

THE FORMS OF CARICATURE: PHYSIOGNOMY AND POLITICAL BIAS¹

MITCHEL GOLDMAN and MARGARET HAGEN

The analysis of techniques of caricature as a system of communication for transmitting information about pictorial subjects has in the past remained primarily the province of artists and their historians, for example, Ashbee (1928), Lynch (1927), Berger (1952), and Rother (1966). More recently, however, perceptual psychologists interested in the history and function of art have turned their attention to the psychological mechanisms which must underlie the successful perception of the caricatured subject. Gombrich (1961) has stressed the role of the creative imagination of the observer in the successful perception of caricatures, while E. J. Gibson (1969) and J. J. Gibson (1954, 1971) have emphasized the crucial role of the information about the subject carried by the caricature itself. J. J. Gibson (1954) argued that caricature was a combination of the techniques of geometric projection and artistic convention. He wrote that it was necessary for the artist to go beyond the projective information about the subject given in the light coming to the eye and to impose certain agreed upon conventions of exaggeration and distortion in the production of a successful caricature. However, Gibson revised this original formulation in a later definition of "picture": "A picture is a surface so treated that a delimited optic array to a point of observation is made available which contains the same kind of information that is found in the ambient optic arrays of an ordinary environment" (1971:31).

Of course, Gibson is referring here only to representative art, to art whose object is the recognizable depiction of objects and scenes from the natural environment. Clearly his definition of pictures excludes the very large class of nonrepresentative, or abstract, art and, indeed, is not even intended to account for all the variables which determine the aesthetic character of a work of art. The purpose of the definition is simply to establish the nature of the optical correspondence between representative pictures and the scenes they depict. The concept of information central to the above definition is Gibson's (1966) major subject matter.

Mitchel Goldman is a recent graduate of Boston University. He wrote this paper as a distinction work senior honors thesis at Boston University under the supervision of Dr. Margaret Hagen. Dr. Hagen is an Assistant Professor of Psychology at Boston University whose research interest focuses on problems in the area of visual perception, particularly the development of pictorial perception.

THE SENSES CONSIDERED AS PERCEPTUAL SYSTEMS

In general, J. J. Gibson has defined optical information about an object as the formless and timeless invariants in the structure of an optic array that specify the distinctive features of the object. E. J. Gibson (1969) has dealt extensively with the concept of distinctive features, which she defines as invariant relations that specify a particular object and allow for its discrimination from a field of similar objects. Distinctive features are relations, ratios, proportions, and gradients which remain invariant across specified sets of transformation. Thus, in this theory, a caricature is successful to the extent that it preserves the distinctive features essential to the discrimination of the particular subject across the exaggeration distortion transformations of caricature. This concept of invariant information or features preserved across transformations is clearly adapted from geometry and has been formalized by J. J. Gibson in his theory of ecological optics.

Invariant Features in Caricature

Perkins (1974) was the first psychologist to begin an analysis of the techniques of caricature within the framework of the theory described above. He hypothesized that caricature recognition is identical to the process of ordinary facial recognition and that caricatures must therefore contain the same attributes as the caricatured face itself or photographs of the subject. On the other hand, as Worth (1977) observed, one might argue that there are grammars and conventions of caricature recognition that are different from those of facial recognition in real life. That is, it may be the case that recognizing pictures demands some of the properties we need to recognize objects in real life, but it may also demand many other things. Still, Perkins's hypothesis that feature correspondence between picture and subject is the core of successful facial recognition in caricatures can be tested by the straightforward process of looking at existing caricatures and their subjects. This is exactly what Perkins did. Through informal observation of caricatures and photos of Richard M. Nixon and through the deletion of various facial attributes from those pictures, Perkins found that the four key properties of caricatures of Nixon's face were jowls, a hairline with bays on either side, a box chin, and a long nose. The omission of these properties or a contraindication of any one seriously degraded the representative character of a caricature. If Perkins is right about the critical nature of these facial properties for the success of a caricature, it should follow that (1) all artists generally use what Perkins calls "the rather necessary" key properties and thus are consistent among themselves in the nature of their depictions; (2) that any individual artist should be consistent in his depiction across time; and (3) that changes over time in the form of the caricature should be a function of changes in the face itself.

