

This dissertation has been 61-4500
microfilmed exactly as received

LA FAVE, Lawrence, 1929-
HUMOR JUDGMENTS AS A FUNCTION OF
REFERENCE GROUPS: AN EXPERIMENTAL
STUDY.

The University of Oklahoma, Ph.D., 1961
Social Psychology

University Microfilms, Inc., Ann Arbor, Michigan

**THE UNIVERSITY OF OKLAHOMA
GRADUATE COLLEGE**

**HUMOR JUDGMENTS AS A FUNCTION OF REFERENCE GROUPS:
AN EXPERIMENTAL STUDY**

**A DISSERTATION
SUBMITTED TO THE GRADUATE FACULTY
in partial fulfillment of the requirements for the
degree of
DOCTOR OF PHILOSOPHY**

**BY
LAWRENCE LA FAVE
Norman, Oklahoma**

1961

HUMOR JUDGMENTS AS A FUNCTION OF REFERENCE GROUPS:
AN EXPERIMENTAL STUDY

APPROVED BY

Muzaffer Sherif

Norman Jacobson

Maurice K. Tennenbaum

W B Lemmon

P. Teske

DISSERTATION COMMITTEE

ACKNOWLEDGMENTS

I wish to thank my many students who, in various ways, helped this experiment; especially William Maesen and Mrs. Eleanor Lee. Also, I would like to thank Professor Alex Lazur of St. Joseph's College; and Morris Friedell, Dr. Elihu Katz, and Dr. Fred Strodbeck, all of the University of Chicago, for their help in the experimental phase of this study. Also, thanks to Dr. Leslie Singer of the University of Indiana, and Dr. James Yarold of the University of Chicago for statistical consultation.

Thanks are extended to the members of my committee in charge, Drs. William B. Lemmon, Maurice K. Temerlin, Percy T. Teska and Norman R. Jackman. Above all, however, I would like to thank my major professor, Dr. Muzafer Sherif, who patiently taught me both the theoretical approach used, and the experimental technique.

TABLE OF CONTENTS

	Page
LIST OF TABLES	v
Chapter	
I. INTRODUCTION AND PROBLEM.	1
II. EXPERIMENTAL PROCEDURE.	27
III. ANALYSIS, RESULTS AND DISCUSSION	42
IV. SUMMARY AND CONCLUSIONS	67
REFERENCES	69
APPENDICES	72

LIST OF TABLES

Table	Page
1. Sets and Permutations Within Class of 20 Jokes	35
2. Strength of Trends in Percentages by Jokes and Hypotheses	48
3. Results of Predictions by Jokes and Hypotheses	49
4. Sign Test of Significance of Number of Jokes Predicted Correctly.	50
5. Humor Judgments by Catholic Group.	79
6. Humor Judgments by Jehovah's Witnesses Group . .	80
7. Humor Judgments by Agnostic Group.	81
8. Humor Judgments by Baptist Group	82

HUMOR JUDGMENTS AS A FUNCTION OF REFERENCE
GROUPS: AN EXPERIMENTAL STUDY

CHAPTER I

INTRODUCTION AND PROBLEM

What is humor? That question has occupied the minds of such renowned philosophers as Plato, Aristotle, Descartes, Kant, Hobbes, Spencer and Bergson. Yet, despite extensive philosophic speculation, a paucity of psychological experimentation upon humor exists.

Why have psychologists been so reluctant to experiment in this area? One can only speculate as to the answer. Perhaps, however, a little speculation might suggest an approach to humor that could help overcome the present research inertia in this area. Speculation is more likely to be of value, however, if the literature is first surveyed.

Upon doing so, it is difficult not to be impressed by two outstanding tendencies in humor theory and research: 1) the frequent equation of humor with laughter; and 2) the relative predominance of psychoanalytically-based studies. Oddly enough, these two tendencies are not unrelated; those of psychoanalytic persuasion seem especially prone to equate

humor with laughter. For instance, in a recent survey, the psychoanalytically-oriented Flugel (1954) seems to treat humor and laughter as though they were synonyms.

The tendency to equate humor with laughter is not restricted to the psychoanalytically oriented, however. Thus, Bergson's (1911) classic book on humor is actually entitled Laughter; Drever (1952) defines "humour" in terms of laughter.

Such an equation of humor with laughter, though, seems of questionable validity. Even the most cursory introspection should convince the reader that he often laughs for reasons other than because he is humored. Thus, a person may laugh when tickled, embarrassed, afraid, releasing tension, pretending to have gotten the point of a joke which went over his head, etc.

And, just as people need not be humored to laugh, so they need not laugh when humored. A humored man may avoid laughing to keep from embarrassing the butt of the joke, to remain unnoticed, to appear sophisticated, etc. Laughter, then, is neither a necessary nor a sufficient condition of humor.

Thus, even the bravest experimental souls might shy away from humor experimentation, if they made the assumption, currently fashionable, that all instances of laughter must be accounted for by a general theory of humor.

But how else can humor be measured--if not through

laughter? After all, psychologists cannot read minds directly; they must infer a humorous state through some behavioral index.

In some instances laughter is, as shall be shown later, a reasonably accurate guide to humor. However, one cannot safely assume a priori that wherever laughter occurs humor is uncovered. Laughter is only an accurate index of humor under certain conditions--and one task of experimental psychology is to help determine just what those conditions are.

There is, however, another way to measure humor--an approach which seems to avoid certain pits into which a humor theory, based solely on laughter, would fall. This other method would define humor as a mental experience, measurable through judgmental indices. Ss would be asked to reflect judgments as to the humor values of certain "humorous" materials through some behavioral index, such as sorting jokes into degrees of funniness. But here too it must be remembered that, as with laughter, so only under certain conditions are sortings accurate reflections of that mental experience called humor. Fortunately, however, the experimental conditions under which judgmental sortings are a reasonably accurate reflection of mental experience are suggested not only in a few judgmental studies on humor, but a good many judgmental studies in related areas as well.

The present study is concerned not to develop a

general theory of humor, but simply to present a judgmental experiment. The foregoing discussion merely outlines the theoretical perspective into which the present work is embedded. Now that this perspective has been suggested, humor may be defined, for present purposes, as a learned mental experience of amusement towards an object.

The above definition suggests the following characteristics: 1) humor is mental experience, not behavior; 2) this mental experience is learned; 3) the particular mental experience involved is one of amusement; and 4) humor implies a subject-object relationship.

Within this general outline, more specific theoretical considerations led to the present experiment:

The judgmental experiment to be presented is based on the postulate that an individual's attitudes are reflected in, and thus measurable by, his sense of humor. Both attitude and sense of humor are viewed as containing cognitive components of mental experience. An attitude is defined as a conceptually-rooted motive of affect towards an object. Following Sherif and Cantril (1947, p. 92), an ego is defined as a constellation of attitudes. The general construct of humor upon which this research is based is: a joke is humorous to the extent to which it enhances an object of affection and/or disparages an object of repulsion, and unhumorous to the extent that it does the opposite. Objects of affection are viewed as positively related to the ego, whereas objects of

repulsion are seen as negatively related. Humor is enhanced when an object in which the individual is ego-involved (towards which he holds a positive attitude) triumphs in a joke and/or a negatively-related object loses, whereas relatively unhumorous are those jokes at the expense of the object of identification, and/or which enhance a negatively-related object.

The hypotheses tested in the present experiment are somewhat more specific than the general principle of humor mentioned above. The greater specificity is partly a function of the introduction of the concept reference group. This concept was first used by Hyman (1942) who meant, essentially, a comparison group.

Muzafer Sherif early advocated use of the reference-group concept. Sherif's enthusiasm for the construct resulted from a broad perspective, which enabled him to relate the reference-group concept to the social sciences on one hand, and to experimental psychology on the other. By reference group Sherif and Sherif (1956, p. 175) came to mean any group with which S identifies himself psychologically.

Sherif noted that the reference-group concept is necessary in modern society because, unlike the effects of simpler cultures, no longer does modern man's conceptual development result in a one-to-one correspondence between his membership and reference groups. Modern man does not

always identify himself psychologically with those groups in which he revolves physically. Hence, the sociological concept of group, or more specifically, "membership group," no longer encompasses the conceptual horizons, or identifications, of modern man. Since his psychological experiences are central to psychology, so reference group is a psychological, not a sociological, concept (Sherif and Sherif, 1956, pp. 176-177).

Sherif next tied reference group with the sociological concept of group norm. When the individual identifies with a group it becomes his reference group; he internalizes its norms as his own ego attitudes (1956, pp. 628-635).

Thus is seen the social origin of important ego attitudes, which anchor the individual's perceptions, and render predictable his judgments of motivationally relevant stimuli. In a classic experiment, Sherif (1935) had earlier demonstrated how individuals form such norms.

Now that the theoretical base for the present research has been sketched, hypotheses may be formulated. First, however, the relevant literature ought to be surveyed. Such studies reveal how they helped shape the hypotheses and the experimental design which follows later.

Survey of Literature

Modern surveys of the literature, some of which are concerned with humor in a more general way than is the present investigation, are found in Flugel (1954), Valentine

(1942), Ghosh (1939), Perl (1933a), and Piddington (1933). More and Roberts (1957) discuss a few experimental studies and some theoretical work of leading philosophers and social scientists on humor; while Rosenberg and Shapiro (1958) relate various humor theories to marginality and Jewish humor.

The question as to whether or not humor has a social character is relevant to the present research (which is formulated within the theoretical context that the ability of an object to provoke a humorous experience is a function of social learning). It seems appropriate, therefore, to briefly mention a few works concerned with the relationship between humor and social character.

Henri Bergson (1911) may have been, as More and Roberts (p. 234) suggest, the first modern philosopher to emphasize the social character of laughter (or humor). Bergson (1911, p. 2) asserted "that the comic does not exist outside the pale of what is strictly human." And for him, the word human meant essentially what today might be labelled social.

Perl (1933b) found that humor appreciation is subject to social influence. Jokes presented to a group are found funnier than jokes judged in private.

Washburn's (1929) study on children reflects the social genesis of humor; both laughter and smiling, she found, arise from social interaction.

Raley and Ballmann (1957), in noting the metamorphosis of humor through adolescence, also suggested its social character: "The most tentative conjecture based on current findings is that, concerning the matter of humor as well as nearly all other personality traits, there are definite and well defined differences that occur due to adolescent development" (p. 22).

After observing that the sex of the examiner exerted a major influence on their experimental results, Doris and Fierman (1956) remarked upon the social nature of this result. Their results indicated to them "that such experimentation should not focus exclusively on the stimulus materials, i.e., jokes or cartoons, per se. The social context appears to be crucial."

All the above studies, regarding the social character of humor, seem to suggest that jokes do not have absolute stimulating value--that the same joke told under different social circumstances, or to a different person, might produce a different effect.

Eysenck (1947) could discover little, if any, consensus as to the relative humor values of various individual jokes or different kinds of humorous material (cartoons, limericks, verses, etc.). Flugel (1954, p. 726) found this result in accordance with that of Heim (1936), Omwake (1939) and Stump (1939). On the basis of these studies Flugel concluded by questioning the justifiability of speaking "in

general terms of a 'good joke'" (p. 726).

