

## Predictability and the appreciation of comedy\*

HOWARD R. POLLIO and RODNEY W. MERS  
University of Tennessee, Knoxville, Tennessee 37916

In order to investigate the role of joke predictability in humor appreciation, Ss were asked to provide punch lines for in-concert routines performed by Bill Cosby and Phyllis Diller. Predictability scores derived from each of these routines were then related to a number of measures of audience laughter—frequency, duration, latency, etc.—produced in response to these routines. In general, results showed no difference in predictability of punch lines and nonpunch lines and strong positive correlations between predictability and the various indices of audience responsivity. These results were discussed in terms of a predictable unpredictability theory of humor.

In one way or another, almost all theories of humor assume that a basic element is unpredictability. In an exhaustive survey of the theoretical literature, Keith-Spiegel (1972) listed no less than 35 different theories which involve principles such as incongruity of ideas, surprise, suddenness, or unpredictability as critical in bringing about a comic or humorous reaction. With so great a theoretical emphasis on predictability, it seems surprising that only a small number of studies have been concerned, even peripherally, with understanding the role of joke expectedness on joke funniness. At the level of data or, more correctly, at the level of operations, what is required are independent measures of joke predictability and joke funniness and, while this seems a straightforward statement of what needs to be done, the actual procedures have yet to be applied or even worked out.

In addition, there is also the problem of context in obtaining such measures; humor is a fragile phenomenon and asking people for judgments of funniness and/or predictability in the context of single jokes (Harrower, 1932; Kenny, 1955) or in odd laboratory situations may not represent a completely adequate or appropriate procedure. It was for this latter reason that Pollio, Mers, and Lucchesi (1972) suggested that a relatively unstructured audience setting would probably represent a more appropriate situation. Prior work had shown that such a setting does provide a good locale for recording and analyzing the responses of individual audience members (as well as of the total audience itself) and further suggested such a situation as a good one in which to study the relationship of joke predictability to joke funniness. Within this context, joke funniness can be defined behaviorally in terms of the duration, frequency, and latency of group laughter, as well as in terms of rated funniness, while predictability can be defined in terms of the success different groups of Ss have writing punch lines for each of the jokes heard. With these

considerations in mind, the purpose of the present study was to determine (1) if Ss in a relatively unstructured audience context could successfully predict punch lines, (2) if this ability to predict punch lines varied across comedians, (3) if joke predictability relates to measures of audience laughter such as latency, duration, frequency, etc., and (4) if all measures of joke funniness correlate with one another as well as with predictability.

### METHOD

In order to analyze joke completions, it is first necessary to define punch lines as unambiguously as possible. For the present study, punch lines were defined with reference to group laughter: Wherever Ss in a prior study (Pollio et al, 1972) laughed, such laughter was noted and served to define the location of a punch line for the present study. In dealing with jokes of the one-liner variety, specifying location is an easy task. In the type of humor performed by Bill Cosby, and to a lesser extent by Phyllis Diller (the comedians of interest for the present study), it is sometimes difficult to decide exactly where a stem ends and a punch line begins, and for this reason the syntactic unit immediately preceding group laughter was used to define the location of punch lines.

#### Subjects

All 44 Ss employed in the predictability part of this study were University of Tennessee undergraduate students, as were all 16 Ss used in the rating part.

#### Stimulus Materials

Stimulus materials consisted of four professionally recorded comedy routines for each of two comedians, Bill Cosby and Phyllis Diller. Three of the Cosby routines were taken from the album "Revenge" (Warner Brothers, No. WS-1691), with the remaining routine coming from the album "Wonderfulness" (Warner Brothers, No. WS-1634). All Diller routines were taken from her album, "The Best of Phyllis Diller" (Verve, No. V-15053). In choosing specific routines, an attempt was made to match them with regard to playing time and number of within-routine punch lines. Two Cosby routines contained five punch lines each and had the following playing times: *The Playground*, 3:21; *Cool Covers*, 4:25. The remaining two routines contained three punch lines each and had playing times of 2:00 for *Planes* and 1:35 for *The Tank*. Two Diller routines also contained five laughs each, with playing times as follow: *The Management*, 4:00; *Cheese and Turkey*, 5:38. The two remaining routines had three punch lines each and playing times of 2:20 for *Lipstick* and 2:30 for *Plastic Surgery*.