The present authors, while acknowledging the insightful and provocative nature of Perkins's exploratory investigation of caricature, take issue with several aspects of his

analysis. First, Perkins has quite freely taken liberties with the concepts of invariants across transformation postulated by the Gibsons. He chooses to speak of "individuating 'properties' or 'attributes' of the face," a translation which we feel may well do violence to the essential relational component of distinctive features. Consideration of faces in terms of feature ratios would more truly have retained the flavor of the relational concept. Second, Perkins offers no objective analysis of the data, and the reader is left to rely on Perkins's own informal observations. Third, Perkins provides no objective evidence concerning the influence of the individual stylistic bias of the caricaturist. Finally, in analyzing how the requirement of recognition influences caricature, he does not address another important issue: how the political climate also exerts an influence.

An Empirical Study of Caricature

In consideration of these points, the purpose of the present investigation was to carry out a systematic empirical test of the hypotheses which follow from Perkins's argument, with particular attention to the hypothesis of consistency within a single artist and that of consistency across artists and time. We wished to test two aspects of consistency: consistency in specific *features* chosen for exaggeration and consistency in *degree* of exaggeration. We hypothesized that any single artist would be consistent with himself, that artists would not be consistent with one other because of variable stylistic bias, and that there would be a lack of consistency across time as a function of shifting political climate. By shifting political climate we mean variation in the degree to which a public figure is evaluated positively and negatively by the public and by the media. Richard Nixon provides a very clear example of a public figure who experienced an increasingly negative political climate from his election to the dénouement of Watergate. So, like Perkins, we have selected caricatures and photographs of Nixon as our data base. This choice will also facilitate comparisons between Perkins's work and the present investigation. Also, we have chosen as our units feature ratios of property magnitude rather than single properties.

The data for the consistency analysis were obtained in the following manner. Five photographs of Nixon from 1973 were measured by two independent judges, and eleven mean feature ratios were obtained, such as length of jowl/vertical head dimension (Perkins's jowl property), width of jowl/length of jowl (box chin), and length of nose/vertical face dimension (long nose). (See Figure 1 for a full presentation of feature ratios.) The particular feature ratios chosen do not exhaust the possibilities but were selected because they seemed to represent obvious candidates for distinctive features and were easily measured. (The larger of the two dimensions was always the numerator.)

The interjudge correlation for these measurements was extremely high ($r = .99$). Then, through a search of news and political magazines and periodicals from 1973, 100 caricatures from 17 artists were obtained. The only con-

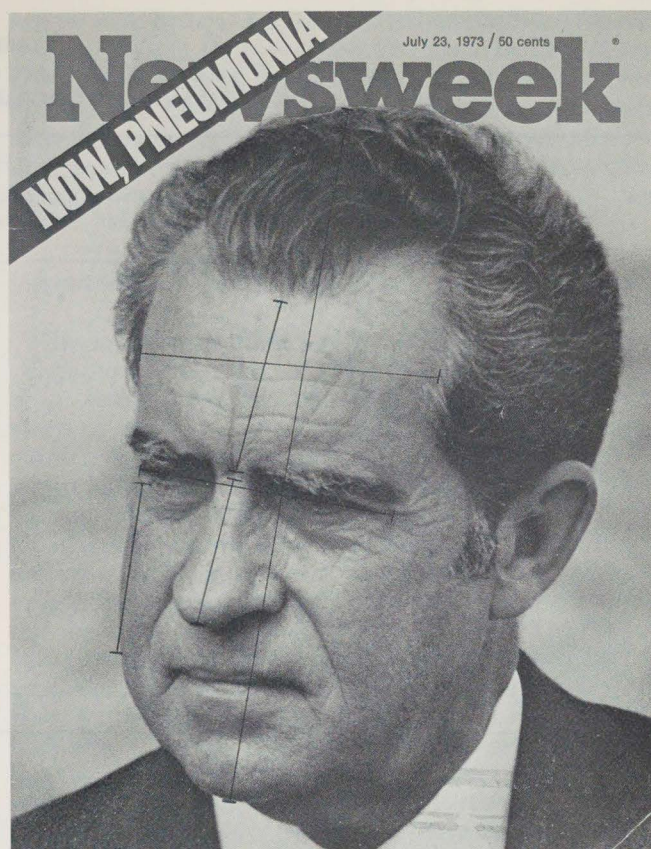


Figure 1 —Cover photograph of Richard M. Nixon with measured feature magnitudes indicated by bars. (Newsweek, July 23, 1973.)

