, where the surprise is the most unexpected,” before springing the punch line. Eastman tells how one particular comedian, Bill Robinson, made jokes out of time itself:

He starts a simple tapping rhythm... and gets the audience going with him in that rhythm, and just as they are going "good," with a most dexterous suddenness, he stops and leaves them in the air. And they laugh. That is what a joke is... getting somebody going and then leaving him in the air. (p. 318)
Incongruity in Humor

As stated earlier, in the humor response, time is theoretically very important. Incongruity theories frequently refer to stages; tension release theories discuss a buildup and sudden release of tension; drive theories refer to an imbalance and a motivation to restore equilibrium, followed by a return to balance; arousal theories discuss the buildup of arousal and its consequences. In addition, some theories have described the spatial configuration of the cognitive processing of humor over time. For example, Koestler (1964) describes the comedic situation as a buildup of expectations accompanied by mounting tension. Then there is an explosion as the individual perceives

a situation or idea, \( L \), on two self-consistent but habitually incompatible frames of reference, \( M_1 \) and \( M_2 \). The event \( L \), in which the two intersect, is made to vibrate simultaneously, on two different wavelengths, as it were. While this association lasts, \( L \) is not merely linked to one associative context, but bisected with two. (p. 35)

Koestler's representation of the cognitive explosion is represented in Figure 3.1.

Koestler suggests that higher forms of sustained humor (such as humorous narratives, comic poems, or satire) rely on "a series of minor explo-

![Figure 3.1](image-url)

Figure 3.1. Koestler's bisociation theory of humor represented in a multidimensional space. Reprinted by permission of the Sterling Lord Agency. Copyright © 1964, 1975 by Arthur Koestler.
This e.
the com after ye.
This is your reading.

COMMUNICATION REVIEWS AND COMMENTARIES

Figure 3.2. Koestler’s representation of sustained humor in a multidimensional space. Reprinted by permission of the Sterling Lord Agency. Copyright © 1964, 1976 by Arthur Koestler.

Paulos (1980, pp. 94-95) also discusses a “mental oscillation” in the joke response:

Thus as a story is being told, the listener oscillates between 1) following it, taking it seriously, and thereby becoming annoyed, and 2) responding to the metacues, and thereby realizing the story is make-believe, and becoming deflated. This phenomenon of mental oscillation partially accounts for the pleasant tension associated with good joke-telling, theatre and play.

Paulos models humor according to catastrophe theory (see Figure 3.3) in three-dimensional space. In the cognitive configuration there is a double layer surface in the center, representing simultaneous meanings of a joke concept. When the punch line of a joke is delivered, there is an interpretation switch, accompanied by a “catastrophic drop” from the upper level to the lower layer or surface. Paulos suggests that the shape of the surface explains the importance of timing. The comedian must know where on the surface the meaning might be located; if the second interpretation becomes too obvious too soon, the joke fails. If the comedian is ahead of the audience, the interpretation switch might not occur.
The representations proposed by Koestler and Paulos have been proposed only theoretically; there has been no empirical evidence to warrant the acceptance of such spatial-temporal representations in the development of a more generalized theory of humor. However, two relevant temporal studies of the cognitive processing of humor have been reported in the literature. Wilson’s (1979) findings suggest that the effect of timing on humor depends on the level of incongruity. In particular, for highly incongruous jokes, greater pauses between joke body and punch line led to decreased funniness.
In another temporal study of the cognitive processing of humor, Goldstein (1970) examined the relation between the affective stimuli in joke material (cartoons) and the time it takes subjects to respond to the cartoons presented. Results suggest that response time is shortest for positively evaluated stimuli. It increases with the complexity of the material and increases at a faster rate for nonsense cartoons than for sexually oriented or aggressive cartoons.

We have suggested, first, that although theories of humor have traditionally been predominantly cognitive or affective in orientation, a synthesis of principles common to both approaches may be necessary in order to describe and explain either type of response. Next, several assumptions common to both kinds of theories were discussed. The first assumption common to many cognitive and affective theories is that the response to humor is a dynamic process, occurring in stages and/or over time. Key elements of the process appear to be incongruity, resolution, and time between the presentation of the incongruity and joke resolution. While cognitively oriented theorists have suggested that a change occurs in the cognitive representation of joke concepts from presentation of joke body to joke resolution, this has not been demonstrated empirically. Affectively oriented theories have proposed that there is a change in arousal or tension from presentation of joke body to joke resolution, but the relations among arousal, tension, and the humor response are still unclear.

