Chapter 3

Media Attribute Research

Early research into the effectiveness of media considered the effectiveness of the new medium compared to traditional approaches, but also began to investigate groups of media which differed along one dimension only, in order to determine the effectiveness of the particular attribute, such as visual motion or pictorial colour.

MOTION

One of the earliest accounts of the experimental investigation of media effectiveness is Freeman's (1924) Visual Education, published 20 years after the first public demonstration of moving pictures by the Lumiere brothers. The results are not as scientifically valid as later research, but they do confirm later findings.

McClusky's experiment represents one of the first comparisons between film and a lecture illustrated with slides. He used a film on the life history of the Monarch butterfly and compared this with two lecture conditions: a slide lecture using eight slides, each illustrating a step in the life cycle; an oral presentation illustrated with two pictures and two blackboard sketches. Each was presented to 20 pupils in grades 6-8 in two schools and lasted 12 minutes. The results failed to show any difference between the methods.

Another of his studies compared the effectiveness of the motion film A Story Of A Mountain Glacier with a 13 minute slide lecture using five slides,

two wall charts and a blackboard illustration. In this case there was a clear superiority for the lecture method.

Brown found similar results when comparing films and filmstrips for teaching factual information about the physiology of seeing in high school. In the filmstrip group discussion was free and questions were asked both by the teacher and the students. A multiple choice test indicated a superior performance for the filmstrip group and Brown concluded that this was because of the greater exchange of comment within the teacher-paced filmstrip group.

Clearly in these comparisons the presentations are not strictly comparable and where difference do occur they may be accounted for by additional opportunities for interaction or presentations which took account of the experimental groups' needs. This is particularly the case with Carson's (1947) study which compared a film on the life of cowboys with a filmstrip and abbreviated filmstrip on the same subject. The abbreviated strip consisted of every third frame of the original 100 frame strip, but in both cases the same commentary was read to the group. Both filmstrips were superior to the film, but as there was no opportunity for any 'greater exchange of comment' in the filmstrip group it is more likely that the actual content was different, and the report indicates that the subject matter 'approximated fairly closely to that of the film strip'! What is fascinating is the difference in performance of the two filmstrip groups. Excising nearly 70% of the filmstrip frames reduced the performance by only 15%, and had the frames been retained on a less arbitrary basis the two versions may have produced a closer result. This tends to confirm findings that suggest that visuals often play a comparatively minor role in conveying information compared to the sound track, especially when the tests used are essentially verbal.

A more satisfactory approach to determining the effectiveness of the visual motion attribute was undertaken by Twyford (1954) as part of the Pennsylvania State University instructional film research programme under taken for the US military, who had used film extensively during the Second World War for training. The topic investigated was methods of riot control, under the title Military Police Support in Emergencies, and introduced the problems of training soldiers to cope with such complex situations as restless, disturbed city populations, agitated groups, mobs and rioting crowds. Real rioting crowds could not be brought readily to hand and film was an obvious substitute. The film had a Hollywood budget and expensive crowd scenes organised in an American city. Twyford does not mention just how expensive the production was, but he points out that during the final approval stages of the film questions were raised about the necessity for such an elaborate production. The question

of simpler and less expensive production methods was raised and Twyford was charged with determining the effectiveness of other methods. Twyford's group suggested an alternative approach using stock film or newsreel coverage of riots, and even the use of still pictures if motion film was not available. The project eventually compared the Hollywood style film with two filmograph versions. A filmograph is similar to a sound filmstrip, but is produced on 16mm film and can be shown on a standard projector. The film is used to convey the sound track but the images are all still pictures. In all three versions the soundtrack was identical. One filmograph was based on the motion film and consisted of individual still frames taken from the original. The second filmograph was made up of stock still pictures of riots taken from news libraries, or simple diagrammatic representations of troop movements. The groups of recruits in the late stages of basic training were tested using a 42 question test with 10% of the questions using pictures. The full motion version scored 4% more than the filmographs, which were equally effective. The difference was statistically significant, but educationally insignificant, at a cost of 50-100 times that of the filmograph productions.

This is one of the earliest well-controlled experiments that shows that motion aids such as film and television will not automatically improve student performance when compared with simpler aids such as filmstrips. The essence of this argument is that film and tv can teach many different groups and subjects about as effectively as traditional methods, but so can simpler aids such as sound filmstrips or sound tapes with booklets, and these simpler aids cost less to produce.

It was argued at this point that although some concepts did not benefit from the motion attribute and therefore could be conveyed accurately by still pictures, some concepts were inextricably linked to visual motion and their essence could only be conveyed by aids capable of showing motion. Perhaps time, motion itself, and certain spatial concepts would benefit from motion aids.

A start was made in this direction by Houser et al (1970). They produced a series of shapes of various sizes generated in a random fashion, with either a smooth or angular outline. Some of the shapes rotated, while others remained stationary. Those that moved were accompanied by the nonsense word GEF and the still shapes by the word FIK. Having seen a series of GEF and FIK exemplars subjects had to label a further group as being examples of either GEF or FIK. There were two groups: one saw the moving shapes in continuous smooth motion recorded on motion film; the other saw a series of slides illustrating discrete changes in the rotation of the shape. The results showed an

89% accuracy score for identifying the motion concept for the motion film group; and 45% accuracy for the group seeing motion as a series of slides, in which the shapes rotated in large movements. The authors concluded that "it is clear that in the case where motion is a defining attribute of a concept it is better to present that concept using motion picture film rather than by a non-motion medium such as slides."

It worked for simple rotating shapes, but would it work for concepts such as space, time or motion in such subjects as biology or physics? The Houser team continued their work with Wells (1973) to investigate the effectiveness of three visual media in teaching concepts involving time, space and motion in biology. The three media were: motion film, slides and sequential still photographs. Sam Postlethwaite, the originator of the Audio Tutorial system, was also involved in the research which posed the question: within a flexible system such as Audio Tutorial, are there guidelines for the selection of specific media to present specific concepts?