The social character in humor is also reflected by differences in the nature of the jokes which various races tell. For instance, Middleton and Moland (1959) studied the content of jokes told by white and Negro university students. They found that Negroes told more "anti-Negro" jokes based on the traditional stereotype of the Negro than did the whites. Negro students also told about four times as many "anti-Negro" jokes as "anti-white" jokes.

Eysenck (1944) studied the relationship between national character and humor. He often found greater differences in the humor content of journals within nations than he did between nations. (His findings hardly seem surprising in that he studied only modern, Western nations. Sociologists often have observed social character to differ more between socio-economic classes within a modern nation than between people from different nations but of the same socio-economic class.)

More and Roberts also studied the social character of humor. They asked SS from three different socio-economic classes to sort cartoons according to degrees of funniness. The authors found differences between classes to "accord fairly well both with expectations on the basis of the humor theory we have advanced and with various social stereotypes" (p. 241).

Middleton (1959) studied Negro and white reactions

to racial humor. In this study, to be discussed more fully later, he found differences between middle- and lower-socio-economic-class Negroes, but not between middle- and lower-class whites.

An important problem in constructing the present experiment, as will be discussed later, revolved around such questions as whether or not well-known jokes could be used, and just how many jokes should be used. The following studies relate to these problems:

Hollingsworth (1911) found, when the same jokes were presented again to the same Ss, that some jokes "waxed" and others "waned." In other words, the waxing jokes were judged funnier upon repetition, while waning jokes lost some of their humorous effect.

Cattell and Luborsky (1957, p. 407) also raised the question of familiarity with jokes and the effects:

The main design of our experiment . . . limits our investigation of such subsidiary questions as that of the effect of familiarity. We could obtain a partial answer, however, by the following analyses: (a) Correlating individuals' scores on 'number of jokes known beforehand' with scores on 'total number of jokes considered funny.' This . . . yielded a negligible correlation of $-.09$. The result does not necessarily prove that familiarity has no effect on funniness of an individual joke; it indicates only that people familiar with more jokes are not inclined to rate the jokes as a whole as more or less funny than do people less steeped in the lore of humor. (b) Secondly, we could find whether the average funniness rating of the jokes rises or falls between the first and second administration. It fell, from a mean of 24.7 to one of 17.2, a fairly considerable fall.

Cattell and Luborsky (p. 407) also noted Eysenck's

(1943, p. 195) finding that jokes already known to Ss tended to be better liked than others. All this led Cattell and Luborsky to the tentative conclusion:

. . . that the subject remembers, and therefore is familiar with, jokes that he especially likes; but that familiarity forced by circumstances (in this case, rereading all jokes) appreciably reduces the funniness of jokes in general (p. 407).

In a more recent study, Lee and Griffith (1960) confirmed the tentative conclusion of Cattell and Luborsky, i.e., that S most readily remembers jokes which he especially likes. Lee and Griffith, however, were not concerned with the effects of joke familiarity upon humor values. Rather, they were interested in the relationship between forgetting of humor and repression. As they stated (p. 436):

The elusiveness of jokes . . . is attributed by psychoanalysis to repression. Humor, like dreams, slips by the censor, who quickly recloses the door. In this study a large number of New Yorker cartoons were paired randomly for each subject who chose the one of each pair which was the funnier. A month later each subject returned to sort the original cartoons from some new ones mixed in with them. Contrary to what should be expected were repression an active force, subjects correctly identified more of those cartoons which had previously impressed them as funny.

After presenting whole series of jokes, Martin (1905) observed both "fun-fatigue" and "fun-accumulation." She concluded (p. 47):

Without doubt very long series of experiments would show fun-fatigue as a controlling factor, but the influence of one comic picture on another as regards the comic impression, investigated in this and other experimental series, and the experiences of every day life as well as the above results, show that fun accumulation may be a controlling factor in a short series.

A large proportion of humor research deals with personality traits and individual differences. Most of these studies are irrelevant for purposes here. A few are pertinent, however, either because of methodological contributions or side issues discussed therein--or because such studies also deal with social data, though often in a disguised language. Thus, some studies most pertinent to the present research, i.e., More and Roberts, Wolff, Smith and Murray, and Roberts and Johnson, are actually couched in personality terms. A "personality" experiment by Cattell and Luborsky contains interesting side issues. However, as the present work focuses not upon personality experiments per se, these personality experiments are discussed elsewhere in this chapter, under diverse headings.

Validity of Judgment Method

The question concerning validity of judgmental ratings in humor measurement is important; what Ss rate as funny and what they consider funny may differ. Therefore, relevant experimental literature becomes highly pertinent.

Perhaps the oldest and most respected area in experimental psychology is psychophysics. The validity of 100 years of extensive experimentation in this field rests upon the basic assumption that S's judgments can be measured through his ratings. When instructions to S are appropriately given, this assumption seems warranted, at least such appears to be the consensus among experimental psychologists. The

psychophysical experiments by Wolff et al. and Roberts and Johnson, through a process of cross checks, tested this validity assumption.

Wolff et al. tested the validity assumption by means of a hidden motion picture camera which recorded S's facial expressions as he responded to jokes. These facial expressions were found to relate to judgments in such a manner as to leave little doubt, in the Wolff et al. experiment, regarding the validity of the psychophysical method. As Cattell and Luborsky (1947, p. 404) remarked: ". . . Wolff, Smith, and Murray . . . found reasonable indications that elaborate physiological or objective measures give results essentially the same as those of subjective report." Each of the several methods of measuring humor used by Wolff et al. may be viewed as a validity check on the others. (See tables Wolff et al. presented; pp. 352 and 360.)

Having also tested this validity assumption, Roberts and Johnson (1957, p. 61) stated:

A question might be raised concerning the validity of the subjects' verbal ratings to the cartoons. It is possible, particularly with a group of mental patients, that what the subjects rate as funny and what they actually consider funny are entirely different. To check this possibility a comparison was made between the laughter and smile responses made by the subjects to the cartoons and their humor ratings. The correlation between these measures was .87, significant at the .01 level. This indicates that those subjects who laughed and smiled the most gave high ratings to the cartoons. Thus we have demonstrated agreement between overt humor behavior and the subjects' verbal reports.

Psychophysical findings, and the essential similarities

between the present work and those by Wolff et al. and Roberts and Johnson, combined with the care taken to keep Ss from awareness of the real purpose of the experiment, should have insured valid data in the present experiment.

Three Related Experiments

Only three experiments on humor seem closely related to the present one:

The first of these was performed by Wolff et al. These authors (1934, p. 343) were concerned with one particular mirth-provoking pattern or "thema," which is "another object (human being) in a disparaging situation." Their purpose was to determine whether differences between the mirth responses of Jews and Gentiles would occur when jokes at the expense of Jews were presented.

Of the 19 Ss, records are complete for only 15 (six Jews and nine Gentiles). Sixteen jokes were selected for presentation; eight disparaged Jews, and eight were control jokes. Each S was isolated in a room and the jokes were shown on a horizontal aperture. Instructions to Ss were as follows:

Keep as still as you can. In this opening . . . will appear jokes, numbered from one to eight. After each joke a star will appear. When it appears, call out loud your rating of the joke on a scale ranging from . . . (-2) to . . . (+4). The score of -2 = very annoying . . . 0 = indifferent . . . and +4 = extremely humorous. Say these loudly so that I can hear you in the next room (p. 350).

Wolff et al. (pp. 359-360) related further procedure:

To clinch our hypothesis in regard to unaffiliated objects we decided to present jokes which disparaged a race to which none of the Jewish or Gentile subjects belonged. Under these circumstances, it was supposed, there would be no marked differences between the responses of the two groups of subjects.

Since the jokes which we chose for the disparagement of Jews were all concerned with covetousness and stinginess in money matters it was a simple matter to transform them into Scotch jokes.

The authors found, as they had before Jewish jokes were changed to Scottish ones, that control jokes were about equally appraised by Jews and Gentiles. But, counter to hypothesis, the Gentiles appreciated the Scotch jokes more than did the Jews--just as had occurred, according to hypothesis, before the Jewish jokes were transformed into Scottish ones.

Since Gentiles found Jew-disparagement jokes funnier than Jews did, and despite the fact that Gentiles found Scotch-disparagement jokes funnier than Jews did, the authors (p. 361) concluded:

The results of these experiments seem to demonstrate that the disparagement of an unaffiliated object is one type of mirth-evoking situation . . . a positive response to such a joke is an index of negative sentiments in respect to the disparaged object.

A second experiment on humor closely related to the present one was performed by Roberts and Johnson. Their major hypothesis, which they maintained was "derived from the theories of humor proposed by Freud (1917, 1928) and G. H. Mead (1934)," is: "The perceived funniness of a humor stimulus is positively related to the degree to which the perceiver is able to

empathize with the characters depicted in the humor stimulus" (p. 57).

The authors next related their procedure and conclusion:

Two judges rated the subjects' empathic associations to each of the cartoons. They rated each response as to the amount of empathy shown for the characters in the specific cartoon on a scale of 0 to 8, 0 indicating no empathy and 8 very high empathic identification. The interjudge correlations, or reliability coefficients for rating the responses to the 12 cartoons range from .57 to .96 with an average of .86.

As a test of the hypothesis the humor ratings of the subjects and the empathic content of their free associations to the cartoon were correlated. An r of .78, significant at the .01 level, was obtained. This indicates that those subjects who rated the cartoons most humorous gave the most empathic responses to these cartoons (pp. 59-60).

The hypotheses of the present experiment are based upon the above two experiments. After the hypotheses had been formulated and pretested, and the experiment begun, a very closely related experiment by Middleton appeared in the literature.

Middleton experimented on Negro and white reactions to racial humor. Each of 50 Negro Ss were matched to each of 50 white on several personal characteristics. These 100 Ss were then presented 18 jokes as a paper-and-pencil test in a classroom situation. After Ss had rated the jokes, Adorno's F-scale for authoritarianism was administered. Four questions, designed to measure the relative acceptance or rejection of four elements of the Negro stereotype, were adjoined to the F-scale.

. . . Negroes found the anti-Negro jokes quite as funny as did the whites. With regard to the anti-white jokes, however, the Negroes reacted more favorably than the whites. These comparisons must be accepted with caution, however, for the important variable of authoritarianism or ethnocentrism is left uncontrolled, and the Negro subjects in general represent a considerably higher level of authoritarianism (p. 178).

In keeping with his hypothesis, Middleton (p. 178) found that among the highly authoritarian persons the whites reacted more favorably to anti-Negro jokes than did the Negroes. Counter to prediction, though, among those with relatively low scores on the F-scale, the Negroes reacted more favorably to anti-Negro jokes than did the whites.

Middleton's (p. 179) expectations for anti-white jokes were confirmed amongst low scorers on authoritarianism; Negroes responded more favorably to anti-white jokes than whites. In the highly authoritarian group, however, counter to prediction, white-Negro differences on anti-white jokes were insignificant.

Middleton (p. 182) also discovered:

. . . little or no evidence to support the hypothesis that persons who accept the validity of the Negro stereotype react more favorably to anti-Negro jokes based upon that stereotype than do persons who do not accept the stereotype.