The role of timing warrants additional research. Many theories suggest that the length of time between presentation of the joke body and presentation of the resolution is critical in determining the magnitude of the humor response, yet few studies have examined this relationship specifically. One study (Wilson, 1979) suggests that there may be an optimal time (neither too short nor too long) that produces the greatest magnitude of humor. This optimal time may depend on the amount of incongruity introduced by the joke body. A second study (Goldstein, 1970) suggested that the complexity of a joke affects the time it takes an individual to respond to a joke, with more complex jokes involving longer response times. The next sections will describe several MMDS models and then will discuss their relevance in the development of a theory of humor.

METRIC MULTIDIMENSIONAL MODELS OF COGNITION

Metric multidimensional models of cognition have been developed by Woelfel and his associates (see Woelfel & Saltiel, 1978; Woelfel, Cody, Gillham, & Holmes, 1980; Woelfel & Fink, 1980) and have been developed
The Structure of Cognition

Metric multidimensional models of cognition all assume that (1) concepts can be represented as being located in a multidimensional cognitive space, and (2) the greater the perceived dissimilarity between two concepts, the greater the spatial separation between them. Such spatial models of cognition have been criticized by Tversky (1977) and Tversky and Gati (1982) on the grounds that empirical data on perceived similarities among concepts sometimes fail to satisfy the axioms of a Euclidean space (axioms that appear to be satisfied by physical space). Of the greatest relevance to us is violation of the triangular inequality. Such violations hold that (1) the greater the perceived dissimilarity between two points A and B, the greater the perceived dissimilarity between A and C; (2) the greater the perceived dissimilarity between points A and B, the greater the perceived dissimilarity between B and C; and (3) the greater the perceived dissimilarity between points A and B, the greater the perceived dissimilarity between A and C. This latter axiom, which is akin to the law of the excluded middle (much as Einsteinian general relativity is tied to the principle of the equivalence of mass and energy), is a fundamental assumption in representing structures in multidimensional space.

Violations of the triangular inequality are especially important for understanding cognition in general and for humor in particular. Woelfel and his associates (Wolfeil & Fink, 1981) and a number of other researchers (Wolfeil & Barron, 1982) acknowledge violations of the triangular inequality but show that such configuration can be represented by means of imaginary dimensions (much as Einsteinian general relativity turns time as an imaginary dimension).
tion in space-time). A space with imaginary dimensions is called a “warped Reimannian manifold.” Such a space is drawn in Figure 3.4.

An alternative approach to such representation assumes that the space itself is Euclidean but that the distances reported may be misinterpreted. The spatial structure is typically created from the interconcept distances on the assumption that either the concepts are points or the distances between concepts are distances from the center of one concept to the center of the other. Kaplowitz and Fink (1983), however, suggest that concepts are best represented as regions that encompass all of the meanings and attributes of a concept (hence regions of different size) and that reported distances may not be center to center. In particular, when subjects report d(A,B), they may be reporting distances from different points within the region defined by concept A than the point within concept B that they use when measuring d(A,C).

Whether the space is genuinely warped or only apparently warped is not, however, crucial to our current concerns. What concerns us is the cause and the consequences of violations of the triangle inequality. It appears that violations of the triangle inequality tend to occur when we examine concepts that contain two or more highly dissimilar attributes or meanings. As an example of the former, Tversky (1977) notes that

Jamaica is similar to Cuba (because of geographic proximity); Cuba is similar to Russia (because of their political affinity); but Jamaica and Russia are not similar at all (1977, p.329).

Thus Jamaica and Cuba would be located close together, Cuba and Russia would be located close together, and Jamaica and Russia would be located far apart. As an example of the latter, Woelfel and Fink (1980) report that the distance between Red and Tangerine is greater than the distance between Red and Orange plus the distance between Orange and Tangerine.

The Dynamics of Cognition

A number of researchers have used spatial models of cognition as the basis for building dynamic models. Woelfel and his associates regard the message “A is like B” as setting up a cognitive force pushing concepts A and B toward each other in the message recipient’s cognitive space. Kaplowitz and Fink (1982) distinguish two different models of this force. Variable mass-impulse models are proposed by Woelfel and his associates and are based on two assumptions: (1) The force created by a message acts like a brief impulse; and (2) messages implicating a concept add inertia to the concept, making it more difficult to accelerate in the future.