They reasoned that after studying the effects of specific media for different concepts, it should be possible for curriculum planners to determine which media should be used when presenting specific concepts. For example, when more than one medium can represent all relevant stimulus dimensions, choosing which to use could be done on the basis of more specific information concerning the interactions of media and type of concept.

Concepts involving the manipulation of time included slow motion pictures of the explosive release of seeds as ripe capsules split open; space concepts included enlargement by photomicroscopy of sequential sections through a corn fruit; and acceleration-deceleration of motion was conveyed using the movements of chromosomes during mitosis, prior to cell division.

The authors reasoned that each visual medium used had advantages and disadvantages for conveying concepts of time, space and motion. For example, the continuous nature of motion pictures creates the illusion of movement over time, changing the focal length or distance to the subject allows variations in space. However, motion films require sophisticated equipment, are costly to develop, maintain and replace, and in general, burden time and budgets. They are also inflexible in that students must view them where and when projection equipment is available.

The slides and photographs in the experiment were taken from the motion film, the sides being shown using a timed slide changer and the still pictures used a visible clock to indicate any time changes.

Figure 3.1: Mean score for 3 concepts and 3 presentation modes

	motion film	slides	sequential photographs
Time	8.52	7.67	7.47
Motion	11.61	11.47	10.67
Space	8.88	9.88	9.72

Audio Tutorial classes tend to be large and a total of 600 students took part in the experiment. The test consisted of 15 objective questions which included photographs and diagrams.

There was no statistical difference when the pooled scores were evaluated, showing that no medium gives superior results for all concepts. For time concepts, motion film was statistically superior; for motion, film and slides were superior to photographs; and for space there were no statistical differences. The authors concluded that motion pictures, slides and sequential still photographs differ in their ability to convey concepts of time, space and motion. And suggested that the medium showing the highest performance, or the least expensive of those for which a difference had not been found, should be used in the production of instructional materials: motion film for time; slides for motion concepts; still pictures for space concepts.

I do not find this to be an acceptable interpretation of the results. Where there is a statistical difference, showing that the groups do differ, the actual differences are small, representing, on average, 6% of the total score. My argument is that the differences are educationally trivial and that if we wish to raise student performance to an **educationally** significant degree we will not find the techniques within these microcosmic manipulations of the educational environment.

Procedural learning, like the learning of motion concepts, was also seen to be a special case. Procedural learning is defined as learning to carry out a series of acts or operations in the correct order. In such cases performance is measured by the time taken to complete the operation and the number of errors made.

The first experimental investigation of the importance of motion as a variable in studies relating to procedural learning was conducted by Roshal in 1949, again as part of the Pennsylvania military research. Three different knot tying tasks were presented in various ways, the ability to correctly tie the knot being measured immediately after viewing. Roshal based his research on the

behaviouristic principle of stimulus generalisation, which stated that during practice a response is set up not only to the stimulus but to a family of similar stimuli, the strength of the response to any particular stimulus being directly proportional to the similarity of the particular stimulus to the stimulus used in training.

On this basis Roshal deduced that for greatest efficiency of learning, the stimuli presented in the film should be as similar as possible to the stimuli which would occur in the actual performance of the desired learning. This implied, according to Roshal, that the film representation should approximate, as nearly as practicable, the details of the actual stimulus in terms of colour, shading, movement etc.

In any perceptual-motor act, there are continual changes in the object and also in the subject's position eg. his hands. From this Roshal argued that if film-mediated learning is to represent the actual stimulus of the test situation, it must represent all changes as exactly as possible. Thus, learning a perceptual-motor act, from films, will be more effective as the film representation approaches a representation of the learner himself performing the acts to be learned.

The experiment used several versions of a presentation showing knot tying, two with full visual motion and two showing successive still pictures. Three knots of graded difficulty were presented: the bowline, sheet bend and Spanish bowline. The films were shown to 4,200 recruits!

Before looking at the results it is as well to point out the value of the pictorial element for this task. The following is a description of the procedure for tying the bowline, taken from the film:

> Form a small loop. Pass the end of the line through this loop, forming a larger loop below. Pass this same end around the long end. Then pass it through the small loop again so that it lies alongside itself. Hold these two parts and pull the long part until the knot is tight.

Roshal's results showed that on the basis of 1 point for each completed knot (a rather crude measure) all the motion film versions were superior for the two more complex knots, but the static versions were equally as effective. Roshal concluded that for simple tasks it may only be necessary to provide a good illustration of the thing acted upon, with the learner providing the intervening movements from his cognitive repertoire. But for more complex tasks a training film will be more effective.

Roshal concludes that his major hypothesis is confirmed and that, with regard to motion, a training film designed to teach a perceptual-motor task will be more effective when it shows all the movements involved in performing the task, than when it shows merely a series of static shots portraying successive stages of the task. I'm not sure that this conclusion is wholly acceptable in view of the fact that one still visual version of the demonstration was as effective as the motion version. However, taking the evidence for all 3 films, the conclusion is justified for the more difficult tasks.

A similar experiment was conducted by Laner in 1954, in which a moving film and a filmstrip, composed of line drawings based on frames from the film, were compared in respect of their efficacy in imparting instruction about a manual task. Their respective effects were tested by making the subject actually perform the task. In his introduction Laner states that the increasing use of pictorial, in place of verbal, methods of teaching and the widely expressed preference for pictorial methods, appears to arise out of a fairly general held assumption that media of communication, which are deemed to represent reality, are more likely to make a greater impact on the learner and produce quicker, greater and more durable learning effects.

