

Smiling As A Measure of Teaching Effectiveness*

G.M. Harrington
Coe College

The frequency with which a teacher smiles in a classroom situation is an aspect of behavior which can be easily and objectively measured. It is possible that this is significantly related to the effectiveness of the teacher and it would appear that further research on the subject would be highly desirable.

Possible significance of smiling

Smiling was observed by Barr to be among the characteristics of good teachers of the social studies. It also is considered by many to be a very important aspect of a nursery school teacher's behavior. Since the frequency with which a teacher smiles can be rather readily and objectively determined, it's of some interest to know whether smiling can be used as one criterion in determining the effectiveness of teachers. Accordingly, the primary hypothesis of this study was that the frequency with which a teacher smiles is associated with the effectiveness of that teacher.

A second hypothesis also arises from the observation that other writers consider smiling to be an important aspect of teacher behavior. Adults tend to respond positively to smiling and similar social cues of affectional warmth. It may be that the frequency with which a teacher smiles is associated with objective measures of that teacher's effectiveness without being associated with objective measures of teacher effectiveness based on actual pupil growth.

Effectiveness of nursery school teachers

To test these hypotheses, seven nursery school teachers were observed for five half-hour periods and the time they spent smiling was noted. The percent of time each teacher smiled is shown in the table together with the ranking of the teachers based on frequency of smiling.

Judges' ratings of the teachers were then obtained in two ways. First four judges were asked to prepare independent rankings of the seven teachers against the criterion: "If you had a nursery school aged child, which teacher would you prefer to have teach your child?" The other judges were asked to establish their own criteria for evaluating nursery school teachers and then arrive at a ranking of the teacher jointly through natural discussion, using detailed protocols and the teachers' behavior. It may be of some interest to note that the rankings of the two judges using protocol material were identical with those the writer would have given after a year of intensive observation of these teachers. The rankings of both groups of judges are reproduced in the following table.

*From **Journal of Educational Research**, 1955, 48: 715-717.

Finally, the seven teachers were ranked on the basis of objective measures of their influence upon children in their classes. These measures were measures of child growth in those areas considered of primary importance in the overall nursery school objectives. The analysis of nursery school objectives, their statement in operational behavior form, and the construction and application of the measures of growth and described in detail elsewhere.⁰

TABLE 1
SMILING AS A MEASURE OF TEACHER EFFECTIVENESS

Teacher	Frequency of smiling		Judges' ranking		Objective ranking
	% of time	Rank	Obsv.	Prot.	
AH1	4.0	1	1	1	2
AH2	3.7	2	3	2	3
AH3	2.0	4	2	3	1
BS1	1.8	5	6	4	6
BS2	2.6	3	5	5	5
BH1	1.6	6	4	6	7
BH2	0.1	7	7	7	4

Using the rank order correlation coefficient Kendall's *tau* the frequency of smiling had a correlation of 0.62 with judges' ranking based on personal observation and a correlation of 0.81 with the judges' rankings based on analysis of the protocol material. In the case of the frequency of smiling and of the two sets of judges, rankings had a correlation of 0.43 with the ranking based on objective measures of child growth. Holding child growth constant, the partial correlation between frequency of smiling and judges' rankings of 0.54 for the observation judges and 0.77 for the protocol judges.

Need for further research on smiling

For the small group of teachers studied, the frequency of smiling would appear to be associated with a teacher's effectiveness as here measured. It was for this group as effective a measure as judge's ratings. These results suggest certain possibilities which indicate a need for further research. First, while these results are based on a nursery school situation, the high correlation of smiling with judges' rankings as compared with objective measures may indicate a substantial source of bias in judgement ratings. If this should be borne out by further study the knowledge would be of considerable use to school administrators who, lacking adequate instruments, must make many personal judgement on teacher competence.

AN OUTLINE FOR CRITIQUING AN ARTICLE

The following list of questions might be used as a basis for evaluating the above research study, or any other research article which you encounter. After the list of questions, the criteria are applied in an informal critique of one of the studies listed in the Appendix. The critique is not intended as a fixed model; it is included here to demonstrate how the questions may be used to evaluate published research.