Kaplowitz and Fink (1982, 1983) give most of their attention, however, to the development of a spring model. This assumes that a message linking
led a “warped”
that the space
nterpreted.
distances on
ences between
ments or differences of
attributes of
distances may
A, B), they may
ion defined by
nly warped is
is the cause
It appears that
ne concepts
ings. As an
uba is similar
ussia are not
uba and Russia
uld be located
) report that
r distance be-
Tangerine.
ognition as the
is regard the
cepts A and
Kaplowitz
force. Variable
icates and are
avages acts like a
ertia to the
ion, however, message linking

concepts A and B sets up a springlike connection between them. The
strength, or restoring coefficient, of this connection is assumed to depend on
the recipient's view of the source and the degree to which the message is
unexpected. Among the important consequences of this model are that (1)
spings created by previous messages serve to anchor a concept to its cur-
rent location; (2) cognitions may oscillate prior to coming to rest at equili-
rium; and (3) stretching or compressing these springlike connections in-
creases the tension within them. This cognitive tension may well be related
to the physiological tension we experience.

In sum, a number of potential models might be utilized to represent
the cognitive dynamics of humor. We assume that the cognitive response to
humor may be explicated by a multidimensional cognitive pattern. Further,
it is proposed that joke messages are like other messages that have been
modeled with multidimensional cognitive theory; thus joke messages
ought like forces to bring together concepts that are not necessarily
viewed as similar. By mapping the response to a joke in a multidimensional
space as that response occurs over time, we can see the motion of concepts
as a result of the cognitive processing of humor.

Relevance of MMDS Models
for the Study of Humor

The first section of this chapter discussed several basic principles com-
mon to many humor theories. One of these principles is that the response to
humor occurs in stages and/or over time. Although several researchers
have proposed spatial representations of such a process (e.g., Koesler,
1964; Paulus, 1980), there has been no empirical test of the dynamic as-
pects of the response to humor. We propose that the use of MMDS models
would enable us to see the change in a cognitive configuration in space and
over time.

Returning to the MMDS models, recall that concepts are assumed to
have location and mass in a cognitive space. Thus the concepts that are
presented in a joke body can be mapped with regard to how far apart (or
dissimilar) they are. Following the spring model, we could consider a joke as
a force that establishes connections between the concepts implicated in the
joke body. Many jokes rely on the use of words that have multiple meanings
(as, for example, puns), and the joke body usually suggests which meaning
of the word is to be understood. As the individual seeks to make sense out of
or to solve the joke, there may be accompanying tension in the cognitive
springs. Concepts implicated in the joke body may thus be pulled closer
together. With the joke resolution, the concepts may change their location
with respect to each other, as there is a sudden realization that another
meaning of the joke’s focal concept is to be understood. The joke resolution

572
may also lead to a released tension in the springlike connections between the concepts.

The location and motion of concepts may occur in a Euclidean space, or it may involve a warping of the cognitive space. The next section of this chapter will discuss models that might represent the cognitive processing of humor. All of these models are testable, given that the location and motion of concepts prior to the presentation of a joke body, following the joke body presentation, and following the joke resolution can all be determined and the violation of the triangle inequality assumption can also be measured.

A second issue we discussed indicated that timing, or the amount of time between the presentation of the joke body and the joke resolution, may be important in determining the magnitude of the response to a humorous message. Theorists and comedians alike have focused on suddenness as a prerequisite for humor. This suddenness has been described variously as a “sudden release,” a “sudden reversal,” a “sudden recognition,” or a “sudden realization.” Holland (1982), who considers playfulness and suddenness the principal conditions for humor, states: “As early as the 16th century people began to point to the suddenness, unexpected, and surprise as indispensable prerequisites of laughter” (p. 32). Further, Holland suggests that “joking gives us pleasure by converting a conscious process full of energy and tension into an easy unconscious process” (p. 49).

Morreall (1983) suggests that humor involves a conceptual shift and that time can be related to that conceptual shift in the following way:

Suddenness in a psychological change, then is a function of the amount of change (the difference between the earlier state and the later) and the time over which the change takes place. For a sudden change there must be a relatively large distance between the two states, and the time separating the states must be relatively short. If the time is short, but the change is small, then the change is not sudden, for the person can assimilate what is happening. And if the change is great but the time is also great, here too the rate of change is slow; the person can adjust smoothly to what is happening, and there is no “jolt.” Knowing in advance that a certain change is about to take place, as we said, has the same effect in reducing suddenness as spreading the change over more time, for the person who is expecting a certain change has already started to adjust to it. (p. 49)