Two hypotheses are of concern here: the **stimulus generalisation hypothesis** is concerned with a correspondence between learning and testing situations, and the **sign similarity hypothesis** is concerned with increasing the reality of the display, irrespective of the test situation and leads to the same emphasis on reality whether the test consists of writing descriptions, drawing diagrams or actual performance.

Laner states that motion film is typical of what is meant by a highly realistic display and that if the assumptions of the sign similarity hypothesis hold good then the properties of continuity and represented motion should give the film a high advantage, particularly in the case involving manual operations forming a skilled performance.

The task used in his experiment consisted of the dismantling, repair and reassembly of a sash cord window. The film group saw a commercially produced film and the filmstrip group viewed 51 still drawings based on actual frames from the film. The commentary was the same for both presentations. The subjects were mainly RAF servicemen, although a group of university students was also instructed, although they only received the film mediated instructions. A model sash-cord window was used for the actual test, and subjects were required to dismantle the window, repair a broken sash-cord and reassemble it. A close record was kept of each individual performance, providing details of responses on each of 23 sub-operations. The mean

percentage operations correct and the mean percentage major errors showed that any differences were not statistically significant, leading Laner to conclude that the two types of display used were equivalent. Inspection of the results shows that in general the mean percentages of correct responses are rather high and errors correspondingly low. This in turn indicates that both displays were exceedingly effective in conveying the task.

Laner concluded, from his detailed analysis, that the performances mediated by the two types of display, film and filmstrip, did not indicate that continuity in pictorial representation or the inclusion of movement representation necessarily produces better results in terms of correct responses.

Clearly from this result we cannot be sure of the efficacy of motion aids even with task performance as the criterion measure. However, two further experiments do offer such an interpretation.

The first was conducted by Silverman in 1959 and its purpose was to furnish experimental evidence of the differences in training effectiveness of two types of overhead transparencies — static and animated. He argued that if animated devices improve the understanding of motion concepts then the number and types of moving parts should be an important variable such that a device with many moving parts would be better served by animated materials.

A second major variable identified by Silverman was the method of testing. He reasoned that it does not follow that the type of knowledge imparted in many training situations can be measured exclusively by means of verbal techniques. If the purpose is to impart knowledge of nomenclature, them verbal paper and pencil tests would be appropriate. If the trainee is required to assemble a device then a non-verbal performance test would be appropriate.

In the study 150 students were required to follow 3 presentations dealing with the loading and firing of 3 different devices: a pistol, a carbine and a rifle. Instruction was by a tape recorded audio presentation synchronised with overhead projector transparencies, consisting of 11, 6 and 5 moving parts respectively. The static presentation utilised the same transparencies but the movements of parts of the device, from one position to another, were not shown. All conditions were essentially the same but with the absence of motion in the static display. Three different tests were used in this experiment:

> test one consisted of multiple choice questions dealing with knowledge of function;
> test two was a written nomenclature test;
> test three was a performance test, in which the subjects were required to load and fire the device and to release the safety catch

when it was set ready for firing. The results for the third test were expressed as the time for completion of each operation.

For the verbal tests there are slight differences in favour of the animated versions (don't forget they are error scores), but none of these approaches statistical significance. For the performance test there are consistent differences between the animated and static transparencies for each device, with the animated transparencies showing the shorter tines. The differences are statistically significant. Silverman predicted such an effect but also points out that he would have expected the greatest difference to occur for the pistol which, you will recall, had 11 moving parts. However, the results indicate that the rifle with the smallest number of moving parts produces the largest difference, nearly fifteen seconds, or 63 percent longer for the static group.

In his conclusion Silverman states that the results of the two paper and pencil tests indicate that there is no difference between animated and static diagrammatic presentations, but that the performance tests did show that the animated transparencies were more effective, and that this result was equally applicable to all devices. From further analysis of the data for the rifle performance he suggests that, when relationships among moving parts are involved, the animated transparencies are clearly more effective.

These results clearly reflect the differential effects of the testing situation, although they do conflict with Laner's non-significant result. This is possibly because we have yet another source of variability within the performance test, which is dependent on the method of scoring the test: either by timing or number of errors. Laner did attempt to record performance times but, with his 23 subsections, it was found to be impossible to give an accurate time of performance if errors were also made during the procedure.

One suggestion for assessment of procedural tasks, such as those studied by Laner, Silverman and Roshal is to determine the number of trials required to reach a perfect performance, with a timed value for the first perfect performance.

A further study carried out by Spangenberg in 1973 moves in this direction in that two tests were used, but the results are confounded because the intervening period was not occupied with a repeated instructional presentation. At the present time there have been no experiments which have implemented the teaching to criterion or mastery performance type of experimental paradigm.

Spangenberg's work is worth considering from another viewpoint in that he succeeded in measuring times for individual or suboperatioms of an overall task, which Laner attempted but failed to do.

In Spangenberg's experiment 40 army personnel were trained to disassemble a machine gun — yet another case of military interest in instructional techniques. Two presentation methods were used: a videotape showing the disassembly technique and a still picture version of this videotape, using the same commentary, but with selected still photographs substituted for the original continuous motion presentation.

The still shots were selected as providing maximum assistance to the learner, rather than to imitate exactly the motion sequence camera angles.

The task was divided into 9 sub-operations, and after the presentation of instruction for each of these the students were required to perform the demonstrated operation. The time taken for completion was noted, as were specific errors. If the subject was unable to complete the operation the experimenter intervened and performed the specific sub operation, this, of course, acted as a form of instruction and confounds any future performance. When all 9 sub operations had been completed the subject was provided with a second identical weapon to disassemble. In this second trial only the number of errors was evaluated.