As mentioned earlier, Middleton (p. 182) found that social class differences in reactions to humor were significant only among the Negroes. But, he (p. 180) explained:

. . . Individuals were classified as middle class or lower class on the basis of whether their parents were in white collar or blue collar occupations. It should be recognized, however, that among university students even those with lower class backgrounds may take the

middle class as a reference group. Consequently these data should be regarded as exploratory.

Middleton seemed to make an ad hoc observation which pointed up the need for the concept of reference group in humor experimentation. He apparently felt that where the family membership group is no longer the reference group (apparently a frequent occurrence in complex, socially mobile, USA), S's sense of humor may actually have reflected his reference--not his membership-nonreference-group.

Middleton concluded:

Some of the hypotheses based upon the theory of Wolff, Smith, and Murray that individuals tend to laugh at the expense of unaffiliated objects and not at the expense of objects for which they possess positive sentiments are confirmed by this study, but as many more must be qualified or rejected. In gross comparisons Negroes did react more favorably than whites to anti-white jokes, but the Negroes apparently found the anti-Negro jokes just as funny as the white (p. 181).

Despite the crucial way in which the above three studies relate to the present experiment, the first two (Wolff et al. and Roberts and Johnson) are clothed in the language of personality, rather than social, concepts. The Wolff et al. article, however, does allude to the social-psychological concept of group. But it uses the concept inadvertently, and nontechnically. (Such use seems hardly surprising, not only because the experiment focuses on personality, rather than social, concepts, but also because the article was published in 1934--before the concept of group regained fashion.)

Wolff et al. were not primarily interested in

experimenting in social psychology, but rather in personality:

To prove the effectiveness of our thema as a mirth-provoking stimulus or, in other words, to show that a subject laughs at his own success and his neighbor's failure, we should make a quantitative comparison between his response to disparaging jokes of which he was the butt and his responses to similar jokes of which a non-affiliated object was the butt. This would be the ideal experiment. We could not, however, devise a feasible scheme for making a butt out of the subject himself, and so we were forced to resort to an affiliated object for comparison with a non-affiliated object (pp. 344-345).

The authors, then, would rather have avoided group identification, or any other identification, and would have preferred to work with jokes which directly disparaged S himself--so little were they interested in group identification per se.

The experiment by Roberts and Johnson, reviewed above, also centered in personality concepts; it was mainly concerned with the relation of S's "empathic identification" to humor.

Although the three experiments reviewed above, including the social-psychologically-oriented experiment by Middleton, lack a close affinity with the present one as regards technical concepts used, the ideas of all three can be recast into the technical concepts used here. When reconceptualized thus, all three experiments seem to have yielded results consistent with predictions made in the present experiment. And yet, though the findings of these three studies seem sensible, the results of Wolff et al. and Middleton appear to contradict those by Roberts and Johnson.

Perhaps these three studies reveal a sparse conceptual development which, omitting important qualifications, makes the results seem contradictory--even though more adequate conceptual treatment would reveal that these two seemingly contradictory sets of results may be treated as two different aspects of a more comprehensive postulate.

The Wolff et al. results were presented in support of the contention that humor is evoked at the expense of an unaffiliated object, i.e., at the expense of an object with which one does not identify. The inadequacy of their conceptualization is indicated below. After finding a significant tendency for Gentiles to find more humor than Jews did in jokes at the expense of Jews, the authors stated:

To clinch our hypothesis in regard to unaffiliated objects we decided to present jokes which disparaged a race to which none of the Jewish or Gentile subjects belonged. Under these circumstances, it was supposed, there would be no marked differences between the responses of the two groups of subjects (p. 359).

Even though the authors' prediction proved wrong, they reasoned ad hoc:

On the face of it it would seem that since the Jews did not laugh as did the Gentiles at the disparagement of Scots, our formulation for this type of thema was incorrect--that is, in so far as an unaffiliated object was concerned. But are we certain that the Jews do not, or did not in this instance, affiliate themselves mentally with the Scots--at least in respect to their attitude towards money? We noted indeed, that one type of affiliation was based upon resemblance--affiliation by similarity--to the extent that subjects who are conscious of possessing a trait enjoy hearing that trait praised in others and dislike hearing it debased. Now, there is evidence that this type of affiliation did occur in the Jewish subjects (p. 361).

The authors then pointed out that evidence for such affiliation is found both in subjective reports, and also in the fact that Scotch jokes were found relatively funnier by Jewish Ss than Jewish jokes were.

Although Wolff et al. seem correct in their hindsight, and by stating such offered excellent suggestion for future research, they cannot be credited with having conclusively substantiated their hypothesis. The scientific method is more than strained when results contradicting prediction are reinterpreted, by reasoning ad hoc, as substantiating the hypothesis. Even though Wolff et al. may well be correct in their analysis, their mistaken prediction on Scottish jokes necessitates further experimentation before their hypothesis can be said to be verified conclusively.

The cause of their mistaken prediction, then, seems to lie not in a faulty hypothesis, but rather in a lack of sophistication as to the choice of the "neutral" group in testing that hypothesis. Yet the lack of sophistication does not appear primarily with respect to experimental design per se. Wolff et al.'s wrong prediction regarding Scotch jokes looks like a result of embedding their hypothesis in an inadequate conceptual framework. The authors apparently were led to assume mistakenly that any "race" which S was not a member of represented for him an "unaffiliated object." (It will be recalled that Middleton made a similar assumption that "lower class" university Ss would not identify with the

"middle class.")

The use of the concept reference group, discussed earlier, might have prevented such mistakes--suggesting as it does why the predictions by Wolff et al. and Middleton were erroneous. Realizing that Jews and Scotch were both stereotyped as stingy, Wolff et al. conveniently substituted Scotch names for Jewish, received a new batch of Jewish and Gentile Ss, then asked Ss to make judgments of "Scotch" jokes. Apparently, however, the Jewish Ss were all too aware that they too, like the Scotch, were stereotyped as holding a group norm of stinginess. Seeing themselves in the "same boat" as the Scotch, the Jews identified with them--thereby decreasing for those Jews the funniness of Scotch-disparaging jokes. Similarly, Middleton failed to realize, until after his results were in, that college students of lower-class origin might come, through collegiate experiences, to identify with middle-class norms.

Roberts and Johnson interpret their findings as indicative that "empathic identification" and humor go hand in hand. Therefore, their thesis appears in contradiction to those of Wolff et al. and Middleton, who suggest that humor results from a lack of identification with the object. This is not to say that Wolff et al. and Middleton were unaware of the fact that humor judgment is not simply a question of identification. In fact, their comments indicate that they were very much aware that humor is also a function of what

is happening to the object of affection--with whether it is being esteemed or abused. However, unlike the present experiment, these authors were not centrally concerned with what was happening to the object of identification. They dealt with jokes which made objects of identification the butt of the joke, and never with jokes which esteemed the object of identification--leaving findings which inversely related humor to identification.

The relationship between these three experiments, as concerns matters relevant to the present experiment, may be summarized as follows: The Wolff et al. and Roberts and Johnson experiments have the fact in common, and differ thereby from the Middleton experiment, that the latter author used social psychological language whereas the former authors conceptualized in terms of personality theory. In all other respects important here, however, the Wolff et al. and Middleton experiments seem similar. First, as mentioned above, both predicted that humor would result from lack of identification, whereas Roberts and Johnson found humor and identification to be related positively. Second, Wolff et al. and Middleton were aware that, under certain conditions, identification could enhance humor, if the jokes esteemed, rather than disparaged, the object of affection. Roberts and Johnson did not make a point that what was happening to the object of identification did matter. Third, both only partially supported their non-identification-with-object

hypotheses, whereas Roberts and Johnson successfully supported theirs. Fourth, both Wolff et al. and Middleton seemed to recognize ad hoc that their failure to conclusively substantiate their non-identification-with-object hypotheses resulted from failure to realize before the experiment that S may identify with a nonmembership group used in the experiment. (Wolff et al. could not readily conceptualize the fact, because they were writing before Hyman originated the reference-group concept. However, Middleton could and did recognize that an earlier concern for reference-group theory might have substantiated his hypothesis.)

But speculation is one thing, experimentation another. If the logic above is sound, and the theory useful, more successful hypotheses than those by the above authors should be possible. The following hypotheses, derived from the conceptual framework discussed earlier, would erase the apparent contradiction between the hypotheses of Wolff et al. and Middleton on one hand, who predicted that humor would result from lack of identification, and Roberts and Johnson on the other, who found humor to result from identification. The present hypotheses imply that identification enhances humor where the object of identification triumphs, whereas lack of identification enhances humor when the object is disparaged. Thus, the results of all these three studies seem reconciled within

the following hypotheses:

Hypotheses

Hypothesis 1 General (H1G)

Jokes tend to be judged as funny by Ss whose reference (identification) group is esteemed, and whose outgroup is disparaged, and to be judged unfunny by Ss whose reference group is disparaged and whose outgroup is esteemed.

H1G may be broken down into the following two sub-hypotheses:

Subhypothesis 1a (H1a). Jokes which esteem S's reference group tend to be judged as funny.

Subhypothesis 1b (H1b). Jokes which disparage S's reference group tend to be judged as unfunny.

The hypothesis of Roberts and Johnson, while agreeing with the prediction in H1a, would suggest that findings pertaining to H1G would not warrant rejection of the null hypothesis; while results on H1b would warrant rejection of the null hypothesis, but in precisely the opposite direction from which H1b predicts. (That is, jokes which disparage S's reference group, by the hypothesis of Roberts and Johnson, should tend to be judged as funny.)

If H1b were rephrased--predicting instead that "neutral" Ss would tend to judge as funny jokes which disparage an outgroup--the consistency of H1b with the hypothesis of Wolff, Smith and Murray would become apparent.

(H1a and H1G remain irrelevant to the hypothesis of Wolff
et al.)

CHAPTER II

EXPERIMENTAL PROCEDURE

Experimental Controls

The conceptual framework used here, more elaborate than that of previous studies, also called for more controls governing the nature of the stimulus materials. These stimulus materials, consisting mainly of written jokes, in many ways are not comparable to the cartoons used by Roberts and Johnson, who indicated very few criteria as their basis for cartoon selection.

Since Wolff et al. used written jokes, a comparison with their stimulus materials is possible. (No comparison will be made with the stimulus materials of Middleton, as nothing important for purposes here is added thereby.)

Designation of Speaker's Group

As mentioned earlier, Wolff et al. did not use the term group in a technical way; they concentrated upon personality, rather than social, concepts. However, Wolff et al. did choose their Ss from two different ethnic classes, Jews and Gentiles; and most Ss probably held the ethnic class they belonged to as their reference group. Yet, in

designating the actors in their jokes, Wolff et al. did not directly mention the group affiliation of each. Rather, they used personal names which merely implied the ethnic group of the speaker. A more rigorous procedure would not assume the ability of Ss always to relate a speaker's name to his group, but would substitute group names, such as Jew, for personal names, like Ikey.

Dialogue

Wolff et al. were only concerned that one ethnic class be disparaged. Each joke in the present study used not only disparages one group, but esteems another (i.e., makes it appear triumphant). All jokes, then, involve dialogue.

"Serious" Jokes

An attempt was made to select jokes which would be taken seriously. Otherwise, S, thinking the speakers were only joking, might detach himself and become humored at "his own expense"--detachment which would likely hurt the present hypotheses.