Recall that both Paulos and Koestler spatially represented the sudden change in the cognitive system due to the humor response. Koestler’s (1964) representation is the “explosion” of the joke concepts on two planes simultaneously (see Figure 3.1). Paulos’s (1980) representation was a “drop” from one surface to another (see Figure 3.3). However, neither Paulos’s nor Koestler’s representation provides a mathematical framework in which to examine the existence of the explosion, the drop, or any other sudden change quantitatively.
by aged readers whose faces are turned in the past, might be any of the immediate consequences of the situation. The reader is expected to be familiar with the underlying principles and to be able to apply them to the specific case at hand. The reader is expected to be able to understand and interpret the data and to make appropriate decisions based on the results.

Regarding the uncertainty in the difference between the actual and expected results:

between the predicted and expected results, the uncertainty in the difference between the predicted and expected results is not quantitatively assessed. The theoretical framework for predicting the difference is based on an assumption that the predicted and expected results are equal. However, the actual results may differ from the predicted results due to various factors such as measurement errors, variability in the data, and other unknown influences. The uncertainty in the difference between the actual and expected results is estimated by calculating the variance of the difference between the predicted and expected results. The variance is calculated based on the distribution of the differences between the predicted and expected results, which is assumed to be normal. The confidence interval for the difference between the predicted and expected results is calculated based on the variance and the sample size. The confidence interval provides a range of values within which the true difference between the predicted and expected results is expected to lie with a certain level of confidence.
stated by Ronald Reagan but not if stated by some other political leader.

Another parallel between attitude change and humor studies is that both consider the effects of prior messages. For example, Kaplowitz and Fink (1982) suggest that information received (and remembered) prior to the reception of a message serves to strengthen the anchoring of the concepts. In humor research, Wilson (1979, p. 66) suggests that a "priming message" may serve to vary the level of the incongruity of a message. To understand the role a priming message plays, consider that some jokes (particularly those involving puns) juxtapose two separate and apparently inconsistent meanings of a focal concept. The joke builds up the expectancy that one of the meanings is to be considered. The punch line then indicates that the other, previously unconsidered meaning is being implicated. Wilson (1979, p. 66) created four levels of initial incongruity through the manipulation of expectancy. Subjects were primed with either synonyms of words in the joke, synonyms of words in the punch line, synonyms of words in both the body and punch line, or with neutral words. Wilson reported that the initial level of incongruity had "an extremely potent" influence on the funniness ratings; a curvilinear relation between the funniness ratings and the initial level of incongruity was found.

The work of the attitude change theorists and the humor researchers suggests that the inclusion of a message, such as a priming message, prior to the presentation of a joke message may strengthen the expectation that one or another meaning is to be understood from the joke body. A priming message will be defined as a message that serves to move or locate an implicated focal concept toward an expected meaning or interpretation and away from the unexpected meaning as revealed by the joke punch line. As an example, consider the following joke:

Q: Why do mice have small balls?
A: Because very few of them can dance.

A priming message is one that would locate or move the concept "balls" (the focal concept) toward the concept "testicles." An example might be "that guy has really got balls." The priming message should significantly affect the cognitive response to humor.

This section has described several multidimensional models of cognition; one of these models, the spring model, might aid in the description and explanation of the cognitive response to humor. This model assumes that concepts have location and mass in a cognitive space, and that messages act like forces to move concepts in that space. We propose that the concepts implicated in a joke (such as a pun) can be located in a multidimensional space, and that the parts of a joke (joke body and joke resolution) will act
like forces to change the location of those concepts with respect to each other. Thus the use of an MMDS model will allow us to test the implications of both cognitive and affective approaches to humor; we may evaluate the cognitive dynamics of humor by looking at these changes over time in models that make explicit assumptions as to the cognitive forces involved in these changes.

The next section further develops the use of MMDS by proposing different models that might be used to examine the cognitive dynamics of humor.

THE APPLICATION OF MMDS MODELS TO HUMOR

The previous section suggested that the concepts within jokes could be mapped in an MMDS space. If we consider the messages presented in the joke body and joke resolution as forces that cause these concepts to change their location, we could examine the cognitive representation of joke concepts prior to the presentation of a joke \( t_0 \), after the presentation of a joke body but before the presentation of the punch line \( t_1 \), and after the presentation of the joke punch line \( t_2 \).