straints on selection were clarity of reproduction and measurability in terms of size. From these 100 caricatures, one caricature was randomly selected for each of the seventeen artists. The eleven feature ratios were again obtained for each drawing. For each feature ratio in every caricature a deviance score was obtained, expressed as the percent deviance from the mean photographic ratio. Thus, if the caricatured feature ratio of width of jowl/length of jowl equaled the mean photograph ratio, the deviance score for this ratio for this drawing was 0 percent. Then, for each artist the 11 feature ratios were ranked from 1 to 11 in terms of least to most deviance from the photographed ratios. These data are tabulated in Table 1. In order to illustrate the technique, we will take the first artist, Cummings, as an example. As one can see from the table, Cummings modified the relation length of nose/length of jowl least of all in his drawing, relative to the magnitude of the relation measured in photographs of Nixon. Thus this relation is assigned a rank of one (1). On the other hand, the relation length of jowl/width of jowl was distorted to the greatest extent in the drawing relative to the photograph, thus receiving the rank of eleven (11). This procedure was followed for all the feature ratios for all the artists.

TABLE 1
RANKS OF DEVIANCE OF FEATURE RATIOS
FROM MEAN PHOTOGRAPHIC RATIOS, BY ARTIST

Artist	Feature Ratios										
	Vert. Head Len. Nose	Vert. Head Len. Jowl	Len. Nose Len. Fore.	Len. Nose Len. Ear	Len. Jowl Wid. Jowl	Len. Nose Eye Eye	Vert. Head Len. Fore.	Len. Nose Len. Jowl	Vert. Head Wid. Fore.	Vert. Head Len. Ear	Vert. Head Eye Eye
1. Cummings	9	10	5	8	11	7	6	1	3	4	2
2. Behrendt	11	10	9	7	3	8	6	5	2	4	1
3. Marlitte	9	11	10	6	2	7	8	1	4	3	5
4. Whitte	11	10	9	8	4	7	6	1	3	5	2
5. Staskyl	10	7	9	8	11	5	6	1	3	4	2
6. Oliphant	9	10	7	8	11	6	4.5	1	2	4.5	3
7. Herblock	9	8	7	9	11	6	3	1	4	5	2
8. Wright	5	11	10	9	6	4	8	3	2	7	1
9. Lurie	8	3	7	5	11	10	1	9	4	2	6
10. Simpson	10	11	8	9	5	7	3	2	4	6	1
11. Hill	11	3	8	10	9	6	2	7	5	4	1
12. Engle	10	11	9	8	4	7	6	1	5	3	2
13. Davis	4	10	2	5	3	9	7	11	6	1	8
14. Haynie	10	4	9	8	11	7	3	5	2	1	6
15. Fisher	10	7	9	8	11	6	5	4	1	3	2
16. Scrawls	11	10	9	8	5	7	6	4	2	3	1
17. Sanders	11	9	10	8	3	7	5	6	4	2	1
Mean Ranks	9.4	8.6	8.1	7.7	7.3	6.8	4.9	3.6	3.4	3.4	2.7

Statistical Analysis

Three types of analysis were performed with these rank data. First, a Kendall coefficient of concordance was computed, $w = 0.597$, $X^2(10) = 101.43$, $p < .001$, indicating a very high degree of consistency among artists in terms of which feature ratios are chosen for exaggeration. Second, a Friedman analysis of variance for ranked data was performed and $F = 83.13$, $p < .001$, indicating that the mean ranks for feature ratios differ significantly from chance ranking. Last, in order to isolate which feature ratios differed significantly from one another in rank, a parametric analysis of variance with multiple post hoc comparisons was performed. The main effect for rank was significant, $F(10, 170) = 19.2$, $p < .02$. Newman-Keuls post hoc comparisons indicated that of the 55 possible comparisons among feature ratios in terms of rank, 31 of these comparisons differed from one another in rank with $p < .05$ at least. This very high number of significant differences is another indication of the high level of consistency among artists in their selection of which features to caricaturize.