Another way to encourage S to take the joke seriously would be by presenting dialogue between enemy groups. (The criteria of the present experiment would not permit S to disparage himself.) A joke at the expense of one's own group may be judged less threatening if told by an ingroup member than by an outsider. For these reasons, only jokes

were used in which dialogue occurred between representatives of two opposing groups. The last line always squelched the opponent, and invariably enhanced the final speaker's own group.

Antagonistic Groups

Of course squelching an "innocent" outsider may not provide ego triumph to the same extent as squelching a member of an antagonistic outgroup. For this reason, Ss were chosen from a population identifying with one of four antagonistic groups. (These four reference groups were selected on the basis of their religious group norms and prevailing social distances between such groups.)

Considering the relative lack of controls by Wolff et al., that they obtained, in such a difficult experimental area, any significant support for their hypothesis may seem surprising. Yet some of their results were significant. They established, through the use of "control jokes," that these differences could not be explained by a general tendency for one group to be more readily humored than another.

Built-In Control

Wolff et al. needed control jokes. With one set of Ss they used only jokes which disparaged Jews, while with another set they employed only jokes which disparaged Scots. Both sets of Ss included only Jews and Gentiles. The Gentiles were never the butt of any experimental jokes. (Unless

some Gentiles were Scotch, but Wolff et al. made no such statement.)

Control jokes were not needed in the present experiment; the control was built right into the experimental design. For every joke which disparaged one group and esteemed the other, another did the opposite. (That is, the design contained both possible permutations.)

For example, suppose a constant tendency exists for Baptists to be humored more readily than Catholics. The tendency would spuriously advantage the hypotheses on pro-Baptist-anti-Catholic jokes. But also it would work to disadvantage by an equal amount on an equal number of pro-Catholic-anti-Baptist jokes. Thus, this extraneous variable, which finds one group more easily humored than another, may be measured, then subtracted out, by a built-in experimental control. This internal type of control is generally considered preferable to the external one of control jokes, which may inadvertently introduce influential new contaminating variables. See Siegel (1956, p. 62). By the same logic, this counterbalanced design is superior to one in which Ss were matched. As Siegel (1956, p. 62) has related:

A matching design is only as good as the experimenter's ability to determine how to match the pairs, and this ability is frequently very limited. This problem is circumvented when each subject is used as his own control; no more precise matching is possible than that achieved by identity.

Education

Since the education variable seemed to need a lower limit, Ss were restricted to a population of high school graduates. Senility, on the other hand, disqualified one Jehovah's Witness.

Nationality

Whether or not a S was reared in the USA also appeared important; Ss were, by and large, limited to long-term residents of this nation.

Social Contagion

Another control concerned the order of each S's class of jokes. Each class was shuffled in an effort to randomize the effects of "social contagion." Shuffling was necessary both because Ss were tested in groups, and to randomize order effects.

Familiar Jokes

Another variable which needed control concerns whether or not each joke had been heard before. As discussed earlier, Hollingworth (1911) found that some jokes "waxed" upon repetition and others "waned." Cattell and Luborsky (1947, p. 407) observed that repetition of jokes reduced their funniness, but that jokes known to Ss prior to the experiment were found more humorous than other jokes. There seems no need to enter into the controversy as to which jokes do what. Of present concern is simply whether or not

a joke has been heard before does effect its humor value, and is a variable in need of control. The best control would have been to eliminate previously heard jokes.

Familiar jokes could be removed by using only original compositions. Unfortunately, the author's limited imagination permitted only half the jokes to be new. Old jokes were changed as radically as possible to prevent Ss from recognizing them. Jokes believed to have "made the rounds," or to be well known, were eliminated. After these three precautions, pretest Ss were asked to sort separately all jokes that they had heard before. Jokes which more than one S reported having heard before were eliminated. Analysis of these pretest results shows that only a small fraction of one percent of jokes currently used were reported as familiar. Nor was any consistent tendency discerned for the same joke, or jokes about the same group, to have been heard before. Through these four precautions, the heard-before variable seems to have been reduced to a negligible one.

Sex

The sex variable showed surprising strength in pretest jokes which contained sexual connotations. How S sorted such jokes was as predictable by his sex as by his religious-group identification. This extraneous variable was controlled by eliminating all jokes which appeared to hold obvious sexual connotations.

Fatigue

Fatigue is another important variable which demanded control. As reported earlier, Martin (1905, p. 47) had observed both "fun-fatigue" and "fun-accumulation." Personal fatigue from reading joke books, and comments from pretest Ss, had also suggested that jokes read in succession, and which are not woven together into a story, soon begin to lose humor value. Under such circumstances, "long-winded" jokes become tedious. It was decided, therefore: 1) to make the jokes as short, simply worded, and easy to follow as practically possible; and 2) to reduce the number of jokes to the minimum required by an adequate experimental design.

Validity

Instructions to Ss disguised the purpose of the experiment. In efforts to insure valid results, a relaxed atmosphere was encouraged.

Attempts were made by the following procedures to insure S that E would not know which results were his: at least five Ss were tested at each session; S was not asked to write his name (E stated that he was only interested in group, not individual, results).

Procedure

Steps In Joke Construction

A three-step procedure was commonly followed in selecting and constructing the jokes to be used: 1) A study

was made of the norms of the religious experimental groups. 2) Norms were chosen which differed for the two experimental groups involved in the dialogue for each joke. 3) An attempt was made to select jokes based on such a difference between norms. For example, the joke regarding the Jehovah's Witness (JW) and flag saluting (Joke 17, Appendix A) was based upon JW and Catholic norms regarding the American flag.

Stimulus Materials

The stimulus materials consisted of an Information Sheet and twenty-five 4 x 6 cards. On five of these cards was typed one of the following: Very Funny, Funny, Indifferent, Unfunny, or Very Unfunny. On each of the remaining 20 cards was typed a joke.

In each of these jokes, dialogue occurred between representatives of two presumably antagonistic religious groups. The last speaker always gave the "punch line," and the victim was always the other speaker. Thus, the group of the last speaker was always the esteemed victor in each joke, whereas the other group was always the disparaged victim.

The class of 20 jokes was divided into five sets of four jokes each, as shown in Table 1. Only four jokes involving JWs occurred because this group was relatively infrequent in the population. More than four JW jokes, it was feared, would both have seemed unnatural, and have aroused suspicion. JWs, Catholics, and Baptists were assumed to identify with Christians--whereas agnostics were not.

TABLE 1

SETS AND PERMUTATIONS WITHIN CLASS OF 20 JOKES

Set	Permutation	Joke	Group Esteemed	Group Disparaged
1	1	4	Catholic	JW*
1	1	17	Catholic	JW
1	2	1	JW	Catholic
1	2	18	JW	Catholic
2	1	14	Catholic	Agnostic
2	1	19	Catholic	Agnostic
2	2	10	Agnostic	Catholic
2	2	20	Agnostic	Catholic
3	1	6	Catholic	Baptist
3	1	13	Catholic	Baptist
3	2	5	Baptist	Catholic
3	2	16	Baptist	Catholic
4	1	7	Baptist	Agnostic
4	1	12	Baptist	Agnostic
4	2	2	Agnostic	Baptist
4	2	15	Agnostic	Baptist
5	1	8	Christian	Agnostic
5	1	9	Christian	Agnostic
5	2	3	Agnostic	Christian
5	2	11	Agnostic	Christian

*JW represents Jehovah's Witness.

This latter assumption was based upon study of the norms of these groups.

In each set of four jokes occurred both possible permutations; one group was esteemed and the other disparaged and vice versa. Further, each group won exactly half the time. For example, Set 1 found the Catholic and JW each victorious twice (see Table 1).

Selection of Ss

Ss were selected on the basis of apparent identification with one of four religious groups: Baptists, Catholics, JWs, or agnostics. An E either selected as Ss those whom he or a trusted ingroup informer considered loyal to one of the four groups. The Information Sheet also asked which group S identified with. Unless S checked the group he was chosen as identifying with, his results were discarded.

Some criteria found useful in the selection of Ss included: 1) group membership; 2) active participation in the group; 3) apparent loyalty or ego-involvement; and 4) checking on the Information Sheet the religion presumed to be preferred. However, S did not need to show characteristics 1 and 2 to be selected. Qualifications 1 and 2 are irrelevant for occasional agnostic Ss who do not belong to agnostic religious groups; there is no "Agnostic Church" per se. However, loyal agnostics were sought who actually did belong to a church in which agnosticism was the norm. Many Unitarian Churches, though not all, satisfied this

condition.

Besides checking the expected group on the Information Sheet, S to be selected needed to check two other items appropriately. First, that he was a high school graduate. Second, S was required to answer yes regarding whether he liked to consider himself a Christian, unless he was chosen as an agnostic, in which case he had to check no.

In this way, a total of 209 final Ss were selected. Included were 51 JWs, 52 Catholics, 52 agnostics, and 54 Southern Baptists (see Appendix C).

Eliminating Ss who did not mark the three essential Information-Sheet questions as expected, furnished reliability checks and probably increased validity.

Validity probably increased by this procedure because: When S gave wrong information on either of the two identification questions at least one of two things was wrong: 1) S gave true information and E had judged S wrongly, meaning S did not identify with that particular group, and his results were invalid; or 2) S gave false information and E had judged S correctly. In the latter case, the method acted similarly to the MMPI "lie scale"--eliminating Ss who were likely to give invalid information on the test itself. (The assumption here is that the S who gave invalid information on one part of the experiment was especially likely to do so on other parts, too. Validity increased by eliminating such Ss.)

Selection of E

To select an ingroup member as E offered better rapport, and eliminated nonloyal Ss. On the other hand, an ingroup E might have caused Ss to feel their loyalty was under test. If so threatened, Ss may have distorted judgments according to E's expectations. The greater Ss' ability to "see through the test," the more likely they would have distorted judgments so as to appear more loyal than they were.

The problems an ingroup E created were resolved by indirect methods, and appropriate instructions to Ss. The feeling of anonymity created by group testing, and anonymous Information Sheets, the indirect method of tapping attitudes, and the misleading instructions to Ss, all were expected to combine to prevent the undesired distortion an ingroup E might otherwise create. Yet, necessary advantages were gained by selecting an ingroup E; namely, better rapport and Ss individually validated with respect to reference groups.

Instructions to Ss

Before presenting the instructions to Ss, the two ways by which these instructions deliberately misled Ss, in order to prevent the distortion and invalidation suggested above, were as follows: First, Ss were told that the experiment was a sociological, rather than psychological, one. This misleading statement was intended to minimize the likelihood that Ss would search for, and possibly find, the real purpose behind this experiment--as they more likely

might have done had they known that the test was concerned with "mind reading." Second, Ss were falsely led to believe that right and wrong answers existed--i.e., that some of the jokes were not really jokes at all. (One purpose here was to ego-involve Ss into gaining a high "sense of humor" score so that they would not conjecture about any more clandestine purpose which the test might, and did in fact, have.)

The instructions to the Ss were as follows:

Take a seat as far away from everyone else as possible.

You are given five label cards. You are to place these five label cards in front of you and from left to right. On your left-hand side put the card which has the words "Very Funny" written on it. Next to this card and to its right place the card which says "Funny." Still further to the right set the "Indifferent" card. The "Unfunny" card should go even further to the right. At the extreme right-hand side should go the last card, the "Very Unfunny" one.