Suppose that with the telling of the joke body, concepts implicated in the joke body move from some initial positions in the cognitive space \( P \) to a new location in the space \( P_2 \). After arriving at the new location, however, the joke resolution is presented and the concepts are pulled to a different location \( P_3 \), which could be the same as \( P_1 \). Thus there will be a difference in the location of concepts from time \( t_0 \) (prior to joke body), \( t_1 \) (following presentation of the joke body but prior to the presentation of the joke resolution), and \( t_2 \) (following presentation of the joke resolution).

This leads to the presentation of several hypotheses common to all of our models.

First, because the message in the joke body acts like a force to move concepts closer together, if the focal concept is a word with more than one meaning (as in a pun), we hypothesize as follows:

\[ H_1: \text{The presentation of a joke body (} t_1 \text{) will cause concepts implicated in the joke body to move closer in the cognitive space than they would be in the absence of the joke body presentation (} t_0 \text{).} \]

The punch line serves as a force to move other concepts closer together. In particular, the focal concept in a joke is moved from the location implied by the joke body to the unexpected location implied by the punch line. Thus we hypothesize:

\[ H_2: \text{The presentation of a joke punch line (} t_2 \text{) will cause concepts implicated in the joke punch line to move farther apart in the cognitive space than they would be in the absence of the joke punch line (} t_1 \text{).} \]
COMMUNICATION REVIEWS AND COMMENTARIES

H_2: The presentation of a joke body and punch line (t_2) will cause concepts implicated in the joke punch line to move closer in the cognitive space than they would be in the absence of a punch line presentation (t_1).

Third, if we assume that each joke component acts like a force to pull concepts closer together, a priming message will also serve to move concepts (those implicated in the priming message) closer together than they would be located in the absence of a priming message. Stated formally:

H_3: The presentation of a priming message will move the concepts implicated in the priming message closer together in the cognitive space than they would be in the absence of a priming message.

Finally, we need to consider the effect of adding a priming message to a joke body. It is possible that a priming message moves the concept with double meaning implicated in the joke (the local concept) even closer to the "wrong" meaning implied by the joke body than it would be in the absence of a priming message. This leads to a fourth hypothesis:

H_4: The presentation of a priming message with a joke body will cause concepts implicated in both messages to move closer together in the cognitive space than they would be in the absence of a priming message.

Next, presuming again that the motion of concepts does occur, we need to consider the effect of this motion on humor. Recall that the humor literature suggests that there is a suddenness requirement for humor to occur. The models that will be presented next consider how this suddenness produces a humorous response.

The Feint Model of Humor

All of the versions of the feint model consider that the joke body may serve as a feint, or trick. This moves concepts implicated in a direction that the punch line shows to be the wrong direction.

The distance variant. This variant of the feint model assumes that the joke body causes humor by increasing the distance between a focal concept and the meaning implied in the joke resolution. This variant has implications for the relation between the time from the joke body to the joke resolution and the amount of humor. Too short a time would not allow for the concept to reach the wrong place and hence should reduce the humor associated with the joke by reducing the distance between the right and wrong cognitive places. Too much time might also reduce the amount of humor, if during the delay other forces (such as those provided by the cognitive responses...
due to the individual's attempt to resolve the joke) pull the focal concept toward P₃ (the resolution position). Stated formally as a hypothesis:

H₅: The magnitude of humor associated with a joke will be curvilinearly related to the size of the time interval between t₁ and t₂; the curve will have an inverted U-shape.

The priming message should not reduce the humor, but if hypothesis 4 is supported, then the priming message might increase the humor in that the distance between the wrong and right places should be even greater. This suggests the following:

H₆: There will be more humor associated with a joke when a priming message precedes it than when a priming message does not; this humor should occur more rapidly than in the absence of a priming message.

The acceleration variant. This variant of the faint model suggests that it is the acceleration the moving concepts exhibit that serves to create the condition for the humorous response. First, assume that shortly after the presentation of the joke body, the concepts are moving at some velocity. Second, assume that with a long delay, or with a priming message, the focal concept has come to a rest (or close to rest), as opposed to having just arrived at the "wrong" place (the meaning implied by the joke body) and still moving in the "wrong" direction. Since acceleration is defined as the change in velocity over time, a focal concept that is still moving in the wrong direction will require a greater acceleration in order to head in the opposite direction than will a focal concept that has been at rest. A small delay between t₁ and t₂ will also result in a lower magnitude of humor, since the focal concept has not yet had time to accelerate enough in the wrong direction. If the priming message causes the focal concept to arrive at its new location earlier, that concept may be at or near rest at the time of the punch line. Hence, contrary to the distance variant, which predicted (H₅) that the priming message will increase humor, this variant predicts the opposite.