The difference in total time is statistically significant, as are 4 of the nine sub-operation times. This means that there is no difference, in a statistical sense, on five of the sub operations. Spangenberg does not provide data for the second trial, that is the one following on directly from the period of instruction plus practice, other than to say that the number of sub-operation errors on the second trial did not differ significantly, with 14 errors for the motion aid and 17 for the still sequences.

The experiment was repeated with 80 subjects in an attempt to understand more fully the differences in motion and non-motion aids. Two further instructional conditions were added which utilised white cueing arrows intended to emphasise critical movements in each sub-operation. There was no difference between the version with arrows and the standard sequences, so I will not discuss this aspect further. However, modifications were made to the still sequences which had proved to be less effective than the motion sequences in the first experiment. The new versions of these still sequences were designed to improve their effectiveness.

Spangenberg's interpretation of the results suggested that overall there is a significant difference in the total time taken to perform the 9 sub operations in favour of the motion sequences. Of the 4 sub operations which had previously

been less effective for the static version, and which were subsequently modified, only two showed the motion condition as superior in the second study. However, three other sequences, which had previously not been significantly different, and which had therefore not been modified, showed the motion sequence as superior; while one now showed the still sequence as superior.

Spangenberg concludes from this that a performance improvement followed the redesign of the still sequences shown to be inferior to the motion sequences in the first study and that this improvement in performance suggests that, following pretesting and re-design, still sequences can provide equivalent learning to some motion sequences. That is, the performance difference between some of the sequences may not be attributable to motion as such. But is this a valid interpretation? If it is, how do we account for the following changes: in sub-operation 1, which now shows a significant difference in favour of the still sequences; in sub-operations 3 & 8 for which the significant difference has now disappeared; as well as sub-operations 5, 6 & 9 which had previously shown no significant differences, but for which significant differences appear in the second experiment?

COLOUR

It is also worth, at this point, considering the experimental evidence concerning comparisons between colour illustrations and their counterparts in monochrome. We will also look at preferences expressed for one or another of the visual displays in order to determine the overall value of adding colour. There are, of course, cost factors involved, which at the simplest level are reflected in the cost of a colour tv set being two or three times that of similar monochrome machines. Overall the implications for education will certainly be centred on cost-effectiveness criteria.

Seth Spaulding reviewed the research on pictorial illustration in 1955 and started by looking at children's preferences. One of the earliest studies mentioned is the 1932 study of Mellinger which indicated that children aged from 6 to 11 preferred coloured, realistic illustrations. This result was confirmed in William Miller's research in 1938 and Marie Halbert's work in 1944. Mabel Rudisill's 1952 research attempted to determine children's preferences for colour versus other qualities in illustrations because she felt that colour is an important factor in determining the cost of illustrations, and hence of text books. She chose five types of illustration.

Three different subjects were used for each of the five illustration types, resulting in a total of 15 different pictures.

1. Uncoloured photograph
2. Coloured photograph
3. Coloured drawing, realistic in form and colour
4. Outline drawing, realistic in form, but outlined in colour without regard for realistic effect.
5. Coloured drawing, conventionalised in form, decorative but unrealistic in colour

Approximately 1200 children aged 6 to 12 were asked to state their preferences and the opinions of 775 adults were also obtained, concerning which of each pair of illustrations they would choose for a child of 6 or 7 years. The general results of the study are:

1. If two pictures are identical in all other respects, most children prefer a realistically coloured one to an uncoloured one.
2. If different pictures include the same subject matter and the same colours, most children prefer the one which is treated in such manner as to give the truest appearance of realism or lifelikeness.
3. If different pictures include the same subject matter, most children prefer an uncoloured one which gives them an impression of reality above a coloured one which does not seem to confirm to reality.
4. If different coloured pictures include the same subject matter, most children prefer a less colourful one which gives a greater appearance of reality above more colourful ones which appear less lifelike.
5. There is an increase in unanimity of those preferences with increase in grade level up to Grade IV. This greater unanimity of preference of older children, as compared with younger children, is believed to be due to the former's greater capacity for discriminating reality.
6. Typical adult opinion over-emphasizes the importance of colour per se and underemphasizes the importance of other qualities in illustrations for children.
7. Photographs of excellent quality, both coloured and uncoloured, deserve much wider use than they are at present being given in illustrations for children.
8. These findings do not justify the statement of a general principle as to comparative preference between photographs and realistic coloured drawings because the examples of the two types used in the present study were not of equal quality.

9. Consideration of the first four conclusions above, together with children's stated reasons for choosing or not choosing certain pictures, suggests:

a) In looking at a picture, a child apparently seeks first to recognize its content.

b) Any picture assuming a certain content proves satisfying to the child in proportion to its success in making that content appear real or lifelike. Whether it is coloured or uncoloured is less important than the appearance of realism.

c) A perfect visual representation of realism includes colour, and colour in pictures proves satisfying to the child in proportion to its success in increasing the impression of realism or lifelikeness.

It is clear that realistically coloured illustrations are preferred to those which are unrealistically coloured and that if the colours are constant, the realistic style is preferred. A very interesting aspect to this research is that, if the subject matter is the same, children prefer an uncoloured illustration, which gives them an impression of reality, to a coloured one that does not conform to reality. Also, there is evidence which suggests that adult opinion over-emphasises the importance of colour per se and under-emphasises the importance of other qualities.

Well, this is an interesting study but there is something missing from it if it is to have any instructional implications. The most important of these, in my opinion, is the lack of any measure of the effectiveness of the various types of illustration in a teaching situation.

The question being raised here is — can we justify the increased cost involved in producing highly realistic colour illustrations when there appears to be no commensurate increase in teaching effectiveness? The next study, reported by Vander Meer in 1954, attempts to throw some light on this matter. Vander Meer's work has been described as demonstrating a rigorous methodology which sets the standard for similar studies. Vander Meer was working at Pennsylvania State University as a member of the Instructional Film Research Programme, funded by the military.