Make sure that there is room beneath all five label cards to put other cards of the same size. These five cards then should read from left to right: "Very Funny"; "Funny"; "Indifferent"; "Unfunny"; and "Very Unfunny." (Note to E.: If blackboard is available, write those five phrases from left to right and remark that this is the correct order as shown.)

You may have noticed that a rubber band is found around each of these five label cards. Leave these rubber bands there for now and I will tell you what to do with them later.

During this sociological experiment, I am going to give you a set of religious jokes. But first, let me tell you about the purpose of this sociological experiment.

We are trying to find out how good a sense of humor most people have. Some of these jokes on religion have been found to be funny by those with a real sense of humor. Other jokes are, strictly speaking, not really jokes at all. No discriminating person would find any

humor in them.

Past experience shows that those who most accurately report just how funny the jokes seem to them usually get the best "sense of humor" scores. So, to do your best, just go by natural instinct in judging how funny the jokes are.

There is no need to put your name down anywhere as we only need to find out how well people do in general, not how you did in particular.

After you have finished, I will give you an envelope to put all your materials in.

I will now pass out to you the set of jokes.

.

You are given a set of jokes on religion. You are to read each joke and then place it under one of those five cards in front of you, depending upon how funny it is. For example, I find this joke to be very funny so I place it under the "Very Funny" card. This next joke doesn't affect me much one way or another, so I put it under the "Indifferent" card. But don't let me influence you. Your jokes are in a different order from my set so judge each joke individually. Go by first impressions as jokes are usually less funny when read a second time. Do not change jokes around after you have sorted them once. Hold up your hand if you have any questions.

.

When you have finished sorting all these religious jokes into their piles, slip the jokes for each pile underneath the label card so that the rubber band goes around the whole pile, like this. (Note to E.: Demonstrate.) Thus, all the jokes in each pile will be surrounded by a rubber band and the top card on each pile will be one of the five label cards. If you have any questions, hold up your hand.

When you finish, hold up your hand and I'll give you an envelope to put your five piles into.

.

We have only one little thing to do and then we'll be through.

I will now pass around an Information Sheet. Fill

out this Sheet. Raise your hand if you have any questions.

.

When you have finished filling out the Information Sheet, hold up your hand and I will check your materials. Do not leave your seat until I have checked all your materials. (Note to E.: Thank Ss for their participation, etc.)

CHAPTER III

ANALYSIS, RESULTS AND DISCUSSION

Treatment of Data

As the hypotheses require only dichotomous data (i.e., non-funny and funny), statistical methods based on assumptions unique to more powerful scales are inappropriate. For this reason, among others, nonparametric statistical techniques were used.

The use of a nonparametric statistic is not quite as disconcerting as it might first appear, however. While power was lost by not using parametric statistics, a nonparametric analysis involves fewer assumptions (cf. Siegel, 1956, p. 32). Since this experimental area is relatively new, it seemed desirable both to minimize the number of assumptions, and to report findings which would give follow-up research as broad an area to work in as possible.

Each hypothesis was given a sign test. However, to test the sign of each joke separately would have failed to account for constant differences between groups in tendency to be humored. Such an extraneous difference could have greatly affected results, however, and ought, therefore, to

have been controlled.

The experimental design had, built into the class of 20 jokes, both sets and permutations. Both possible permutations were equally represented within each set, and an equal number of jokes (twins) occurred within each permutation. Following the logic of analysis of variance, irrelevant differences between groups were eliminated by detecting the magnitude of constant error thus introduced, and subtracting this amount from the result.

In predicting each joke, this constant error was taken out by comparing both groups, not merely with each other, but also with their own means for all four jokes within the particular set involved (as is shown in a formula presented below).

The method was developed appropriate to the following conditions: First, a "none-some" scale with dichotomous data. Second, each joke was given equal weight. Third, each group was given equal weight (though the number in each group was not always equal).

The jokes sorted under five labels were treated as follows: The indifferent label, representing (by definition) the point of origin, was assigned a zero (nonpositive) value. Positive values were given (arbitrarily) to the piles on the humorous side of the origin; while the labels on the unhumorous side were assigned nonpositive values. Thus were the five piles bunched into two categories: those to which

positive values had been assigned, and those to which non-positive numbers had been given. The positive (humorous) grouping included the Very Funny and Funny labels, while the nonpositive (nonhumorous) category involved the Indifferent, Unfunny, and Very Unfunny piles.

The constant error, discussed earlier, was then removed as follows: First, the proportion of Ss within the appropriate group was found who assigned a positive (humorous) value to each of the four jokes within the set. Second, the mean of these proportions was determined for the four jokes within the set. (This mean proportion represented fe in the positive column for the joke in question.) Third, the proportion of Ss within the appropriate group was found for the joke in question. (This proportion represented fo in the positive column for the group in question.) Fourth, the appropriate formula below was applied:

Let G_e represent the group which is esteemed in joke i .

Let G_d represent the group which is disparaged in joke i .

Let G_n represent the two neutral groups, which are neither esteemed nor disparaged, in joke i .

Let P_{Ge} , P_{Gd} , and P_{Gn} represent the proportion of Ss, from G_e , G_d , and G_n , respectively, who assign a positive value to joke i , i.e., who place it under either the Very Funny or Funny label. [G_n invariably involves a combination of two groups.

However, the number of Ss in each of these two groups may not be exactly equal, and yet each group must be given equal weight. Therefore, P_{Gn} represents not the proportion of all Ss in G_n who assigned a positive value to joke i .

Rather, P_{Gn} represents the mean proportion of the two groups; i.e., $P_{Gn} = (P_{Gn1} + P_{Gn2})/2$.]

Let $\overline{P_{Ge}}$, $\overline{P_{Gd}}$, and $\overline{P_{Gn}}$ represent the mean proportion of Ss from G_e , G_d , and G_n , respectively, and with regard to joke i , who assign a positive value to the four jokes within set j .

Then, the following predictions follow for each joke in line with the hypotheses:

from H1G, Formula H1G:

$$(P_{Ge} - \overline{P_{Ge}}) > (P_{Gd} - \overline{P_{Gd}})$$

from H1a, Formula H1a:

$$(P_{Ge} - \overline{P_{Ge}}) > (P_{Gn} - \overline{P_{Gn}})$$

from H1b, Formula H1b:

$$(P_{Gn} - \overline{P_{Gn}}) > (P_{Gd} - \overline{P_{Gd}})$$

According to theory from which the hypotheses were derived, the quantity on the left-hand side of $>$ in H1G should be positive and that on the right-hand side should be negative. If so, the hypothesis predicts correctly for the joke in question. If vice versa, the result would read $<$, and the hypothesis would predict wrongly for that particular joke. However, one side of the inequality may give the

expected sign, but the other side the reverse. It was decided, in such a rather equivocal event (as the inequality indicates), that the hypothesis would be supported for that particular joke if the correct sign were of greater magnitude than the incorrect--and that H1G would be found wrong for that joke if the opposite occurred.

Each joke was tested separately in order to rule out constant error which irrelevant differences between jokes introduced. Each hypothesis was tested by whether the number of jokes it predicted correctly was significantly above chance. (While H1G predicts for all 20 jokes, H1a and H1b each relate to only 16. The discrepancy results in that, since the Christian represents a composite of three experimental groups, Set 5 jokes involve no neutral groups.)

Abstract of Experimental Design

This experiment may be conceived in three major dimensions: (1) N Ss; (2) sorted i jokes; (3) into k categories of funniness. The funniness dimension (3) may be dismissed for purposes here by simply saying that jokes were judged with respect to five categories, and the data were treated on a "none-some" scale.

Each of the other two dimensions may be subdivided into four levels of abstraction: With respect to the joke dimension (2), there was a Class (of 20 jokes), divided into Sets (5), which consisted of Permutations (2), which included Jokes (2, i.e., twins). In decreasing order of abstraction,

there were: Class, Sets, Permutations and Jokes.

Ss too may be conceived in four abstraction levels. Three of these levels are: N (which included all 209 Ss); Groups (each of which ranged from 51 to 54 Ss); and S himself. On four of the 20 jokes a "supergroup" was involved, which was at the second highest order of abstraction. This group was the Christian (consisting of 157 Ss), which included all Baptist, Catholic and JW Ss. In descending order of abstraction there were: N; "Supergroup"; Groups; and Ss.

Results

Tables 2 and 3 show the results of predictions by jokes and hypotheses. All 20 jokes came out as predicted on H1G; 12 of 16 on H1a; and 13 of 16 on H1b. (For reasons given earlier, the four jokes belonging to Set 5 did not show any results on H1a and H1b.)

Table 4 shows the significance of the number of jokes predicted correctly for each hypothesis. Since tests of all three hypotheses were significant at less than the .05 level, every hypothesis was substantiated.

As discussed earlier, subhypotheses H1a and H1b tried to answer certain questions raised by prior experiments. Not only were these subhypotheses expected to be substantiated, but by roughly equal amounts. If this expectation is correct, differences between results obtained by H1a and H1b should not have warranted rejection of the null hypothesis.

TABLE 2

STRENGTH OF TRENDS IN PERCENTAGES BY JOKES AND HYPOTHESES

Joke	Formula H1G	Formula H1a	Formula H1b
1	+25.1	+25.7	- .6
2	+39.9	- .7	+40.6
3	+30.8		
4	+24.1	+ 9.0	+15.1
5	+41.3	+50.1	- 8.8
6	+40.4	+ 7.7	+32.7
7	+44.2	+40.8	+ 3.4
8	+26.0		
9	+32.0		
10	+20.1	+ 7.0	+13.1
11	+26.9		
12	+42.7	+35.5	+ 7.2
13	+27.4	- 2.9	+30.3
14	+12.5	- 2.5	+15.0
15	+47.0	+11.1	+35.9
16	+26.8	+19.1	+ 7.7
17	+12.6	- 6.4	+19.0
18	+11.5	+ 8.5	+ 3.0
19	+27.9	+12.9	+15.0
20	+20.2	+22.6	- 2.4

TABLE 3

RESULTS OF PREDICTIONS BY JOKES AND HYPOTHESES

Joke	H1G	H1a	H1b
1	+	+	-
2	+	-	+
3	+		
4	+	+	+
5	+	+	-
6	+	+	+
7	+	+	+
8	+		
9	+		
10	+	+	+
11	+		
12	+	+	+
13	+	-	+
14	+	-	+
15	+	+	+
16	+	+	+
17	+	-	+
18	+	+	+
19	+	+	+
20	+	+	-
Right:Wrong Ratio	20:0	12:4	13:3

TABLE 4
SIGN TEST OF SIGNIFICANCE OF NUMBER OF JOKES
PREDICTED CORRECTLY

Number of Jokes in Predicted Direction	Number of Jokes Predicted Wrongly
H1G	
fo = 20	fo = 0
fe = 10	fe = 10
df = 1; one-tail: $p < .000001$	
H1a	
fo = 12	fo = 4
fe = 8	fe = 8
df = 1; one-tail: $p < .04$	
H1b	
fo = 13	fo = 3
fe = 8	fe = 8
df = 1; one-tail: $p < .02$	

Inspection of Tables 3 or 4 clearly indicates that such was the case; obviously no significant difference exists between 12:4 and 13:3, when $f_e = 8:8$.