Interpretation of the Analysis

The three types of analysis allow us to conclude several things about our consistency hypotheses. First, we have shown that there is a very high level of agreement from artist to artist in terms of which feature ratios are to be distorted in the caricatures. Generally speaking, a feature ratio greatly distorted by one artist will also be greatly distorted by the others, and a ratio little modified by one artist is relatively untouched by the others. Of course this

is not true in each and every case, but statistically the level of agreement is overwhelmingly significant. Second, we have shown that the relative degree to which a particular feature ratio is distorted in caricatures is very stable from artist to artist. Individual feature ratios tend to stay in the same ranked position as one goes from artist to artist. The most distorted tends to remain the most distorted, the least, the least, and those ratios in between tend to maintain constant positions in the ranking. Although this is but another way of saying there is very high agreement among artists, the interpretation goes even further: The agreement across artists on what to distort even extends to this finer level of analysis of individual feature relations.

INDIVIDUAL FACIAL FEATURES

A comparison of high- and low-ranked feature ratios also allows for the specification of feature as independent from feature ratio as the source of exaggeration. It should be noted, however, that the specification of a particular feature as a source of distortion always implies underlying feature ratios. That is, a long nose is long only with respect to other dimensions of the face. A long nose will stand out as a feature per se, rather than a component of a ratio, only if it functions in multiple ratios as a source of high-ranked deviance. Comparisons of pairs of feature ratio ranks provide support for Perkins's argument that the jowls and nose function as major distorted features with good consistency across artists. When these two features occur in the same ratio, the ranked deviance from the photograph is very small (3.6), but when either occurs in

conjunction with another minor feature, such as the vertical dimension of the head, the feature ratio deviance rank is very high (jowls/head 8.6; nose/head 9.4). That this is not due to increase or decrease in the vertical dimension of the head is evident from the low ranks occupied by other ratios with this dimension such as eye-eye/vertical head = 2.7 and length ear/vertical head = 3.5. Such pair comparisons, as well as the three convergent analyses for consistency, provide considerable evidence in support of Perkins's general argument for consistency among artists in terms of features chosen for caricature. It should be noted, however, that while there is significant agreement among artists about *what* to exaggerate there is little agreement about *how much* to exaggerate. Mean percent deviance from photographed ratios is 53 percent but the means for individual artists range from 12 percent for Davis to 86 percent for Fisher. Presumably such variability is due to the artists' individual bias and style.

Statistical Analysis

The related questions of consistency within an artist's work and across time were addressed by similar types of analysis. For these types the data base was generated by multiple caricatures done by five prolific artists in 1972 and 1973. The five artists were selected solely on the basis of relative frequency of published drawings observed in the initial sampling of 17 artists (see Figure 2). They were Oliphant, with seventeen drawings; Herblock, with nine; Wright, with twelve; Lurie, with ten; and Haynie, with five. For each of these five artists for both years, mean percent deviance from photographs was computed and ranks assigned to the features ratio from 1 to 11 for least to most deviance. To address the question of self-consistency, Spearman rank order correlation coefficients were calculated for each of the five artists and the following rho values were obtained: Oliphant: $r = .964$, $t(10) = 10.8$, $p < .001$; Herblock: $r = .75$, $t(9) = 3.4$, $p < .01$; Wright: $r = .855$, $t(9) = 4.94$, $p < .001$; Lurie: $r = .44$, $t(9) = 1.48$, $p < .10$; Haynie: $r = .97$, $t(9) = 12.6$, $p < .001$. Thus, of the five artists tested, four showed significant correlations between 1972 and 1973 rank orders for feature ratio distortions, $p < .05$. Only Lurie failed to show this significant correlation. Consequently, grouped mean ranks for 1972 and 1973 also are positively correlated, $p = .957$, $p < .01$, reflecting consistency of treatment from year to year. However, whereas the rank orders of feature ratios, in terms of relative degree of deviance from photographed ratios, are highly correlated between 1972 and 1973, the absolute size of the percent deviance or distortion increases from 1972 to 1973. Of the eleven feature ratios listed, only one does not change. Of the ten that change nine increase in distortion relative to the photographed ratios, and only one decreases. The results would occur by chance with $p = < .011$, so this nearly uniform increase in distortion reflects a very real change in degree of caricature. In order to test the significance of the magnitude of this change, a t-test was performed on percent of distortion of each