\*This study was supported in part by Grant MH-20442 from the U.S. Public Health Service. Some of these data were reported previously in a Master's thesis (Mers, 1973), done under the direction of H. R. Pollio.

**Table 1**  
**Response Measures for All Comedy Routines**

Comedy Situation		Percent Agreement		Group Measures			
Comedian	Routine	Nonjokes	Jokes	Funniness	Duration	L/S	Latency
Cosby	Playground	41	24	3.81	3.32	.24	.62
	Planes	18	12	3.50	2.40	.14	.90
	Cool Covers	34	62	3.50	3.34	.27	.58
	Tank	14	21	3.50	3.07	.13	.80
Mean		27	30	3.58	3.03	.20	.73
Diller	Management	0	15	1.75	2.62	.15	1.78
	Lipstick	32	30	2.56	2.27	.16	1.83
	Cheese and Turkey	30	14	2.19	1.92	.09	.94
	Plastic Surgery	9	6	2.19	1.63	.12	1.60
Mean		18	16	2.17	2.11	.13	1.54

In addition to punch lines, other lines which produced no laughter were also specified. These lines were located in routines such that the stem or preceding material was of approximately the same length as joke stems. A further constraint on these nonlaughter lines involved the location of the actual punch lines themselves. The routines which contained five jokes had two nonlaughter lines omitted, while routines with three jokes had one nonlaughter line each.

#### Procedure

**Predictability Judgments.** Stimulus materials were presented both in written form and on audio tape. A typed script of each routine was presented so that each stem (joke and nonjoke line) was situated on a separate sheet of paper stapled together to form a booklet for each routine. In addition, a separate response sheet was provided on which Ss wrote their completions. To the right of each space were the words "yes" and "no", the purpose of which will be explained shortly.

The initial intention was to run two groups, using all eight routines, with Group I listening to the four Diller routines first and Group II the four Cosby routines first. After a few pilot Ss, however, it became clear that four routines for each comedian took too much time. In order to decrease each S's time in the experiment, four experimental conditions were employed, with the actual playing time for each comedian cut by using only two routines for each. To evaluate possible order effects, counterbalancing was used, thereby yielding four groups delineated by routines and order of presentations.

The experiment was conducted in a classroom with Ss seated at individual tables. A tape recorder was placed at the front of the room with an external speaker facing Ss. Instructions were read aloud at the beginning of the session in which Ss were told what the experiment was about and that their task was to write punch lines for joke stems presented. Taped routines were played through the speaker and stopped immediately prior to each punch line. All Ss were given time to write out completions. When finished, the actual punch line was played and the tape stopped after that point. The Ss were then asked to make a judgment as to whether each punch line "meant the same thing" as the comedian's punch line, and instructions required a yes or no response to this question in the appropriate place on their answer sheets.

**Funniness Ratings.** In addition to these four groups, two other groups listened to the same routines and were asked to rate them on a scale from 1 (not funny) to 5 (very funny). These groups were run with order of comedians reversed for each of the groups.

#### Group Laughter Measures

Group measures, such as duration, latency and frequency of laughter, were taken from earlier results reported by Pollio, Mers, and Lucchesi (1972). These values were obtained from

four different groups—two for Cosby and two for Diller—containing between 13 and 15 Ss each. All values were obtained from oscillographic recordings of group laughter, with the exact procedures specified in the chapter by Pollio et al (1972).

## RESULTS AND DISCUSSION

### Predictability of Joke Endings

In order to evaluate joke predictability, two independent raters first categorized all completions produced. In doing this they were asked to use the same criterion as did Ss in the initial groups, i.e., did the punch line mean the same thing as the comedian's punch line? Inter-rater agreement scores for the three raters (the original S and the two new raters) were computed and results indicated that independent raters agreed with each other 87% of the time and with each original S's judgment 86% and 88% of the time. For this reason, yes or no values were defined as agreement between any of two of the three raters. Since counterbalancing procedures were used, an examination of order effects was undertaken and indicated little or no difference due to order. All judgments, therefore, were pooled, producing 22 Ss in both Cosby and Diller groups.

Following this, a mean percentage of yes responses for all punch lines in a given routine was obtained and used as the measure of predictability for that routine. These values are presented separately for jokes and nonjokes in the Percent Agree column of Table 1, where it can be seen that jokes and nonjokes were equally predictable for both comedians, although some differences did appear between comedians. For Cosby, jokes are more predictable than nonjokes in two routines and less predictable in two. The Diller material presents a similar pattern: Two routines produced a pronounced difference between predictability scores for jokes and nonjokes, and, as with Cosby, these differences were in opposite directions for each routine in question.