Another possible comparison concerned the four Christian-Agnostic jokes in Set 5. While this comparison seemed theoretically important, it was not formally incorporated as an hypothesis. The existence of enough comparisons to make a prediction was only discovered after completion of the experiment. Though Table 3 found all four of these jokes were judged as predicted, $(1/2)^4$ means $p > .05$, an insignificant result. It later occurred to the author, however, that, since the Christians represented a composite of three experimental groups, three predictions could be made for each joke.

Therefore, from Formula H1G for each of the four jokes there occurred three predictions: JW vs. Agnostic; Catholic vs. Agnostic; and Baptist vs. Agnostic. Analysis of these results indicated that all 12 (3×4) predictions were correct--a result significant on the one-tailed sign test at $p < .00025$.

Discussion

The highly significant result of comparisons between agnostics and the three religious groups is theoretically interesting to the social psychologist. It indicates that groups (antagonistic towards each other, as the significant results supporting H1G indicate) who come to identify at a

higher level of abstraction (Christian), against a common antagonist (agnostic), make similar judgments (rather than opposite, as occurred on jokes which caused these groups to be opposed).

What the present experiment seems to indicate rather clearly for the first time, is that humor judgments are a function of reference groups. Further, at a more general level, it supports Hobbes' superiority theory of humor. The humor value which each joke holds for S is directly related to the esteem of his reference group, and is an inverse function of its disparagement. Neutral groups, who are not ego-involved in the joke, tend to assign to it a middle position--finding it neither so funny as does S whose reference group is esteemed by it, nor so ^{UN}funny as does the S whose reference group is disparaged.

The present results suggest that of the three prior experiments which were related to the present one, the two by Wolff et al. and Middleton were on the "right track"--even though their hypotheses, to the effect that humor is enhanced by nonidentification with the butt of the joke, were not fully substantiated. The apparently contradictory hypothesis of Roberts and Johnson, however, to the effect that humor is positively related to identification, was substantiated. Yet the present results suggest that only an inadequate conceptual framework allowed Roberts and Johnson to substantiate their major hypothesis. Influenced as they were by the theories of Mead and Freud, to the effect

that humor is a direct function of "empathic ability," Roberts and Johnson used a personality-trait-type theory which failed to consider the conditions under which the empathic identification occurred. Were their hypothesis adequate, the present experiment should have substantiated neither H1G nor H1b. In fact, H1b results should have been exactly opposite from what they were; the group disparaged in H1b would more frequently, not less frequently, have found these jokes funny than would the neutral, non-ego-involved, groups. After all, the disparaged group obviously held more "empathic ability" towards these particular jokes than did the neutral groups. Results in H1b, then, according to the empathic-ability theory, should have been similar to H1a, as the identical jokes were involved. H1b results, therefore, seem clearly to refute the "empathic" ability theory of Mead and Freud.

Although both the major and the two subhypotheses were substantiated, the fact that the test of the major was strikingly more significant than H1a and H1b, respectively, seems to require some explanation: First, even if every joke had been predicted correctly in H1a or H1b, since each predicts only 16 jokes, whereas H1G predicts 20, the result would have been 1/16 less significant than for H1G--i.e., $(1/2)^{20} = (1/2)^{16} \times (1/2)^4$, and $(1/2)^4 = 1/16$.

Second, H1a and H1b had "bad luck" on "close calls"--losing most of them. (H1G was not "lucky," however; every

one of its 20 predictions was in the right direction by over 11 per cent. See Table 2.)

Third, only the Catholic group failed to displace as predicted. The Catholic and neutral groups were paired against each other in 12 of 32 predictions involving either H1a or H1b. Six of those 12, precisely what would be expected by chance, went counter to prediction. Of those predictions in which one of the three non-Catholic groups was paired against the neutrals, however, only one of 20 proved wrong. (That is, exactly 10 times as many jokes, percentagewise, directly involving the Catholic, as opposed to another, group, failed to follow prediction.)

Each of six times the Catholic group sorted the "wrong" way on H1G the probability was only about 1/20 that the opposing group would also judge "wrongly"; the total probability for these three opposing groups on H1G, roughly, was less than 6/20. However, in all six cases on H1G the opposing group came out strongly enough in the predicted direction to more than offset the wrongly predicted Catholic result. On H1a and H1b, however, there was no opposing group to offset "wrong" Catholic results.

Catholic results were disappointing, not only because this group was the only one which did not judge strongly in line with predictions, but also since pretest findings on Catholics provided no basis for expecting this result. Pre-testing, however, was done in the East and Southwest, while

the experimental testing on Catholics was performed in the Midwest. Eastern Catholics were pretested primarily on Set 1 jokes. Maesen, the pretester, pointed out why he was not surprised that Eastern Catholics judged more as predicted on Set 1 jokes than did Midwestern Catholics: "They (Eastern Catholics) all were much more familiar with the JWs (than Midwestern Catholics were), because that's (East) where they (JWs) have their headquarters, and the local (Catholic) clergy speak out against them (JWs) more often."

Catholic Ss were pretested on Set 3 jokes in the Southwest, with results encouragingly in line with predictions. However, that Southwestern Catholics living in the Southern Baptist Belt might be much more acutely aware of their minority-group status than the more numerous Midwestern Catholics, was not considered. It is not surprising, then, that Midwestern Catholics, probably feeling little animosity towards Baptists, failed to judge Set 3 jokes as predicted.

Experimental results on Midwestern (Northern Indiana) Catholic Ss show that one particular experimental session of Catholic Ss accounted mainly for the negative Catholic results. This experimental session was the largest, consisting of 20, mostly Caucasian, nursing students. Several smaller groups of Negro Catholics, however, displaced very much as predicted, offsetting considerably the results of these nursing students. One reason for the above difference seems to be racial, as expressed by a Negro E, who tested

Negro Ss: "Negro Catholics around here feel a rivalry with Baptists, because most Negroes are Baptists. Negro Catholics also feel a rivalry with JWs, since there is a large Negro JW Kingdom Hall in this city." (For example, at about the same time, a Negro Catholic student of the author's volunteered that she was experiencing marital difficulties over religion, her husband being a member of this same JW Kingdom Hall.)

The agnostics, the group which the author was far more worried about than the Catholics, actually displaced more in line with predictions (only one of eight jokes directly involving agnostics went counter to prediction on H1a or H1b). This finding is especially interesting theoretically in that about 75 per cent of agnostic Ss are Unitarians, and Unitarians pride themselves on, and are often recognized for, intellectualism, liberalism, and free thinking. Catholics, on the other hand, are frequently stereotyped as anti-intellectual, conservative, and dependent upon authority. Yet, Catholic Ss in this experiment showed somewhat more independence in their judgments than did agnostics.

As this reversal of expectation may well be due to sampling peculiarities, it should not prompt quick generalization. However, these results do strongly suggest that if agnostics are nonconformists, they are conforming nonconformists. In other words, the S who deviates from societal convention may be gravitating towards a smaller unconventional

circle in which he may be more of a conformist than is the conventional person within the larger circle of society.

--Indeed, judgmental variability within extremist groups may be no more, and sometimes less, than within more moderate, conventional groups. (Muzafer Sherif has often made this observation in his lectures.)

As indicated earlier, all 12 predictions on Set 5 jokes in HIG (for each of the three Christian groups, as opposed to the agnostic) came out as predicted. The Catholics, judging least in line with the hypotheses of the three Christian groups, might be expected to differ from the agnostics on Set 5 jokes by the smallest amount. However, the JWs, not the Catholics, had been expected to show the weakest trend on Set 5 jokes, and did. But why?

JWs, while considering themselves Christian, think the so-called Christian churches practice "Christendom" (a derogatory term for JWs), rather than "Christianity." Under these circumstances, JWs might hold towards the word Christian an attitude ambivalent--favoring the term only when it is applied to themselves. If, however, the JWs feel ambivalence towards the word Christian, their judgments should not support HIG on Set 5 jokes as strongly as those of other Christian groups.

An occasional S presented results which good reason seemed to suggest were invalid. The scientific method appears to demand, however, that no S be eliminated by ad hoc

reasoning, no matter how convincing. Only Ss who violate criteria for selecting Ss (as determined before the experiment) can, and in fact must, be eliminated.

Two Ss in particular, who sorted to the detriment of the hypotheses, and whose results, though apparently invalid, must be counted, include the following: A21, who appeared to "see through" the purpose of the experiment. This agnostic asked E, after taking the experiment, if its "purpose was to see if we would laugh at ourselves."

Another sorting, probably invalid, was that of JW7, who placed all 20 jokes in the Indifferent category. Maesen, one of the JW Es, was not surprised, and was especially interested in this result, because he had known that this particular S was a "frustrated would-be overseer" of JWs. In other words, JW7 is suspected of having made no attempt to discriminate between the jokes--such results as his seem to be symptoms of frustration or hostility.

One "agnostic" S was eliminated for failing to check herself as an agnostic on the Information Sheet. Yet, this same S, in the presence of an agnostic E, and several other Unitarians, had called herself an agnostic--revealing, thereby, a discrepancy between her public and private behavior. The history of the Unitarian Church shows cause for this discrepancy.

Several years ago, an agnostic E pointed out, the name Christian Register was changed to Unitarian Register,

"because members didn't like the word Christian." (Yet, some "old-fashioned" Unitarian Churches still appear to favor the word; thus, no attempt was made to gain agnostic Ss in those particular places. Furthermore, as this agnostic E added, older members of the more agnostic Unitarian Churches are still inclined to consider themselves Christians.) It is among the "younger set" in the more liberal Unitarian Churches that agnosticism is fashionable. In short, this S, apparently influenced by the "group pressure" and public scrutiny of a younger set of Unitarian agnostics, did refer to herself as an agnostic. However, under anonymous experimental conditions, when she may have felt alone with her conscience, she could be what she privately was--and could not bring herself to check agnostic.

Suggestions for Future Research

Now that the present hypotheses have been substantiated, it is possible to reason backwards--using them as tentative assumptions upon which to guide future research. One way in which these results can steer future research relates to differences between twin jokes. (Twin jokes refer to those two jokes which occur within each permutation within each set. Such twins are, for purposes of this experiment, identical.)

Had the hypotheses not been verified, twin-joke comparisons would not reveal as much--differences between twins per se do not tell which twin is the "abnormal" one. Yet,

substantiated hypotheses can be used to serve the function of "background"; if one twin acts in line with the hypotheses, and the other does not, the one which does not may be considered abnormal. What is called for, then, is the detection of essential ways in which the abnormal twin differs from the normal. (Theoretically-rooted suggestions for future research, based on differences between twins, will be presented later.)

One suggestion would be an experiment to determine whether reference or membership groups are more important influences on humor judgments. Ss for whom the group in question is a membership-nonreference group could be compared with Ss for whom the group is a reference-nonmembership group. For instance, college students who have not yet been accepted into a fraternity, but identify with it, would represent reference-nonmembership-group Ss. Fraternity members, who remain so only for ulterior motives, would be considered membership-nonreference-group Ss. Based on theory used here, when an enemy group disparages the reference-nonmembership group, Ss should sort the jokes as the disparaged group did in the present experiment; when Ss' reference-nonmembership group is esteemed, they should sort like the esteemed group. Membership-nonreference-group Ss, however, should sort jokes not greatly different from the way the neutral groups in the present experiment did.