feature ratio by year relative to photographs from the two years, and $t = 3.18$, $p = < .01$. The mean percent distortion for 1972 was 56 percent; for 1973, 61 percent. Thus the increase noted in the sign test is also a significant increase measured parametrically, that is, not only the directionality of the changes but their size is also significant at $p = < .01$.

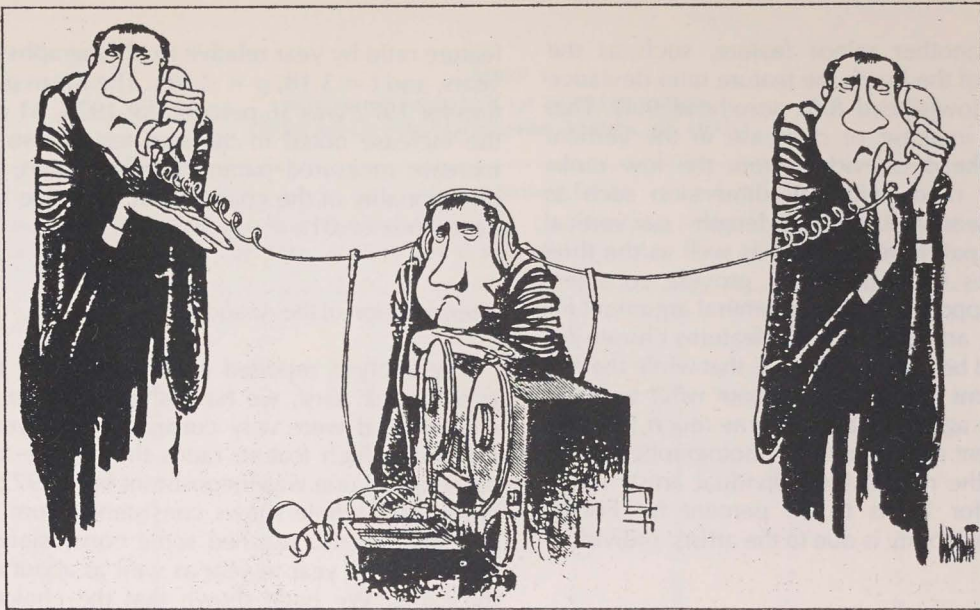
Interpretation of the Analysis

The analysis reported above allows for the following conclusions: First, we have shown that four of the five artists tested were very consistent from year to year in terms of which feature ratios they chose to distort. Because only Lurie was inconsistent from 1972 to 1973, the group as a whole shows consistency from year to year. Second, we also reached some conclusions about what changes from year to year as well as about what remains the same. We have shown that the choice of *what* to exaggerate is quite constant across artists and across time, but we have also shown that the *degree* of distortion varies considerably from artist to artist and from year to year. We found that the mean degree of distortion varies from 12 percent for Davis to 86 percent for Fisher. In addition, when we looked at all five selected artists together, we found that of the ten mean feature ratios which changed from 1972 to 1973, nine of them *increased* in the degree of distortion relative to photographs.

This increase in distortion could be due either to alteration in political climate (which underwent a very rapid change vis-à-vis Nixon from 1972 to 1973) or to real physical changes in Nixon's face due to aging, strain, or fatigue. In order to control for the effect of this type of change, photographs from both 1972 and 1973 were measured and compared for physical change in feature ratio. Of the nine feature ratios which increased in the caricatures from 1972 to 1973, only three were observed to change in photographs—the three involving jowl length. Measurements of photographs indicate a 30 percent increase in jowl length from early 1972 to 1973. It may be argued, however, that this finding still leaves the six jowlless feature ratios, which increased in distortion from 1972 to 1973, a function of increasingly negative political climate rather than of real changes in Nixon's face. Thus this result runs counter to any simplistic assumptions about caricature as a function of true facial features. The finding with respect to jowl size, however, shows that artists are also very sensitive to real changes in the subject's features.