The first experiment involved 500 students, 14 and 15 years old. One half of the students saw colour versions of 5 films, whilst the other half simultaneously saw black and white prints made from the original colour materials. The films were commercially produced titles including: Maps are fun; How man made days; Rivers of the Pacific Slope; Snakes, and Sulphur & its compounds.

Two types of tests of perceptual and conceptual learning were developed for each film: non-verbal and verbal.

Non-verbal tests were necessary, Vander Meer reasoned, because some films lend themselves to the use of pictures, diagrams or samples as supplements to the verbal questions or statements. For example, in the "Rivers" film a map with rivers identified by numbers was included in the tests. In the film titled "Sulphur and its components" a set of sealed glass vials containing sulphur compounds were used in conjunction with multiple-choice questions. A similar arrangement included colour photographs of snakes as test items for the "Snakes" film. Twenty-eight pictures of 14 snakes, not all of which appeared in the film, were used as test items, together with several landscape slides showing habitats of several snakes.

The Verbal tests represented the larger number of items and were of the conventional multiple choice type in which the learner was required to select a response in terms of completing a purely verbal statement or answering a verbal question. The same tests were used as pre- and post- tests and as measures of retention.

Vander Meer identified several common reasons for using colour in selected instructional films. These were:

1. Colour may be an important cue in learning what the film is intended to teach. eg. a film to teach the identification of poisonous snakes of the United States might be made in colour if the colour of the snakes is an important cue in identifying them.
2. Contrasting colours in graphic presentations could be used for emphasis to make certain things stand out and one example that Vander Meer gives is that different pipes in industrial processes could be identified using contrasting colours which might suggest certain characteristics of the process.
3. The third reason was that colour may be pleasing to the learner and its aesthetic appeal may have an indirect effect in promoting greater learning even though the colour itself provides neither important meaningful cues nor emphasis.

As each test item was developed it was designated as relating to one of the previously mentioned reasons for using colour in film production. Three quarters of all test items were in the last category, the use of colour was purely aesthetic. In addition test items were classified as to whether the source of their correct answer was in (a) film commentary but not the pictorial content, (b) the

pictorial content but not the commentary, or (c) both the commentary and the pictorial content. This meant that Vander Meer obtained 7 scores or each test: a total score, 3 part scores related to various uses of colour in films and 3 part scores related to the source of the information, in other words whether it came from the commentary, the pictures, or both

Students were also asked to rate each of the films they saw on a like - dislike scale. A separate measure of the two groups liking for another film "Birds" indicated that they were substantially the same enabling Vander Meer to conclude that any differences during the experiment could be attributed to the presence or absence of colour.

Two experiments were conducted by Vander Meer and the results for the verbal tests indicate that in only one case is there a statistically significant result in favour of the colour film version!

Unfortunately mean scores are not given in the research report and so we are unable to judge whether the differences are significant in an educational sense. An especially interesting result here is that after the 6 weeks delay the differences between the two groups are more noticeable, giving rise to statistically significant differences for 3 of the films.

Figure 3.2: Verbal Tests: Colour vs B&W

	C	BW	t-Ratios 1 Immediate Recall Tests	t-Ratios (1) Delayed Recall Tests
Maps Are Fun	114	137	-1.13	-.41
Rivers of the Pacific Slope	108	135	.69	1.28
Properties of Sulphur	112	150	1.99*	2.65**
Snakes	115	144	-.29	3.04**
How Man Made Day	117	145	-.63	2.08*

1 minus sign indicates that the difference favours the black and white versions

* significant at the 5.0 per cent level of confidence

** significant at the 1.0 per cent level of confidence

Figure 3.2: Non-verbal tests: Colour vs B&W

	C	BW	t-Ratios Immediate Recall	t-Ratios Delayed Recall
Identification of Snake Slides	111	131	-1.01	.29
Identification of River Locations on Outline Map	94	112	-2.62*	.59
Identification of Materials in Glass Vials. (Sulphur forms and distractors)	133	169	-2.67*	

* Significant at the 1.0 per cent level

The non-verbal test results, reverse the statistically significant results for the verbal tests with two of the three films favouring the b & w version. However, the differences do not persist and the delayed recall test indicates no difference between the two versions. The three non-verbal tests are also rather unreliable tests, according to Vander Meer's tables which show them to have much lower test reliability than the verbal tests.

The mean rating scores for the two versions of each film favour the colour group, although the difference is only statistically significant for the Sulphur film.

However, the group viewing the b & w films show a lower rating for the "Birds" film, which was presented in colour to both groups and we may conclude that this group clearly had a tendency to rate films lower in the first instance. Another interesting result is that none of the experimental colour films was rated by the colour group as being liked as much as the pre-experimental film.

A second experiment omitted the "Maps" film and did not measure the effects of delayed recall. For measures of immediate recall there were no differences between the 2 groups, confirming the results of the first experiment.

Figure 3.4: Colour vs B&W, second experiment

	Sex	N	Items No.	Mean B & W	Mean Colour	Critical Ratio
Score Total	Fem	107	185	99.7	100.0	.2
	Males	92		96.6	100.4	.6
Items Answered In Video Only	Fem	107	34	100.4	99.7	-.4
	Males	92		98.0	102.0	2.3*
Items Answered In Audio Only	Fem	107	82	101.5	98.7	-2.l*
	Males	92		99.3	100.0	.4
Items Answered In both Audio and Visual	Fern	107	69	98.4	101.8	2.4*
	Males	92		100.8	99.3	-.9
Items Showing No Obvious Use of Colour Other Than Aesthetic	Fem	107	140	100.0	99.9	-1.3
	Males	92		99.5	100.3	.6
Items In Which Colour was used to Emphasize Differences	Fern	107	31	98.4	101.5	1.7*
	Males	92		99.5	100.2	.3
Items in which Colour is an Intrinsic Learning Cue	Fem	107	14	101.5	98.5	-1.3
	Males	92		100.6	98.6	-.9

* Significant at the 5.0 per cent level

The combined results for immediate recall for the four films is given in terms of the 6 test item categories for both male and female subjects. Overall, there is no consistent pattern of sex differences. For items in which colour was used for a specific purpose there is only one significant difference in favour of colour — and what is especially fascinating is that those items in which colour is an intrinsic learning cue show a tendency towards the superior performance of the b & w films, although the differences are not statistically significant.