The fraternity illustration suggests that the theory employed in the present experiment, if useful, ought also to

apply in dimensions other than religious. For instance, loyal Democrats ought to find jokes which disparage Republicans and esteem Democrats funnier than those jokes which do the reverse; and ego-involved Republicans ought to do the opposite. Political parties, races, Greek-Independent-student rivalry, ethnic groups--these are a few other dimensions in which the current hypotheses could be tested.

Besides changing the dimensions, future research might also vary the identification levels. Modern man seems to identify at abstraction levels both higher and lower than those used presently. For instance, the dimension used here is religion, and three of the experimental groups were called upon, by the nature of the stimulus material (jokes), to identify at two different levels of abstraction: the "group" level (i.e., Catholic or Baptist of JW), and the "supergroup" level (i.e., Christian). However, S might be presented with a joke which causes him to identify at a level still higher than Christian. Thus, a joke involving dialogue between a "religious" person and an "atheist" might cause Christians, Hindus, and Moslems to identify against the common atheistic enemy. The religious level of abstraction is a higher one than the Christian level, since the former includes non-Christian groups as well. Likewise, the Christian level is higher than the Catholic, etc. These different levels within a particular dimension might be viewed as concentric circles--the bigger the circle, the higher the level of

abstraction.

Within the "Group" circle could be placed smaller circles representing intragroup or interindividual status striving. For instance, a joke disparaging a Catholic Priest might more often be funny to Catholic non-Priests, if told by a Catholic non-Priest, than the same joke would be to members of another religious group.

The typical socialized adult in modern society seems to carry around in the "back of his mind" numerous dimensions and levels of identification. Yet, these levels and dimensions do not operate simultaneously and all at once to govern his judgments. If they did, life for him would be but chaos and confusion. Rather, which dimension and which level "pops into mind" at a given moment seem to be the weighty factor in judgmental determination, and to act as a function of the nature of the stimulus condition (joke). (Such theory is outlined as above, because its implications for humor research seem too numerous to enable specific experiments to be appropriately elaborated here.) Also, concentric-circle-levels-of-identification theory is supported by results on Set 5 jokes, which, as mentioned earlier, actually demonstrated that three Christian groups, who displaced differently on sets suggesting intra-Christian rivalry, actually sorted similarly when the other speaker became an outsider (agnostic).

Students of humor, puzzled by the apparent ability of some people to laugh at themselves, commonly suggest that

perhaps two types of humor--"ethnocentric," and "philosophical"--exist. Ethnocentric humor is viewed as the common, vulgar type, and seems to fit well with Hobbes' superiority theory of humor (the general notion from which the present hypotheses are derived). Philosophical humor, however, is looked upon as what is "really" meant by a sense of humor, and only comes to a select few of sufficient age and wisdom.

But is such a dichotomy (suggesting, as it seems to, that two different sets of psychological principles are required to explain humor) necessary? Or does it but reflect some rather loose thinking by these theorists? Which question receives a "yes" answer has broad implications for future research. But how could Hobbes' superiority theory account for the ability, often reported, of some people to laugh "philosophically" at themselves?

First, if no humor is involved, Hobbes' theory has no explaining to do. This study opened by questioning the assumption that laughter and humor are one-to-one related. Perhaps a person laughs at his own expense, not because he is humored, but since he does not want others to think him a "poor sport." Humor research is recommended, therefore, concerning the social conditions under which humor and laughter are, and are not, correlated. But how could the superiority theory explain genuine humor at one's own expense?

Second, perhaps a person is humored at "his own expense" only because the joke is not costing his ego

anything. At least several ways seem to exist in which such detachment can occur: (A) His membership group, as suggested earlier, may not be his reference group. In such an instance, holding no ego-involvement in the membership group, S may feel no loss at its disparagement. In extreme instances, S may actually despise the group, but not feel that he can withdraw from membership safely. In this case, and based upon superiority theory, S might actually find more humor in seeing "his" group made the butt of a joke than would a neutral outsider enjoy seeing S's group disparaged. (B) S may hold conflicting reference groups. In this event, if one reference group attacks another, he might find the joke funny--perhaps identifying more strongly with the attacking group than the disparaged, so that a lesser ego loss is surmounted by a greater gain. (C) The level of identification which certain nouns in the joke cause to gain conscious awareness may be different from what the humor theorist suspects. For example, as discussed earlier, perhaps a Catholic S may find a joke funny in which a Catholic Priest is the butt, not because S judged a Catholic (part of his own ego) to be attacked--but rather since S, a non-Priest, conceived that only Catholic Priests were attacked. (Suggestions for future research involving (A) and (C) have already been discussed.) (D) S may be humored at the expense of what he conceives to be, not his present self, but his former self--as when a man is humored by a faux pas he committed as

a child.

The above discussion indicates testable reasons as to why Ss who "laugh at themselves" may appear to, but not really, contradict Hobbes' superiority theory of humor. However, as the present experiment indicates, if the group disparaged is S's reference group, rather than simply his membership-nonreference group, and if the victor belongs to another, preferably enemy, group, then S is unlikely to find the joke funny. Rather, support of H1b shows that disparaged groups were especially likely to find these jokes unfunny. Of course, if the joke is told by an ingroup member, and if S feels no ego-involvement with the group anyway, then he might be as likely as most non-ego-involved outsiders to find the joke funny. (But such were not the conditions of the present experiment.)

At least one other way (E) appears possible for S to be humored at "his own expense" without violating the superiority theory--that is, when the joke itself seems so incongruous, ridiculous, or absurd that S cannot take it seriously (cannot be threatened by it). Sherif and Hovland (1953), and La Fave and Sherif (1959) have demonstrated that attitudinal statements close to S's own position tend to be assimilated, i.e., displaced towards his position, while statements further away tend to be contrasted (displaced away from his own position). Perhaps statements still further removed from S's own position will seem "so far out in left field"

as to appear humorous.

For instance, an "insult" may seem so unjust that it is judged as humorous, rather than insulting--a kind of humor sometimes referred to as irony. This type of joke (which seems more directly related to an incongruity, rather than superiority, theory of humor) may be judged equally funny by the group disparaged and the group esteemed--since neither feels a real ego loss or gain from it, the joke being too ridiculous to be taken seriously.

But how can such jokes be identified? The present experiment suggests a method: analyze the "abnormal" member of twin jokes. One might suspect that especially funny jokes (those with high humor medians) would show a lack of variability between groups; such jokes, ridiculous or incongruous enough that every group finds them funny, would fail to support the hypotheses. (However, no jokes in the present experiment failed to come out as predicted on H1G.) Another type of joke which might not be displaced differently by antagonistic groups is one in which the group is disparaged or esteemed for an "unrealistic" reason, i.e., a reason which seems to have no connection with the norms of that group. (Lack of "realism" in some jokes may help explain why some "abnormal" twins are found on H1a or H1b.)

An experiment could be set up similar to the present one, except that the twins would not remain twins; one former twin would disparage a given group for no "good" reason, while the other would belittle a major value of that group.

CHAPTER IV

SUMMARY AND CONCLUSIONS

A "superiority theory" of humor was tested by relating humor judgments to reference groups. Hypothesis H1G states that jokes tend to be judged as funny by Ss whose reference (identification) group is esteemed, and whose outgroup is disparaged, and to be judged unfunny by Ss whose reference group is disparaged and whose outgroup is esteemed. Also expected to support the major hypothesis are jokes which cause Ss to identify at the "supergroup" level. Hypothesis H1a: jokes which esteem Ss' reference group tend to be judged as funny; H1b: jokes which disparage Ss' reference group tend to be judged as unfunny.

Four experimental groups (Catholics, Jehovah's Witnesses, Southern Baptists, and agnostics), each containing over 50 Ss, were asked to sort 20 jokes on religion into five degrees of funniness. Each S was chosen only if both a participant observer and S himself indicated the appropriate one of these four groups as S's reference group. (These four particular groups were selected because each has traditionally felt unfriendly towards the other three.) A built-in

experimental control means that each joke which esteems one group and disparages another is counterbalanced by a joke which does the opposite.

A sign test substantiated all four predictions; and, on H1G all 20 jokes were judged as predicted ($p < .000001$).

Results show that conflicting findings among previous studies regarding the relationship between humor and identification are resolved through a more rigorous experimental design--and that the concept of reference group, applied within the context of a "superiority theory" of humor, leads to such success in predicting humor judgments as earlier studies were unable to demonstrate.

REFERENCES

- Bergson, H. Laughter: An essay on the meaning of the comic. New York: Macmillan, 1911.
- Cattell, R. B. and Luborsky, L. B. Personality factors in response to humor. J. abnorm. soc. Psychol., 1947, 42, 402-421.
- Doris, J. and Fierman, Ella. Humor and anxiety. J. abnorm. soc. Psychol., 1956, 53, 59-62.
- Drever, J. A dictionary of psychology. Baltimore: Penguin, 1952.
- Eysenck, H. J. An experimental analysis of five tests of "appreciation of humor." Educ. Psychol. Measmt., 1943, 3, 191-214.
- Eysenck, H. J. National differences in sense of humor. Charact. Pers., 1944, 13, 37-54.
- Eysenck, H. J. Dimensions of personality. London: Kegan Paul, 1947.
- Flugel, J. C. Humor and laughter. In G. Lindzey (Ed.), Handbook of social psychology v. 2. Cambridge: Addison-Wesley: 1954. Pp. 709-732.
- Freud, S. Wit and its relation to the unconscious. New York: Moffat, Yard, 1917.
- Freud, S. On humor. Int. J. Psychoanal., 1928, 9, 1-6.
- Ghosh, R. An experimental study of humor. Brit. J. educ. Psychol., 1939, 9, 98. (Abstract)
- Heim, A. An experiment in humor. Brit. J. Psychol., 1936, 27, 148-161.
- Hobbes, T. Collected works. In W. Molesworth (Ed.) v. 4. London: Bohn, 1940.

- Hollingworth, H. L. Experimental studies of judgment of the comic. Psychol. Rev., 1911, 18, 132-156.
- Hyman, H. The psychology of status. Arch. Psychol., 1942, No. 269.
- La Fave, L. and Sherif, M. Placement of items on a controversial social issue. Pre-publication report, Institute of Group Relations, University of Oklahoma, 1959, 1-16.
- Lee, Joan C. and Griffith, R. M. Forgetting of humor: Repression? Amer. Psychologist, 1960, 15, 436. (Abstract)
- Martin, Lillian J. Psychology of aesthetics: Experimental prospecting in the field of the comic. Amer. J. Psychol., 1905, 16, 35-116.
- Mead, G. H. Mind, self and society. Chicago: University of Chicago, 1934.
- Middleton, R. Negro and white reactions to racial humor. Sociometry, 1959, 22, 175-183.
- Middleton, R. and Moland, J. Humor in Negro and white subcultures: A study of jokes among university students. Amer. Soc. Rev., 1959, 24, 61-69.
- More, D. M. and Roberts, A. F. Societal variations in humor responses to cartoons. J. soc. Psychol., 1957, 45, 233-243.
- Omwake, Louise. Factors influencing the sense of humor. J. soc. Psychol., 1939, 10, 95-104.
- Perl, Ruth E. A review of experiments on humor. Psychol. Bull., 1933, 30, 752-763. (a)
- Perl, Ruth E. The influence of a social factor upon the appreciation of humor. Am. J. Psychol., 1933, 45, 308-312. (b)
- Piddington, R. The psychology of laughter: A study in social adaptation. London: Figurehead, 1933.
- Raley, Agnes Lucile and Ballmann, Christine. Theoretical implications for a psychology of the ludicrous. J. soc. Psychol., 1957, 45, 19-23.
- Roberts, A. F. and Johnson, D. M. Some factors related to the perception of funniness and humor stimuli.