SUMMARY

The present investigation undertook an empirical analysis of several of the consistency assumptions which seemed to follow from Perkins's model of the caricature process. We looked at the question of consistency of distorted features in the work of 17 artists during 1972 and found a very high degree of concordance among artists in



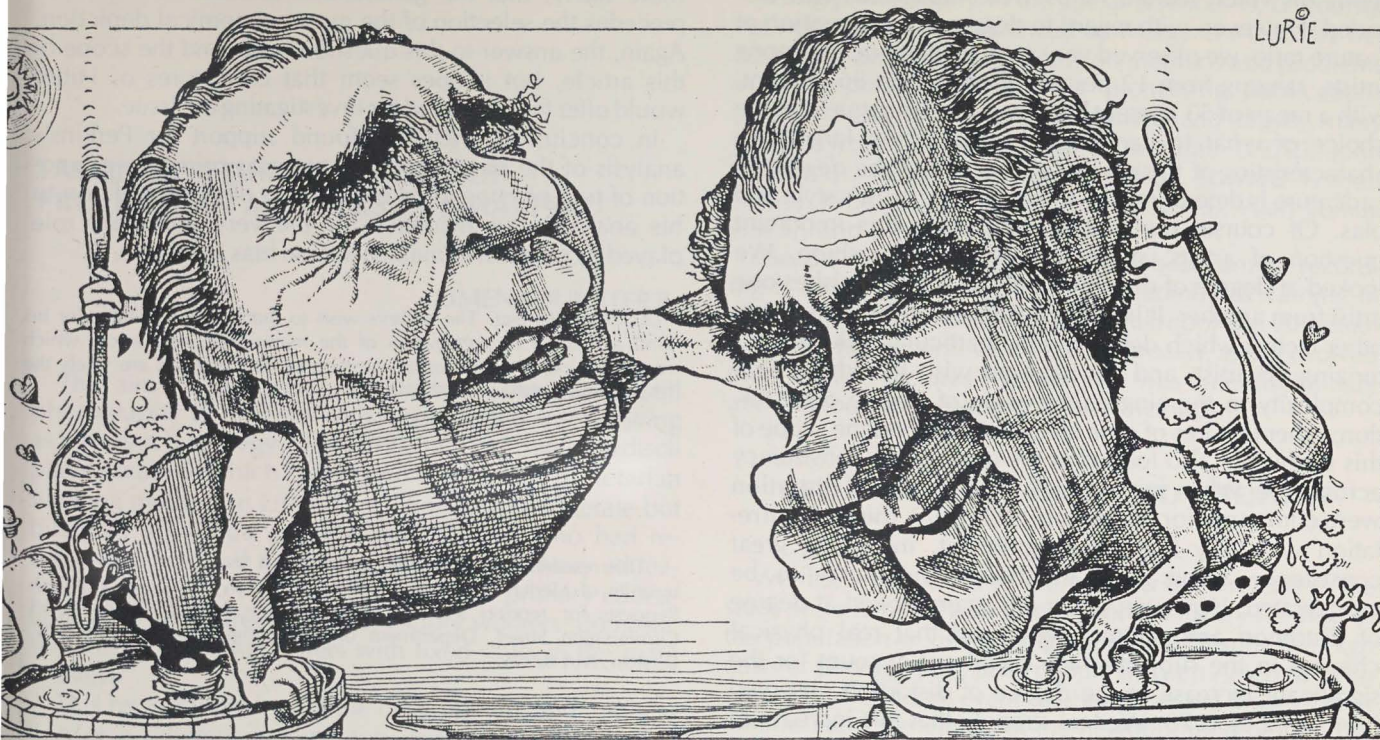
DON WRIGHT—MIAMI NEWS

Time, July 30, 1973



© Herblock in Washington Post
Newsweek, May 7, 1973

Figure 2 —Caricatures of Richard M. Nixon, 1973.



Hanan R. Lurie

Newsweek, June 25, 1973



Oilphant © 1973 Denver Post

A question of priorities: More for the Pentagon, less for the poor?

