A meta-analysis procedure was used to review research on pictorial effectiveness which focused on the use of static iconic visuals in instructional materials. The purpose of this exploratory study was to provide a means for forming future hypotheses based upon a quantitative aggregation of past research. The study was concerned with differential instructional effectiveness, as measured by cognitive dependent outcomes attributable to illustration iconicity. Both automated information retrieval and manual search procedures of selected indices and references were used to identify a final sample pool of 121 data sets. Four major variables--illustrations, pacing, grade level, and achievement--and five physical attributes of treatment illustrations--production, shading, context, embellishment, and chroma--were coded. A jackknife technique was used to average 2,607 effect-size values across all studies and all variables. Results showed that illustrated treatments were more effective than verbal treatments, illustrations were most effective with secondary students, and externally paced illustrations were more effective than internally paced illustrations; color illustrations were more effective than black and white. (LMM)
TITLE: Finding the Rose Among the Thorns: Some Thoughts on Integrating Media Research

AUTHOR: Jay F. Angert
          Francis E. Clark
Background to the Problem

Through the years there have been frequent calls for research synthesizes or integrations (Broudy, 1970, 1972; Clark & Angert, 1980; Kuhn, 1962; Petrie, 1976; Randhawa, 1978). Narrative literature reviews have been quite prevalent in media research, as exemplified by Reid and MacLennon's (1967) review of 350 instructional media comparisons, and Schramm's (1977) review of "big" and "little" media, to name but a few.

Several researchers, most notably Dwyer (1978), also have synthesized in a qualitative fashion the results of research on the realism/relevant cue controversy. As is typical of narrative literature reviews, many of the conclusions which have been drawn on this subject are vague and lack quantitative precision (Clark & Angert, 1981). The series of studies undertaken by Dwyer (1978) and colleagues is unique for its consistency, numerous replications, and longevity in investigating pictorial stimulus complexity as it relates to static instructional visuals. Since guidelines for instructional visual design rely heavily on the conclusions drawn by Dwyer and possibly on the conflicting conclusions about pictorial effectiveness drawn by Samuels (1970) and others, and since none of the reviews of this body of research have employed sampling methodologies which would ensure the inclusion of most of the relevant studies, a quantitative review of pictorial effectiveness was deemed to be a useful adjunct in guiding future instructional design and empirical efforts.

Meta-analysis (Glass, 1977) is the most recently developed methodology for accomplishing quantitative research integration. Meta-analysis
research is not new to media research integrations. As one example, Cohen, Ebeling, and Kulic (1981) published a meta-analysis of 74 studies of visual-based college instruction. Unfortunately, media variables in this study were still conceptualized in terms of presentation modes or technologies. A statistical integration of pictorial complexity, with variables defined in terms of iconic symbols and coding elements, was still needed.

The Purpose

It was felt that the use of meta-analysis procedures, as described by Glass (1977), would add quantitative precision to future hypotheses and also would help reaffirm the linkage between communication theory and educational technology. Although Glass has suggested that this methodology is particularly well suited for resolving controversies arising from conflicting research results, the philosophical position which guided this investigation was that research integration is best considered a form of exploratory rather than confirmatory research. Accordingly then, the purpose of this study was to generate hypotheses based upon identified differences among the coded variables resulting from a research integration. This integration was limited to the body of research utilizing static iconic visuals in instructional materials. The major research question that guided this investigation concerned the differential instructional effectiveness, as measured by cognitive dependent outcomes (Bloom, Engelhart, Furst, Hill, & Krathwohl, 1956), attributable to illustration iconicity. The influences of subject grade level, stimulus exposure control, and time of testing were examined with respect to this effectiveness.

Procedure

Locating Reports of Research

Bibliographic search procedures used to locate pertinent studies involved both automated information retrieval and manual searching of
selected indices and references known to contain extensive pictorial
citations. Computer searching of selected data bases within the DIALOG
network was performed on two occasions. Four data bases were searched:
Educational Resources Information Center (ERIC), Psychological Abstracts,
Sociological Abstracts, and Dissertation Abstracts. A total of 558
potentially relevant citations was derived from automated retrieval
methods. The automated search procedures were augmented by a manual
search of data-based indices, annual meeting and convention programs,
and selected references. Upon completion of both the automated and
manual search procedures, an initial sample pool of 720 research
reports was identified.

Sample Selection

The pool of 720 research reports was reduced to a final sample of
121 data sets (from 151 separate reports of research) by means of a
three-tiered screening procedure. Each of the 720 reports was obtained
and read in full. These screening procedures resulted in the elimination
of non-experimental reports, of research not concerned with iconic treat-
ments or cognitive achievement, and of research with abnormal subjects,
inadequate picture descriptions, faulty research designs, or insufficient
data.