- J. soc. Psychol., 1957, 46, 57-63.
- Rosenberg, G. and Shapiro, G. Marginality and Jewish humor. Midstream, 1958, 4, 70-80.
- Sherif, M. A study of some social factors in perception. Arch. Psychol., 1935, No. 187.
- Sherif, M. and Cantril, H. The psychology of ego-involvements. New York: Wiley, 1947.
- Sherif, M. and Hovland, C. I. Judgmental phenomenon and scales of attitude measurement: Placement of items with individual choice of number of categories. J. abnorm. soc. Psychol., 1953, 48, 135-141.
- Sherif, M. and Sherif, Carolyn W. An outline of social psychology. (Rev. ed.) New York: Harper & Bros., 1956.
- Siegel, S. Nonparametric statistics for the behavioral sciences. New York: McGraw-Hill, 1956.
- Stump, N. F. Sense of humor and its relationship to personality, scholastic aptitude, emotional maturity, height and weight. J. gen. Psychol., 1939, 20, 25-32.
- Valentine, C. W. The psychology of early childhood. London: Methuen, 1942.
- Washburn, Ruth W. A study of the smiling and laughing of infants. Genet. Psychol. Monogr., 1929, 6, 405-537.
- Wolff, H. A., Smith, C. E. and Murray, H. A. The psychology of humor. J. abnorm. soc. Psychol., 1934, 38, 345-365.

APPENDICES

APPENDIX A

The 20 Jokes¹

Joke 4. Catholic: Is it true that in your religion every man is a minister?

Jehovah's Witness: Certainly! What is so funny about that?

Catholic: I was just thinking of what a large congregation each of you must have.

Joke 17. Catholic: Why are you tied to this telephone pole?

Jehovah's Witness: Some devils did it just because I refused to salute the American flag. Untie me Brother!

Catholic: Do you not believe that refusal to salute the flag is the only true way?

Jehovah's Witness: Yes, and the Truth shall make you free.

Catholic: Then you need no help from me. Let the Truth, which you have followed, set you free.

Joke 1. Catholic: I thank God that I don't question anything. I accept things on faith.

Jehovah's Witness: What? You mean you are thankful for your ignorance?

Catholic: I am!

Jehovah's Witness: Well then you have a lot to be thankful for.

¹The 20 jokes are presented in the order shown in Table 1.

Joke 18. A Jehovah's Witness and a Catholic met on a narrow walk:

Catholic: I never make way for a fool.

Jehovah's Witness (letting Catholic pass): Don't you? I always do.

Joke 14. Agnostic: I have noticed that Catholics are quite stupid when it comes to matters of religion. For example, can you tell me the difference between an agnostic and a Buddhist?

Catholic: I think so. A Buddhist is going into Purgatory whereas an agnostic is going straight to Hell.

Joke 19. Agnostic: Have you heard the story about the argument between the Catholic and the jack ass?

Catholic: No, but I'll bet in your version of it the agnostic won.

Joke 10. Agnostic: What is an archbishop, a successor to the peaceable apostles, doing marching at the head of a train of soldiers?

Catholic Archbishop: Aye, but I do not head these soldiers as an archbishop, but as a prince of Strasburgh.

Agnostic: Then if the prince of Strasburgh should go to the devil, what will become of the archbishop?

Joke 20. Catholic Priest: You are an agnostic aren't you. You don't know whether or not there is a God.

Agnostic: That is true.

Catholic Priest: Well, why don't you become a martyr so you can go to Heaven and find out?

Agnostic: You are a priest aren't you. You want to save souls from Hell.

Catholic Priest: Of course.

Agnostic: Well, why don't you go there?

Joke 6. Baptist: It looks to us Baptists like you Catholics make too much fuss over the Virgin Mary.

Catholic: It is not surprising that natural things should look strange to you Baptists. After all, who can see clearly when he spends so much time with his head under water.

Joke 13. Baptist: Why do you call that fish you just caught a Baptist fish?

Catholic: Because they spoil so fast after you take them out of water.

Joke 5. Catholic: Where was your religion before Luther?

Baptist: Did you wash your face this morning?

Catholic: I did sir.

Baptist: Then where was your face before it was washed?

Joke 16. Catholic: We have a very good reason for being opposed to birth control, a fact which you don't seem to realize.

Baptist: I know why. You are opposed to birth control because the only converts you can win are your own children.

Joke 7. Agnostic: Do you really believe that silly story about Jonah being swallowed by a whale?

Baptist Preacher: When I get to Heaven, I will ask Jonah.

Agnostic: But supposing he is not there?

Baptist Preacher: Then you will have to ask him.

Joke 12. Agnostic: Why do you allow me to see you while you deny yourself to all your friends?

Ill Baptist Minister: I feel confident of seeing my friends in the next world, but this may be my last chance of seeing you.

Joke 2. Baptist Minister: Hell contains nothing but roulette wheels, whiskey and chorus girls.

Agnostic: Oh death, where is thy sting!

Joke 15. Baptist Minister: Why don't you ever come to hear me preach?

Agnostic: I have conscientious scruples against going to places of amusement on Sunday.

Joke 8. Agnostic: It may interest you to know that I am not sure that Heaven exists.

Christian: (Nods head, goes on reading.)

Agnostic: You don't realize what I'm trying to make clear.

I want you to understand that I don't for a single, solitary moment believe that such a place as Heaven exists.

Christian: All right, all right, go to Hell! Only don't make so much fuss about it.

Joke 9. Agnostic: If Jesus Christ returned, his only accomplishment would be to get crucified again.

Christian: Perhaps. But the accomplishment would not be his alone. It would be with the help of skeptics such as you.

Joke 3. Christian: Why did so many Christians come over to America?

Agnostic: To worship in their own way, and make others do the same.

Joke 11. Christian: An agnostic is a blind man groping in a dark cellar for a black cat that isn't there.

Agnostic: Yes, and a Christian would find it.

APPENDIX B

Information Sheet

Age _____ Sex: Male _____ Female _____

Nation most of your life was spent in: United States _____

Other _____

If answer to the above is the "United States," mention
states you have lived in, and number of years lived
in each one: _____

Education: (Check level of highest education)

Less than 12 years _____

High School Graduate _____

Years of College = 1 _____ = 2 _____ = 3 _____ = 4 _____

College Graduate _____

Some Graduate Work _____

Religious Preference: Check the religion which you prefer
above all others:

Presbyterian _____

Atheist _____

Catholic _____

Humanist _____

Episcopalian _____

Other _____

Baptist _____

Agnostic _____

Church of Christ _____

Jehovah's Witness _____

Methodist _____

Do you like to consider yourself a Christian? Yes _____

No _____

APPENDIX C

TABLE 5

HUMOR JUDGMENTS BY CATHOLIC GROUP

Joke	VF	F	I	UF	VUF	Totals
1	10	19	10	11	2	52
2	27	20	5	0	0	52
3	0	12	16	9	15	52
4	6	19	12	12	3	52
5	6	17	6	14	9	52
6	7	20	12	9	4	52
7	15	26	7	2	2	52
8	21	22	6	2	1	52
9	1	7	20	11	13	52
10	0	11	12	15	14	52
11	6	11	19	11	5	52
12	11	22	11	5	3	52
13	6	20	10	8	8	52
14	6	17	8	15	6	52
15	12	21	6	10	3	52
16	5	14	7	8	18	52
17	2	15	15	11	9	52
18	13	21	6	7	5	52
19	12	20	11	7	2	52
20	<u>16</u>	<u>12</u>	<u>5</u>	<u>15</u>	<u>4</u>	<u>52</u>
Totals	182	346	204	182	126	1040

TABLE 6
HUMOR JUDGMENTS BY JEHOVAH'S WITNESSES GROUP

Joke	VF	F	I	UF	VUF	Total
1	13	21	14	1	2	51
2	9	21	8	11	2	51
3	0	16	19	11	5	51
4	1	4	18	22	6	51
5	4	8	19	11	9	51
6	3	14	20	9	5	51
7	6	11	13	17	4	51
8	3	20	12	13	3	51
9	0	8	19	17	7	51
10	1	3	24	17	6	51
11	2	11	19	14	5	51
12	6	13	18	11	3	51
13	5	22	12	10	2	51
14	2	9	23	13	4	51
15	10	14	15	9	3	51
16	5	17	13	8	8	51
17	0	3	15	22	11	51
18	15	17	14	5	0	51
19	4	7	22	13	5	51
20	<u>3</u>	<u>19</u>	<u>12</u>	<u>15</u>	<u>2</u>	<u>51</u>
Totals	92	258	329	249	92	1020

TABLE 7
HUMOR JUDGMENTS BY AGNOSTIC GROUP

Joke	VF	F	I	UF	VUF	Total
1	8	16	17	6	5	52
2	22	21	3	4	2	52
3	10	23	7	6	6	52
4	6	8	22	8	8	52
5	2	13	19	11	7	52
6	6	13	16	8	9	52
7	4	28	6	9	5	52
8	6	21	12	9	4	52
9	1	3	22	15	11	52
10	5	13	17	10	7	52
11	13	19	11	6	3	52
12	4	23	15	8	2	52
13	7	11	14	12	8	52
14	3	10	17	12	10	52
15	15	24	9	2	2	52
16	7	7	20	10	8	52
17	2	12	14	14	10	52
18	4	28	10	5	5	52
19	3	11	15	15	8	52
20	<u>18</u>	<u>17</u>	<u>8</u>	<u>7</u>	<u>2</u>	<u>52</u>
Totals	146	321	274	177	122	1040

TABLE 8
HUMOR JUDGMENTS BY BAPTIST GROUP

Joke	VF	F	I	UF	VUF	Total
1	8	18	21	5	2	54
2	3	14	13	17	7	54
3	2	7	16	15	14	54
4	4	15	24	7	4	54
5	13	28	10	2	1	54
6	0	1	5	21	27	54
7	26	25	2	0	1	54
8	25	18	6	1	4	54
9	16	17	15	4	2	54
10	6	9	22	10	7	54
11	2	9	19	16	8	54
12	22	23	6	2	1	54
13	0	1	6	25	22	54
14	7	14	19	10	4	54
15	1	8	19	15	11	54
16	14	15	16	5	4	54
17	4	15	19	13	3	54
18	16	16	12	7	3	54
19	7	16	21	6	4	54
20	<u>4</u>	<u>11</u>	<u>13</u>	<u>15</u>	<u>11</u>	<u>54</u>
Totals	180	280	284	196	140	1080