H₇: There will be less humor when the joke body is preceded by a priming message than when there is no priming message.

The jerk variant. A third variant of the faint model is the jerk variant. "Jerk" is the first derivative of acceleration; this version suggests that the rate of change in acceleration causes the humor. Assume that with a long delay, or with a priming message, the focal concept has come to a rest, while with a shorter delay the focal concept is still, as a result of the force from the
message, accelerating in the wrong direction. The shorter delay (without a
priming message) will result in a greater jerk, and hence a greater level of
humor, than if a long delay or a priming message were given.

In both the acceleration and jerk variants, but not the distance variant,
the maximum amount of humor arises from a sudden reversal in direction.
These variants lead to an additional hypothesis:

\(H_0\): The greater the angle between the \(t_1 - t_2\) motion of the focal concept and the
\(t_2 - t_3\) motion of the focal concept, the greater the magnitude of humor
associated with a joke. (The angle will be restricted to be between 0° and
180°).

The three variants of the joint model have very different implications as
to the optimal time lag. The distance variant suggests that enough time be
allotted for the focal concept to make the entire journey to its new location,
P_2. The acceleration variant suggests that the joke resolution should occur
when the focal concept has reached its maximum velocity in the wrong di-
rection (i.e., on its way toward P_2). For an undamped oscillatory trajectory,
this maximum is reached halfway between the two extreme points of the
trajectory, which are P_1 and P_2. (For a damped system, the maximum is
reached before the halfway point.) For the jerk variant, humor is maximized
if the joke body occurs when acceleration in the wrong direction is at a max-
imum. This occurs as the joke body is just beginning to move the focal con-
cept toward P_2. (For equations and graphs that suggest why the points that
have just been made are true, see Kaplowitz & Fink, 1982.)

In short, of the three variants, the distance variant suggests the greatest
time lag for optimal humor and the jerk variant suggests that the optimal
time lag is extremely small, a prediction that makes the jerk variant seem
rather questionable.

The Resonance Model

Juxtaposing the fact that timing is important to humor with an oscillat-
ing spring model of cognition suggests another possible process. Humor
may be maximized if the punch line resonates with the oscillatory motion set
up by the joke body and/or priming message.

The joke body, we assume, sets up a spring that moves the focal con-
cept toward a new location P_2. The focal concept is assumed to be con-
nect ed to other springs, which anchor it to its original location, P_1. After
moving toward P_2, it will oscillate about its new equilibrium and, hence,
there should be a point at which it is moved back in the direction of P_1.

Let \(t\) moves the joke body outside from P_2. Hence, the optimal
How model? In velocity osci-
llating in the acceleration are de-
defined by the maximum dis-
stance between the joke body
from its initial location. This
spring system is expected to
be the resonant frequency \(f_0\).

\(H_0\): The optimal point
This hypothesis

Warp Model

A third logical shift of the spatial
realization warping o
relevant.
Let us suppose, consistent with the faint model, that the punch line moves the focal concept in a direction that is close to 180° from the direction in which it was moving as a result of the joke body. Resonance occurs if the outside force is in phase with the oscillations that are already in process. Hence, resonance is obtained if the punch line comes just at the point at which the focal concept is instantaneously at rest and is changing direction.6

How does the resonance model compare with the variants of the faint model? In this model, humor is maximized if the punch line occurs when the velocity of the focal concept is zero and the focal concept is already accelerating in the direction in which the punch line will pull it. Hence, unlike the acceleration or jerk variant, for maximum humor, the velocity and acceleration are decidedly not at maximum values when the punch line occurs. This model is somewhat similar to the distance variant, in that the point of maximum distance is the point at which the focal concept will begin to oscillate back toward P1. Unlike the distance variant, however, this model does not predict that the increase in distance, which might result from a priming message, enhances humor.

This model does, however, have certain implications as to the relation between incongruity and timing in humor. The optimal delay between the joke body and the punch line is the time it takes for the focal concept to go from its initial location, P1, to the point at which it starts to oscillate back. The time it takes is one half of a period of oscillation. The period of oscillation of a spring system is a decreasing function of the restoring coefficient of the spring. Kaplowitz and Fink (1982, 1983) moreover, propose that more unexpected, disconfirmatory messages produce springs with smaller restoring coefficients. Hence, we have hypothesis 9:

H9: The more the joke body disconfirms prior expectations, the larger is the optimal time delay between the joke body and the punch line.