1. What is the problem?
 - a. Is it significant?
 - b. Is it properly delimited?
 - c. Is it clearly described or implied?
 - d. Are assumptions and limitations recognized?
2. What is (are) the hypothesis (hypotheses)?
 - a. Are hypotheses made explicit?
 - b. Are hypotheses clear?
 - c. Does the study have practical or theoretical value?
 - d. Does (do) the hypothesis(hypotheses) lend itself(themselves) to empirical testing?
Can data be obtained?
3. What are the methods of procedure?
 - a. Is the method appropriate to the hypothesis?
 - b. Do procedures follow an orderly, logical sequence?
 - c. Is there evidence of review of previous studies to indicate context of this study in related body of knowledge?
 - d. Are research procedures described in detail to permit verification or replication?
4. What sources of data are identified?
 - a. Are data-gathering devices appropriate?
 - b. Are measuring devices valid and reliable?
 - c. Are sufficient amounts and appropriate kinds of data obtained?
5. What information is provided about the subject?
 - a. Are the students studied representative cases?
 - b. What evidence is given about sampling techniques?
 - c. Are there sufficient numbers of subjects (or observations)?

6. How are the data analyzed?
 - a. What logical consequences were deduced from hypotheses?
 - b. Was an appropriate significance test applied to the findings?
 - c. What was the nature of the analysis? (Verbal description, statistical manipulation, tables of frequency counts, etc.)

7. What important conclusions are presented?
 - a. Were the conclusions sound, that is in accordance with the data?
 - b. Did the conclusions conform to good logic; were they “internally consistent,” free from self-contradiction?

8. How do you evaluate the written report?
 - a. Is the style clear, objective, and “reliable?”
 - b. Is the length too short, “about right,” or too long?
 - c. Are the tables and/or figures helpful in clarifying the presentation?

AN INFORMAL CRITIQUE OF A PUBLISHED STUDY

Reference: G.M. Harrington. “Smiling as a Measure of Teaching Effectiveness,” *Journal of Educational Research* 1955, 48:715-717.

Problem. Although not explicitly stated, the implied problem of this study seems to be “How can teaching effectiveness (at the nursery-school level) be determined?” or more specifically, “Is the frequency of smiling an index of the effectiveness of the (nursery-school) teacher?” Although some critics may not agree, most teachers would probably believe that this problem is significant. The title of the study is somewhat misleading since the nursery-school level is not included (In fairness, it is never possible to include all limitations in a title.) The reader gets a superficial impression that smiling may be an index of general teaching effectiveness. Careful reading of the study removes this impression. It is conceivable that smiling is more important at some levels (especially when pupils are younger) than at others. The problem is not clearly described but is more or less implied. Assumption and limitations are not made explicit.

Hypotheses. The primary and secondary hypotheses of the study are explicitly stated. Operational definitions of the variables, however, are not clear, nor is the relationship between the variables. This is to say, which comes first: “smiling” or “teaching effectiveness?” Does the teacher smile because she is effective or is she effective because of the frequency of smiling? Which is the dependent, which the independent variable? Is the functional relationship the order “If a nursery school teacher smiles she is likely to be a more effective teacher?” or “If a nursery teacher is more effective, is she likely to smile more frequently?” Or are the two variables the concomitant (joint) effect of a (hidden) third variable? (For example, confident, secure teacher whose self-image permits her to accept herself in the teacher role and which enables her to meet

her basic needs may be both effective and may smile more frequently. The question seems to be quite complex.) Smiling is perhaps an outward manifestation of an inner feeling but some persons may smile *because of* (a complex of factors) while others smile *in spite of* (a complex of factors).

What is a smile, anyway? Can observers agree on an operational definition of a smile? The measurement used in this study is the percentage of time smiling as seen by observers who remain unidentified. Is the *length* of time a valid indicator or is it possible that a short spontaneous period of chuckles or laughter is more therapeutic than a longer period or “Cheshire Cat” smiling? Are there no occasions when smiling is an indication of something other than healthy responses? There are all types of smiles; sardonic smiles, “Mona Lisa” smiles, etc., How can one distinguish between the “happy” smile and the “professional” smile?

Shifting to the other main variable, how is teaching effectiveness measured? Three criteria are presented in the study: (1) four judges’ composite rankings of the nursery teachers using the global criterion of preference for teaching one’s own child; (2) two judges’ ranking arrived at by consensus after discussion following the reading of protocol material on the behavior of each teacher; (3) ranking based on objective measures of child growth. The limited description in the research article tells nothing about the judges and their qualifications and very little about the objective measures of child growth. (In fairness to the author, the original source may include more adequate descriptions of the judges, the protocol material, and the objective measures.)