Sample Coding

Four major variables of interest to this study (illustrations, pacing, grade level, and achievement) were coded. Treatment illustrations
were coded independently with respect to five physical pictorial attri-
butes: production, shading, context, embellishment, and chroma. Pacing
referred to the locus of control over subjects' exposure to treatment
illustrations, with "external" coding referring to group-paced (experi-
menter-controlled) instruction, and "internal" coding reserved for self-
paced (programmed) materials. Grade level was coded in three ways.
Kindergarten through college and non-college adult grade levels were
coded independently and with respect to two different categorization arrangements. Achievement was coded in two ways, based upon the time of retention testing.

**Quantifying and Equating Study Outcomes**

The ultimate purpose of the meta-analysis procedures in the present study was to achieve a quantitative aggregation of findings from the various reports of research. To this end, a basic index of achievement effect, called $d$ (Cohen, 1969) or effect size (Glass, 1977) was first calculated for every treatment group comparison. All told, 2,607 effect-size measures were generated from the available data.

**Data Analysis Procedures**

Data analysis was performed in three distinct stages. The first stage completed the establishment of the data set through an aggregation of both effect-size and pseudo effect-size values (Glass, 1977). The second stage of the analysis sought to identify potentially significant main effects and interactions among the numerous levels of the coded variables. Since the 47 levels of the 11 independent variables contained far too many main effects and interactions to examine each individually, multiple linear regression was used to pare these comparisons to a manageable level. Individual effect-size values served as the units of analysis. The third data analysis stage involved using Tukey's jackknife method (Glass, 1977) to calculate a grand mean and separate means for main effects and significant interactions.

**Results**

By using the jackknife technique to average the 2,607 effect-size values across all studies and all variables, a grand jackknifed mean effect-size of .51 was derived. This figure indicated that the post-test achievement score of the average subject receiving some form of illustrated treatment was about one-half standard deviation higher than the post-test achievement
score of the average control group member receiving only the verbal treatment. This jackknifed mean was nearly identical to the grand mean effect-size of .52 reported in Clark and Angert (1981).

The effect-size statistic is more appropriately considered to be a measure of practical significance than statistical significance, and accordingly, all mean differences between levels of main effects and interactions were interpreted in terms of practical significance. Cohen's (1969) operational definitions of small, medium, and large effect-size differences were applied to the results. A $d = .2$ represented a small difference in means, $d = .5$ a medium difference, and $d = .8$ represented a large difference.

Final results were presented in the form of tables of jackknifed mean effect-sizes for all main effects and for six potential interactions. The effect-size values were interpreted in terms of the size of the mean differences between the predictor variables and verbal control groups, and in terms of the size of the mean differences between levels of the coded variables. The size of the mean differences was equated with levels of practical significance as suggested by Cohen (1969).

Jackknifed mean differences favored colored over black and white illustrations, and illustrations within a context over those out of context, at a small significance level. Medium significance level mean differences were found to favor externally paced over internally paced illustrations, and unembellished drawings over embellished drawings. Insignificant mean differences were noted between shaded and unshaded illustrations, and between immediate and retention testing and between modes of representation (photo versus drawing). The 7-12 and 9-12 grade level groupings attained the highest mean effect-size value. In general, illustration effectiveness increased from K-12 and dropped off markedly for college level subjects. Potential interactions were noted and reported for grade level versus time of testing, grade level and pacing, grade level and chroma, pacing and embellishment, pacing and representation, and representation versus chroma.
Results of the study suggested that:

1. Illustrations were most effective with secondary level subjects, a condition not altered significantly by differences in pacing, chroma, or time of testing.

2. Externally paced illustrations were more effective than internally paced illustrations, a condition not significantly affected by differences in mode of representation, embellishment, or grade level.

3. Colored illustrations were more effective than black and white illustrations, a condition especially noticeable with chirographic pictures, and with college and non-college adult learners.

4. Mean differences between modes of representation (in favor of chirographic) were significantly more pronounced for externally paced illustrations, and for colored illustrations.

In general, the results obtained in this study were consistent with the previously reported results of a meta-analysis dealing solely with the past research of Dwyer and colleagues (Clark & Angert, 1981). Since the present investigation was not conclusion-oriented, no definitive statements regarding the realism/relevant cue controversy could be derived. These results do, however, provide a means for forming future hypotheses based upon a quantitative aggregation of past research.
Selected References