This hypothesis is consistent with Wilson’s (1979) findings about the effect of incongruity and timing on humor.

Warp Models

A third set of models are warp models, which propose that the psychological shift that results in humor is due to the sudden warping or dewarping of the space defined by the concepts. We believe that warping results from the realization that the relevant concepts have double meanings and dewarping occurs from the realization that only one of these meanings is relevant.
Just as arousal theories differ as to whether humor is caused by the increase in arousal or the subsequent decrease, so warp models vary as to whether humor results from the creation of, or the disappearance of, the warp. Below we present several variants of the warp model, which differ as to the shape of the space at different points of time in the joke process.

**EEN variant of the warp model.** This proposes that the space is Euclidean both before and after the joke body but becomes Non-Euclidean with the resolution of the joke. It is based on the assumption that prior to the resolution, only one meaning of the focal concept is being considered. With the presentation of the resolution, however, the individual suddenly sees two meanings for the focal concept at once. Again, utilizing the mice joke, recall that the cognitive space warps when there has been a violation of the triangle inequality assumption that

\[ d(A, B) + d(B, C) \geq d(A, C). \]

After the presentation of the resolution, the distance between testicles (T) and dance (D) will be great, but the distance between testicles (T) and balls (B) will be small, and the distance between balls (B) and dance (D) will be small. Thus,

\[ d(T, D) \geq d(T, B) + d(B, D) \]

and the only way this can be represented on a planar surface is if the space warps. In this model, if this warping is sudden, the message is perceived to be humorous.

Hypotheses associated with the warp model can be stated as follows.

First, the existence of the warp can be investigated:

**H12.** When a joke body and its resolution are presented, there will be a greater violation of the triangle inequality for the joke-relevant concepts than when a joke body is presented without a resolution, or when neither joke body nor joke resolution has been presented.

Since we have stated that the sudden warping results in humor, we may also hypothesize about the magnitude of that humor. As Morreall (1983, p. 79) suggests, suddenness is a function of the relatively large distance between two states and the relatively short time separating them. Thus:

**H11.** The greater the rate and size of the violation of the triangle inequality for joke-relevant concepts, after the resolution, the greater the humor experienced.

**ENE variant of the warp model.** The second warp variant is the ENE (Euclidean, non-Euclidean, Euclidean) variant. This model proposes that concepts are normally located in a Euclidean space, but with the presenta-
tion of the joke body (without resolution) the force of the message causes
the focal concept to be viewed ambiguously, creating a non-Euclidean space
(one in which we have violations of the triangle inequality). Next, the joke
resolution reduces the ambiguity, forcing or directing thought to one particu-
lar meaning of the focal concept, and the cognitive space is once more
Euclidean. Thus the space could be described as Euclidean at \( t_0 \) (prior to
joke body), non-Euclidean at \( t_1 \) (following joke body but prior to joke reso-
lution), and Euclidean at \( t_2 \) (following joke resolution). This suggests the
following hypothesis:

\[ H_{12} \quad \text{Joke-relevant concepts will be located in a non-Euclidean space following}
\text{the presentation of a joke body and prior to the presentation of a joke reso-
lution, but in a Euclidean space at other times.} \]

Further,

\[ H_{13} \quad \text{The greater the change in the size of the violation of the triangle inequality}
\text{for joke-relevant concepts between joke body and joke resolution, the}
greater the magnitude of humor.} \]

**NEN variant of the warp model.** This variant suggests that the space
begins as non-Euclidean, become Euclidean at \( t_1 \) (with joke body presenta-
tion), and returns to non-Euclidean at \( t_2 \) (with joke resolution). For example,
perhaps concepts are ordinarily located in a space that takes in all meanings
of a focal concept simultaneously. The joke body then serves to reduce the
ambiguity. Using the mice joke as an example, we would find that balls,
testicles, and dance are represented in a non-Euclidean space prior to the
presentation of the joke body. The joke body then directs thought to one
meaning of the word “balls” (testicles), and the resulting configuration is
represented in a Euclidean space. Presenting this as a hypothesis:

\[ H_{14} \quad \text{When a joke body is presented, there will be a smaller violation of the trian-
gle inequality than there will be without the joke body presentation.} \]