Vander Meer concluded that colour appears to have contributed significantly to learning of only those items in which colour was used to emphasise differences, though he fails to note that this is so only in the case of female subjects.

In terms of preference rating the second experiment shows two significant results both in favour of the colour version.

When correlation coefficients were calculated to determine the colour factor in total liking for a film and the difference in learning gain attributable to colour the results led to the conclusion that there is little relationship between degree of preference for colour and the learning gain attributable to it.

The conclusions reached by Vander Meer are:

1. The use of colour in instructional films which may superficially seem to 'call for colour' does not appear to be justified in terms of greater learning on the part of those who view the films. If colour is to be used effectively in films there must be careful preproduction consideration of the probable psychological impact of specific uses of colour upon the learner.
2. The contribution of colour in film seems to be related more to the retention of learning than to the immediate acquisition of learning. Those who view colour films may not learn more than those who view the same films in black and white but they are likely to forget less of what they learn than those who view the black and white films.
3. While 'liking' for a film and learning from the film are probably positively related the influence of colour in determining such liking is not great enough to warrant its use as a means of increasing liking and therewith increasing learning. The 'aesthetic value' of colour as a contribution to learning effectiveness appears to be less than that of the intrinsic appeal of the subject matter.
4. Sex differences with regard to preferences for and learning from colour films are slight.

These clearly indicate the lack of consistent differences attributable to colour. In reading the report, with its total and partial test scores, related preference measures and the separation of data according to the sex of the subject, one gets the feeling that the exercise has been a rather traumatic one for the investigator.

Don't forget that this report represents a pioneering study, much of the work was conducted during 1949-50 at a time when new colour materials were making the use of colour educational films feasible. Before World War 2 such materials were only available for large volume cinema films. I am sure that Vander Meer together with many of his colleagues working on the Pennsylvania Film Research Programme, expected to see large gains in favour of the colour film group and were surprised when they did not appear.

In this context I would now like to comment briefly on a similar project undertaken at Yale University and reported in 1958 by May and Lumsdaine. In chapter two of the "Learning from Films" report they discuss the effects of pictorial quality and colour. The studies discussed represent preliminary efforts to assess the importance of factors generally regarded as entering into the degree of polish or quality of the pictorial component of teaching films, both factors being related to the cost of producing and printing films.

The Yale team produced a colour film 'Seasons', which dealt with the causes of seasonal change, and was to be used to investigate the efficacy of colour instructional films. During the production phase a story board was produced as a guide for the eventual production of animated and live colour footage. The story board consisted of very crude b & w pencil sketches for each scene. In order to aid in visualising the content of the final film a so-called pencil test running reel was made by photographing these sketches on motion picture film in the planned sequence.

May and Lumsdaine emphasise the fact that this crude pencil-test version was never intended to be a teaching instrument - it was merely to give the producers a better idea of sequences and presentation, as they might appear in the final version. Such a pencil-test version is quite inexpensive and an estimate of total cost at the time was about ten percent of the total cost of standard 16mm colour-film productions of instructional sound motion films.

The authors go to great lengths to describe the poor production standards of the pencil test version, describing it as extremely crude, with a few jerky animation effects, almost all visuals being pencil sketches, the photographic quality being quite poor by any usual standards of technic al or artistic merit. The reason for these denigrating descriptions is pure bathos because the authors

proceed immediately to describe how, before the final film became available a silent print of the pencil test version was shown to a sixth grade class, with a staff member reading the commentary. The post test scores were later compared with the performance of a similar class who viewed the full colour version and the astounding result was, as you've probably guessed that the learning from the crude, jerky b & w version on was substanially as great as from the full colour version. These surprising results, which must have shocked producers, sponsors and researchers were at first only accepted as being very tentative, because the groups compared were not selected to assure equivalence or provide any valid measure of error.

In order to be more definite concerning the unexpected result a further experiment was conducted, for, as the authors say, the implication that a crude penciltest film might be substantially as effective as a polished colour film costing many times more to produce was provocative enough to justify obtaining further comparative data under better controlled conditions.

The second experiment utilised 4 classes of fifth grade pupils, each class being randomly divided in half; one half viewing the pencil version, the other viewing the colour version. Immediately after seeing the films pupils took a 91 item test on knowledge of the subject matter.

Lumsdaine concluded that the difference between the two mean scores is so small that it is interpreted as the result of mere chance fluctuations — in other words they are not statistically significant. The result of the experiment clearly pained the investigators who attempt to deny their clear implications. Accordingly they argue that the results may be interpreted merely as showing that such

Figure 3.5: Pencil test running reel (PTTR) vs standard colour film

Number of Test items correct (Total 91)	PTTR	Standard Film
70 - 79	4	7
60 - 69	8	8
50 - 59	14	7
40 - 49	11	17
30 - 39	6	6
29 or less	7	5
Mean Score	49.0	49.6

a crude presentation can at least equal in teaching effectiveness the more polished colour version (costing more than ten times as much to produce).

In the ironic conclusion the authors suggest that the clearest case for use of colour is where colour cues are essential for a discrimination that is to be learned — for example in learning signal code flags or the identification of minerals, flora and fauna etc.