There is considerable doubt that the variables (of smiling and teaching effectiveness) are measured validly in this study. Although more information would probably clarify the nature of the measurement used, it is more likely that considerable work on how to arrive at more satisfactory ways of measuring the complex variables is needed before hypotheses such as those suggested by this study can be satisfactorily tested.

Methods of procedure. The hypotheses suggest that a correlation design may be appropriate for this study. The Kendall *tau* rank-difference correlation (which permits a partial rank correlation) was used. Only one empirical research study was cited along with three references suggesting the importance of smiling at nursery or kindergarten levels. Although the original source of information (Harrington’s doctoral dissertation) may have more nearly adequate description of the procedures, the judges, and their qualifications, the meager description given in this account is quite insufficient for replication.

Sources of data and subjects. Although observation of behavior is a method which has been used for many years and which is necessary for some questions, observation is not *per se* reliable. It is even questionable that smiling can be adequately measured since it is somewhat qualitative. The questionable validity and reliability of the measure of teaching effectiveness have already been mentioned.

It is noted that the sample of teachers observed is very small - seven in all. No information is given about the method of selecting or whether the teachers are representative of a larger specified population. Although the findings based on such a small sample may be interesting, they are likely to be very unreliable.

Analysis of data. The logical consequences deductible from the hypotheses would seem to be as follows: "The more frequently a teacher smiles (as measured by rank among a group of seven teachers), the more likely she will be selected by judges to be the teacher of the judges' child." Or, if one wishes to assume that the criterion is valid, "The higher the rank of the teacher in smiling, the higher the rank of the teacher in teaching effectiveness as measured by several criteria." Analysis of the data consists of presentation of a table showing the rank of the seven teachers on four variables (as well as percentage of the time smiling) and the results of five correlation coefficients (Kendall's tau, and in two instances, the partial correlation). No significance test is applied. With a sample of 7, the test of the null hypothesis that two variables are not related (that *tau* equals zero) involves five degrees of freedom. (This is equivalent to stating that one restriction is placed on each of the two sets of ranks). Checking with a table of correlations with five degrees of freedom reveals that a correlation must be $\pm .754$ (two-tailed test) to be significant at the .05 level and $\pm .874$ to be significant at the .01 level. Only two correlations met the .05 criterion, that of smiling and the judges' ranks of descriptive protocol materials (.81), and the same correlation with ranking of child growth held constant (.77). It is noteworthy perhaps that the "objective" criterion of judgement of written materials produced a higher correlation than the "subjective" criterion of "choice of teacher based on personal observation." Still, can we trust the "cold, black words on the white paper" any more than the "warm, smiling bodies?" It is doubtful.

Conclusions. The author concludes somewhat modestly, "for the group of teachers studied, the frequency of smiling would appear to be associated with a teacher's effectiveness as here measured." It is granted that the correlations (at least of the judges' ranking of the protocol material) are statistically significant and relatively high. In view of the very doubtful validity of the measurements, this conclusion seems to be unwarranted. It is even more dubious that this knowledge (or the relation of smiling to effectiveness) will be of "considerable use to school administrators." Still, the correlations are high and positive and perhaps provocative. They certainly are in accord with common sense. It seems more likely that the smiling, however, is itself an effect. If teachers are able to smile, if teachers have an inner feeling of well-being, in short, if teachers have something to smile about, perhaps that fact is not unrelated to whatever may be involved in being effective at the nursery school level.

Written report. One wonders if the study should have been published at all. It is certainly not too long, but if it were worthy of publication, many more details should have been given. The study has considerable interest value and the subject is provocative.

Summary. The principal good points of this study may be briefly summarized as follows: The hypotheses are explicit. The study touches upon a topic of considerable interest and social and practical significance. The research strategy (correlation) seems to be appropriate. The chief weaknesses lie in the paucity of information about the procedures, the lack of information about judges and their qualifications, the inadequate definitions of the basic variables (in operational terms), the extremely small sample, the lack of significance tests, and the over generalizations. The title, although provocative, seems to be misleading.