The resolution, however, reminds the receiver that there is more than one
meaning for the focal concept (here, the word “balls”); thus the joke-rele-
vant concepts return to a non-Euclidean space. If this change is sudden, the
receiver will find the joke humorous. In any case, the degree of warp associ-
ated with the joke resolution should be greater than that associated with the
joke body. The NEN model is consistent with hypothesis 10 presented ear-
lier for the EEN model, and with hypothesis 13; however, the NEN variant
postulates that the change in the size of the violations of the triangle inequality is
Table 3.1
Comparison of Some Different Warp Models as to the Geometry of the Cognitive Space Represented

<table>
<thead>
<tr>
<th>Model</th>
<th>Before Joke Body (t₁)</th>
<th>After Joke Body (t₂)</th>
<th>After Resolution (t₃)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EEN</td>
<td>Euclidean</td>
<td>Euclidean</td>
<td>non-Euclidean</td>
</tr>
<tr>
<td>ENE</td>
<td>Euclidean</td>
<td>non-Euclidean</td>
<td>Euclidean</td>
</tr>
<tr>
<td>NEN</td>
<td>non-Euclidean</td>
<td>Euclidean</td>
<td>non-Euclidean</td>
</tr>
</tbody>
</table>

in the opposite direction from that proposed by the ENE variant. Table 3.1 lists the warp models that have been proposed and shows how they differ with regard to the geometry of the cognitive space at different times in the joke's unfolding.

**SUMMARY AND FUTURE RESEARCH**

This chapter has presented some of the assumptions and hypotheses that are shared by both cognitive and affective orientations to humor and that have yet to be tested empirically. One of these assumptions was that the response to humor is dynamic, and that the timing between the presentation of the joke segments (i.e., joke body and joke resolution) is important in determining the magnitude and type of response to humor. We next discussed MMDS models of cognition and suggested ways in which these models might be utilized to specify and test some aspects of humor. Next, we developed three models that might be utilized to represent the cognitive dynamics of humor. These were the feint model, which had distance, acceleration, and jerk variants; the resonance model; and the three variants of a warp model.

We have raised several questions for future research by posing hypotheses derived from our MMDS models. Below are some additional questions that have been posed regarding humor and that await translation into our MMDS models; space precludes our elaboration here.

(1) We have discussed humor of a particular kind—that induced by incongruity. To what extent may all humor (e.g., sexual, aggressive) be so modeled? Can all differences in humorous stimuli be modeled using the tools we have provided?
Incongruity in Humor

(2) How may we represent the cognitive processing involved in attempting to get the joke? Is this cognitive Brownian motion, or is something more systematic occurring?

(3) What are the developmental implications of our model? Children's humor appreciation is different from that of adults; how may the models reflect this?

(4) How can these models be used to explain the process of humor creation, as opposed to humor appreciation?

(5) Do the models successfully reflect individual differences in humor appreciation? The spring model implies that all individual differences reflect differences in concepts, their locations relative to other concepts, strength of linkages to these other concepts, and cognitive friction. Can all individual differences be translated into these elements?

Considering the cognitive processing of humor in a multi-dimensional framework allowed us to develop precise and clear models. It is possible that one of the models fully explains humor. It is possible that humor results from some combination of the processes discussed. It is possible that one of the process models best explains the humor response in one kind of joke and another model best explains humor in other kinds. And it is possible that none of the models discussed is ever correct. But in any case, we have now moved beyond verbal theory and have constructed testable alternatives by using a physicalistic analogy. More research is needed, and so is more theory (e.g., full development of a tension release model of humor). We hope we have contributed to the latter.

NOTES

1. Although humor and laughter have been of interest for many centuries, only within the present century have empirical studies been reported, and most of these have been within the last twenty years. Prior to this time the study of humor has been basically philosophical in its approach.

2. Although purported connections between communication functions and humor have been made, evidence to support such assumptions is as inconclusive as the basic humor theory research. For example, although politicians and advertisers often suggest that humor enhances persuasion, research suggests that humor does not significantly influence persuasion (Markiewicz, 1974).

3. What we now call disconfirmation was called discrepancy by Kaplowitz and Pink (1982).

4. This assumes that the joke is "given."

5. Note also that both the variable mass impulse model and the spring model suggest that priming messages should make the focal concept more resistant to acceleration—the former by changing its mass and the latter by increasing its connections to other concepts.

6. In an empirical test of a spring model of cognition, Kaplowitz, Pink, and Bauer (1983) found the period of oscillation to be 13.5 seconds. Half of this is 6.75 seconds, which is not far from some of the longer time lags used by Wilson (1979).

REFERENCES


Incongruity in Humor