Looking at overall mean predictability scores for Cosby and Diller (jokes and nonjokes), it appears that Cosby is the more predictable of the two, although an examination of individual values shows Cosby's overall

score is inflated due to the single routine *Cool Covers*. With nonjokes there again appears to be a difference between the two comedians, although variability among routines makes it difficult to make any clear-cut between-comedian statement.

**Group Measures**

The next three columns of Table 1 present group measures evoked in response to Cosby and Diller, and indicate that on all measures Cosby's material was funnier than Diller's. This relationship is particularly striking considering the total distribution of values: In all cases mean ratings produced by Cosby were higher than any produced by Diller. With this particular group of Ss and routines, Cosby is the funnier comedian by a reasonably large margin.

Results presented in Table 1 show that the latencies (elapsed time from termination of punch line to onset of group laugh) are consistently longer for Diller than for Cosby. The values for duration present a similar pattern but in an opposite direction, and only in one case, *The Management*, does a Diller routine produce a mean duration slightly longer than that produced by any Cosby routine. Durations are significantly longer for Cosby than for Diller groups.

The data presented in Table 1 also contain a column labeled L/S (laughs per sentence) which was computed by dividing the total number of laughs in a given routine by the total number of sentences in that routine. For this reason L/S values can be construed as a measure of how frequently each routine induced laughter with routine length taken into account. As with funniness ratings and duration, Cosby's laughs per sentence (L/S) values were consistently higher than those produced by Diller.

**Intercorrelation Among Response Measures**

Table 2 presents the pattern of product-moment correlations for all measures except the percent agreement of nonjokes. These correlations were computed over the eight pairs of values provided by each of the four routines for each comedian. In terms of theoretical concerns, the measures L/S, duration (D), and funniness (F) all represent measures that should correlate with each other and should show some consistent relationship to predictability (P). We should also expect a consistent relationship between each of these measures and latency (L). The results presented in Table 2 do reveal a rather consistent pattern: L/S, D, and F all correlate positively with each other and with P, and all four measures do correlate negatively with L.

An examination of the appropriate scatter plots for each of these correlations indicated that caution is necessary in interpreting some of the values. Of the 10 correlations, only those involving L/S and P, D and P, and D and L/S provide appropriate numerical estimates of the relationship between the two variables in

**Table 2**  
**Intercorrelations Among All Response Measures**

	Predict-ability	L/S	D	F	Latency
Predictability	—	.81**	.66*	.43	-.40
L/S		—	.79*	.59	-.44
Duration			—	.72*	-.62
Funniness				—	-.81**
Latency					—

\*p < .05      \*\*p < .01

question. Of the remaining correlations, all four involving funniness (except for that between F and P) result from overall differences in the F ratings produced by Cosby and Diller groups. An examination of individual plots indicated no clear-cut relationship within the four pairs of values for either comedian, although there was an obvious and strong relationship due to differences between comedians. Even the -.81 correlation between L and F does not present a true picture of within-comedian patterns, depending instead upon cross-comedian differences.

For this reason, a careful examination was made of the remaining correlations, all of which involve L. An examination of the appropriate scatter plots did indicate a clear and consistent pattern: All correlations involving Cosby's routines produced negative values, while all correlations involving Diller's routines produced positive values. More specifically, the correlation coefficients for Cosby were: L and P, -.78; L and D, -.92; and L and L/S, -.94 (all ps < .05). Comparable values for Diller were +.36, +.48, and +.95, respectively, with only the last value significant at p < .05.

Correlational results thus indicate that predictability correlates positively with all measures of a group's responsiveness to a comic routine (L/S and D) and either negatively or not at all with latency. Latency correlates negatively with all other measures for Cosby and either slightly or strongly positively with all other measures for Diller. Mean funniness ratings do not provide a well differentiated measure, and, therefore, correlations involving this metric are equivocal even though their pattern is similar to that produced by other more differentiated measures such as D and L/S.

Results of the present study indicate that in only two of four Cosby routines and in only two of four Diller routines were jokes more predictable than nonjokes, and in only one of these was the difference pronounced. In addition, present results show that Cosby is a more predictable comedian than Diller, although absolute differences are not large. Concerning nonjokes, comparisons show basically the same pattern: Cosby's nonjokes are slightly more predictable than Diller's.