Newsweek, February 12, 1973



Haynie—Louisville Courier-Journal

Newsweek, November 12, 1973

terms of which feature ratios were most and least distorted. However, with regard to degree of exaggeration of feature ratio, we observed very great discrepancies among artists, ranging from 12 percent distortion to 86 percent, with a mean of 53 percent. Thus we wish to argue that the choice of *what* to caricature is determined largely by characteristics of the subject's face, while the *degree* of caricature is determined by the individual artist's style and bias. Of course, our analysis addresses this important question of artists' styles in a simplistic fashion. We looked at degree of distortion only as it distinguishes one artist from another. It is perfectly clear that there are many other factors which determine the particular style characterizing an artist, and we have no wish to reduce such complexity to the single dimension of degree of distortion; other features of style are simply beyond the scope of this study. We also looked at the question of consistency across time. When features ranked for degree of distortion were compared for the period 1972-1973, the rank correlation was very high and significant, indicating great consistency across time in choice of features to be exaggerated. However, when we again looked at *degree* of distortion across time, we found that real physical changes in the subject's face could not account for the significant increase in the degree of distortion observed from 1972 to 1973. We feel strongly that this increase in distortion was largely a function of the increased negativity of the political climate surrounding Nixon in 1973. It is not possible, of course, to directly test this conclusion, but an indirect method has been suggested by Worth (1977). If our reasoning is correct, that an increasingly negative political opinion increases the degree of distortion in caricatures of a subject, then the opposite of this reasoning should also be true. That is, as the political climate around some public figure becomes increasingly positive, the degree of distortion in caricatures of the subject should decrease and be, at least in principle, testable.

We also do not wish to appear to be arguing that simple exaggeration of feature ratios is all there is to caricature. If this were true, then increased distortion would be the unique and inevitable result of increased negativity of public opinion, which it is not. As Worth (1977) rightly observed, it might also be noted that by the end of World War II, all one needed to caricature Hitler was one diagonal line and one horizontal line underneath, and everyone understood the representation of the hair and the mustache. This increasing economy of line as a function of familiarity cannot be accounted for by the above hypothesis. At the same time it is not clear that increasing economy runs counter to increasing distortion. It may be that the two processes are parallel in time or,

more likely, that the generalized increase in distortion precedes the selection of the most economical depiction. Again, the answer to this question is beyond the scope of this article, but it does seem that caricatures of Hitler would offer fertile ground for investigating the issue.

In conclusion, then, we found support for Perkins's analysis of the process of caricature as primarily a function of true physiognomy but wish to offer, in addition to his analysis, our evidence for the very important role played by political climate, personal bias, and style.

Acknowledgment. The authors wish to thank David Perkins for his kind assistance in preparation of the manuscript. Any errors which remain in interpretation of his position or data analysis are solely the responsibility of the present authors.

NOTE

¹ This research was supported in part by a grant from the National Institute of Mental Health to the second author, #1R01MH27947-0. Requests for reprints should be sent to the second author at 64 Cummington Street, Department of Psychology, Boston University, Boston, MA 02215.

REFERENCES CITED

- Ashbee, C. R.
1928 *Caricature*. New York: Scribner's.
- Berger, O.
1952 *My Victims: How to Caricature*. New York: Harper Brothers.
- Gibson, E. J.
1969 *Principles of Perceptual Learning and Development*. New York: Meredith.
- Gibson, J. J.
1954 *A Theory of Pictorial Perception*. *Audio-Visual Communication Review* 1(13).
1966 *The Senses Considered as Perceptual Systems*. Boston: Houghton-Mifflin.
- 1971 *The Information Available in Pictures*. *Leonardo* 4(27).
- Gombrich, E. H.
1961 *Art and Illusion: A Study in the Psychology of Pictorial Representation*. Princeton: Princeton University Press.
- Lynch, B.
1927 *A History of Caricature*. Boston: Little, Brown.
- Perkins, D.
1974 *A Definition of Caricature and Caricature and Recognition*. *Studies in the Anthropology of Visual Communication*. 2(1):1-23.
- Rother, E.
1966 *Drawing Caricatures*. *Design* 67(4).
- Worth, S.
1977 Private communication to the author.