In films and printed materials colour differentiation is also, obviously, advantageous, according to May and Lumsdaine, when multiple colour codes keep otherwise confusing visual elements separable, for example in complex electrical schematic diagrams.

Well, some of these elements were investigated by Vander Meer, for example the identification of fauna, without any observable effects and others were investigated in a later study by Kanner and Rosenstein in 1960.

The eleven lessons used in their study are listed and it is worth considering whether or not you would expect differences between b & w and colour displays for the different topics, before you look at the results:

1. **Complex Wave Analysis:** Colour used in charts to emphasize wave differences
2. **Cathode Coupled Multivibrator:** Schematic diagram and voltage waveshapes used. Little use of colour.
3. **RF Signal Generator:** No colour
4. **RF amplifier:** Various coloured chalks used on green board to explain circuitry
5. **IF amplifier:** Various coloured chalks used to explain circuit operation. Colour code of IF transformer shown in colour.
6. **Foster-Seely Discriminator:** No colour.
7. **Phase Sensitive Rectifier:** Multi-coloured schematic used for emphasis and differentiation in explaining circuit operation
8. **Automatic Range Tracking:** Colour used in black diagrams and animated chart.
9. **Photo Materials for Copying:** Colour used to illustrate proper selection of copying material
10. **Introduction to Colour:** Coloured photos, charts and diagrams used extensively
11. **Photographic Films and Characteristics:** Extensive use of coloured photos.

The study was undertaken to evaluate the need for colour rather than b & w instructional television in the US Army. The report indicates that reliable colour television equipment was available but that the cost was higher than monochrome equipment although costs were expected to fall with technological developments as indeed they have done. Nevertheless the report continues, colour television equipment will probably always be more expensive to procure and maintain than monochrome equipment and, therefore, to justify the cost of using colour television for training it would seem reasonable to expect some return in teaching effectiveness and learning not possible with monochrome television.

The experimental study beamed the eleven lessons from a mobile colour television facility into two classrooms. The randomly divided subjects viewed the same programmes at the same time. The difference between the two groups was that one viewed the lessons on colour receivers while the other group viewed them on monochrome receivers.

Immediately following a lesson the subjects were tested using multiple-choice questions. Every effort was made to incorporate colour items into the tests, in other words items in which colour seemed to play a relevant role in understanding or answering the question. An example would be showing various coloured resistors and asking for their value in ohms.

A total of 368 trainees took part in the experiment and pairs of subjects were matched on electronics aptitude or general technical scores and then were randomly assigned to one of the two experimental conditions. Under the study conditions each instructor simultaneously taught both the colour and monochrome sited classes and where colour seemed to be important, the colours were pointed to and describe so that the monochrome groups would hear the names of the colours being used

The authors emphasised that colour was used largely as it was found in ongoing conventional classroom instruction

The major question in this study was whether colour or monochrome television produces any difference in learning, and the results clearly do not provide evidence for such a difference.

Ten out of 11 comparisons show no significant differences and the single statistically significant result is considered to be unimportant in view of the overall picture and small differences in test performance. The overall mean scores are remarkably similar, bearing out the results of May and Lumsdaine.

A further analysis was made to see what differences, if any, were achieved by trainees of different aptitudes. The groups were divided in half, based on aptitude scores, and the results separately analysed. Only two statistically significant differences emerged: one favouring colour for the low aptitude

group, the other favouring monochrome for the same group. The authors' statement, that there is a statistically significant difference in performance of high and low aptitude groups with respect to the type of television used, is clearly erroneous. There is a slight tendency for the high aptitude group to learn more from monochrome and the low aptitude group to learn more from colour, but the results have no significance and are purely chance fluctuations. It is clear from the results that it is aptitude rather than type of television instruction which pedicts the amount of learning. As with the Vander Meer experiment, Kanner and Rosenstein included questions where colour seemed important as a pre-requisite to answering correctly. A further analysis was made of trainee performance on these and other non-colour items. The results, again, showed that there is no difference in performance.

The authors state in their closing discussion that the general finding of no important differences in trainee learning was expected for those lessons where colour played a minor role. They continue to say that what is of interest is that this result prevails where colour on a priori grounds would seem important. In particular, they reason, it would be expected that in those test items where colour seemed critical, for example, the reading of colour coded electrical materials, the colour television group would learn better. The explanation

Figure 3.6: Colour vs B&W

Lesson	No Items	B&W Score	Colour Score	Difference
Foster Seely Discriminator	20	12.56	12.75	0.19
Phase Sensitive Rectifier	17	11.40	11.20	0.20
Complex Wave Analysis	20	15.05	14.54	0.51
Cathode Coupled Multivibrator	20	8.73	9.07	0.34
IF Amplifier	20	16.00	15.62	0.38
RF Amplifier	12	8.88	8.79	0.09
Automatic Range Tracking	15	5.80	7.50	1.70*
Signal Generator	18	13.07	13.18	0.11
Introduction to Colour	16	8.54	7.21	1.33
Material for Copying	15	11.28	10.95	0.33
Film Characteristics	15	6.97	7.44	0.47
Overall means	**188**	**118.28**	**118.25**	**0.03**

* significant at 0.05 level

offered by the investigators is that where colour seemed important the names of the colours were inserted or referred to by the instructor. When reading colour coded resistors monochrome viewers would hear that the colour red had a certain value but only see it as a shade of grey. It seems reasonable to assume that for this type of learning, hearing the association of a colour and a particular value is adequate for later seeing this colour and remembering its value. In other words, for most people colours are sufficiently familiar so that words can be substituted for them in many learning situations and transfer can occur when colour is used in the performance situation.