The major reason Cosby is more predictable than Diller has to do with the comic format used by each comedian. Cosby's material can be characterized as story or narrative; the title of a routine announces a theme,

and what follows is a logically illogical story dealing with that theme. As the story progresses, more information is conveyed and, therefore, more cues as to what the punch line might be about. Diller, on the other hand, relies more on one-liners. In this type of humor there is a short stem for each punch line, with the total routine little more than a collection of one-liners. Although jokes are loosely related to a central theme, there are fewer cues as to how the total narrative will go than is true for Cosby. In addition, Diller often laughs "raucously" during her routines and this laughter frequently either immediately precedes or follows a given punch line. Many times it is unclear as to whether the audience is laughing at the punch line, the raucous laugh, or, more likely, at some combination of both.

Turning now to the correlations, results indicate that the three measures, duration, L/S, and humor ratings—measures which can be seen to form a single factor perhaps best described as "total funniness"—all correlate positively with each other and with predictability. This positive trend between predictability and all three measures of funniness seems opposed to expectations derived from a surprise or incongruity theory if one holds that unexpectedness is a necessary condition for evoking laughter or other humorous reactions. The positive relationship found in the present study, however, is in accord with findings of an earlier study in which it was found that jokes rated as being predictable were also rated as funniest (Kenny, 1955).

What this suggests is that, within the context of single jokes such as used by Kenny or within the context of more extended routines such as used in the present study, joke predictability and *not* joke unpredictability turns out to correlate with all measures of joke funniness. Although it is not clear as to what type of theoretical principle to invoke, present results do suggest that Ss laugh longer and more frequently at jokes that are easy to anticipate than at jokes that are difficult to anticipate. It is not so much that we laugh when we are surprised but rather when we are able to anticipate an unpredictable event, and, therefore, are *not* surprised. From this point of view, laughter is a partial exclamation of achievement rather than an expression of surprise over incongruity.

Latency results, while somewhat unclear, also seem compatible with this analysis. As may be remembered, the major finding was that latency correlated either

negatively or not at all with all other measures. A further and more refined analysis, however, indicated that correlations involving predictability and latency produced different results for Cosby and Diller, with Cosby groups producing strong negative correlations and Diller groups producing moderate to strong positive correlations. The negative correlations for Cosby are easy to explain: Unpredictable endings take longer to interpret and, therefore, both latency and agreement values represent the single factor of joke difficulty. If both represent joke difficulty, we ought to expect them to correlate strongly with each other and with all other measures, which in fact they do.

For Diller, response latency seems to indicate something different, and this something has to do with Diller's comic style, a style which involves a good deal of raucous laughter—what the back of her record album describes as a "bray"—either preceding or following punch lines. By and large, students listening to Diller laugh not only or even primarily in response to punch lines, but just as often in response to her bray. Since such laughter by actual count either precedes or follows at least half of the punch lines, it seems reasonable to expect that Ss often wait a second or two in anticipating Diller's laugh before laughing themselves. For this reason, Diller evokes more laughs (L/S) under conditions where audience members take a longer time between punch line and laugh. The positive correlation, then, between latency and other measures of response to humor relates specifically to Diller's style and would seem to have little or nothing to do with joke difficulty and/or joke predictability.

## REFERENCES

- Harrower, M. R. Organization in higher mental processes. *Psychologische Forschung*, 1933, 17, 56-120.
- Keith-Spiegel, P. Early conceptions of humor: Varieties and issues. In J. H. Goldstein, and P. E. McGhee (Eds.), *The psychology of humor*. New York: Academic Press, 1972. Pp. 3-39.
- Kenny, D. I. The contingency of humor appreciation on the stimulus-confirmation of joke-ending expectations. *Journal of Abnormal & Social Psychology*, 1955, 51, 644-648.
- Mers, R. The role of predictability in humor appreciation. Unpublished Master's thesis, University of Tennessee, 1973.
- Pollio, H. R., Mers, R., & Lucchesi, W. Humor, laughter, and smiling: Some preliminary observations of funny behaviors. In J. H. Goldstein and P. E. McGhee (Eds.), *The psychology of humor*. New York: Academic Press, 1972. Pp. 211-239.
- Shurcliff, A. Judged humor, arousal, and the relief theory. *Journal of Personality & Social Psychology*, 1968, 8, 360-363.

(Received for publication July 6, 1974.)