Stephen Cox produced a survey of the research into the effects of colour in learning from film and television in 1976 and from the results of twenty or so studies concludes that overall there is no marked difference in learning from colour or black and white film or television.

However he further points out that this does not mean that colour has no effect, but rather that our measuring instruments are too insensitive for measuring such effects and that it is probable that the use of purely objective tests for recall of selected information has masked the true effects of colour on the learner. According to Cox, it would seem that any such effects are not to be found in the learning of central information from a programme, but probably work at other levels where measurement is difficult or insensitive.

What Cox is saying does make sense in the context of a micro analysis of an individual's cognitive and affective responses to stimuli, however, I believe that the effects concerning the learning of central information from a programme are of major concern and I agree with Kanner's 1968 conclusion that because colour tv is more expensive and does not appear to provide any learning advantage its use should be restricted principally to medical facilities. He did, however, have the foresight to suggest that this conclusion may be swept aside if it becomes apparent that monochrome television equipment will be phased out by manufacturers. This is rapidly becoming the case, especially in the educational sphere and it is probable that all tv equipment is now colour. This means that despite the fact that no advantages accrue from learning from colour materials, educational establishments must bow to commercial factors and pay the increased costs associated with colour materials.

Before looking in greater depth at the reasons for the lack of differences I should point out how these research results have influenced the development of modern computing and how they are still relevant to decisions that have to be made when looking at computer systems.

It is fair to say that the face of desk top computers changed dramatically in 1984 with the introduction of the Apple Macintosh computer. It was an

extraordinary machine for a number of reasons. In particular was the astonishing idea that colour was to be eschewed in favour of a high definition monochrome screen. In part this decision was taken because the tasks for which the machine was designed — the handling of WISIWIG text —did not required colour, and in part because the cost of adding colour to the machine would have significantly increased its cost. Colour did eventually arrive, and the cost of the machines was very high. The colour decision for education is still relevant when considering portable computers: monochrome laptops can be 50% of the cost of colour versions. Most of this book has been prepared on such a machine — there will be no colour in it, so how could I justify the additional cost? Now, games, that's a different matter!

Dwyer took this research forward in an attempt to improve visualised instruction. He planned and carried out his programme of systematic evaluation of the effects of a variety of pictorial types over an extended period of time, involving 100 separate studies with a total population of 23,000 students.

The first piece of research in this programme is recorded in Dwyer's 1967 research report: Adapting Visual Illustrations for Effective Learning, published in the Harvard Educational Review. He describes an experiment which compares four different audio-visual presentations of the same material. The commentary is the same for each presentation but the pictorial image is different. If you look at the figures below you will have some idea of the differences.

Initially, only black and white illustrations were considered. The four conditions used in the experiment were:

1 Oral/verbal presentation. The students in this group saw no accompanying illustrations, but the name of the parts of the heart mentioned were projected on the screen.
2 Abstract linear representations (Fig 3.7a — later called simple line drawing, black and white). The form and relative locations of the parts of the heart were shown as they were mentioned in the oral presentation.
   The line drawings used in this presentation were similar to instructional drawings used in many current science books.
3 Detailed shaded drawing presentation (Fig 3.7b). Here students viewed detailed shaded drawings to complement their instruction The illustrations were more complete than the simple line drawings, and they represent the heart and its related parts more realistically rather than merely identifying and locating them as in the simple line drawings.

Figure 3.7a

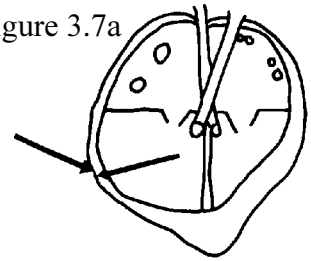
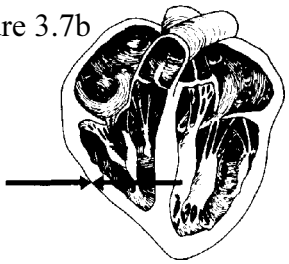


Figure 3.7b



4 Realistic photographic presentation. Students receiving this treatment viewed realistic photographs of the heart.

From the variations in the pictorial materials it is evident that Dwyer is concerned to about the relative effectiveness of illustrations with increasingly more detail. The important question is: do students learn more if illustrations are more realistic i.e. do the available cues add up to produce a more effective visual aid? This is sometimes called the 'Sign Similarity Hypothesis', suggested by C.R. Carpenter, as follows: the sign similarity hypothesis states that a film (or other visual aid) with signals, signs and symbols having a high degree of similarity or iconicity to those of real life will be more effective than films whose signs etc. have a low degree of similarity.

TEST MATERIALS

Hartman (1961) acknowledged the work of Carpenter in his discussion of the 'Stimulus Generalisation Hypothesis' which suggested that: learning of related and unrelated information simultaneously presented by multiple channels increases as the situation in which the learning of the information is tested becomes more similar to the situation in which the learning was presented. Dwyer was aware of Hartman's work and was also familiar with Lefkowith's research, which suggested that audio-visual research must pay particular attention to the testing situation and the test materials.

Lefkowith (1955), on the basis of stimulus generalisation theory, suggested the following hypothesis, in connection with instructional film research:

> As the iconicity level of the pictorial material stimuli in the testing method becomes more similar to the iconicity level of the pictorial stimuli in the teaching material, the recall scores will become higher.

> Note: *iconicity may be defined as the degree to which pictorial signs or symbols are similar to the object signified (Morris, 1946).*

Lefkowith argued 'that it was not too uncommon to find someone evaluating the relative effectiveness of a filmstrip and/or a motion picture by a completely verbal test. Further it seemed to him: 'both unrealistic and invalid to assess the relative contribution of these two visually oriented teaching methods by a completely non-visual test.'

In order to test his hypothesis Lefkowith designed an experiment in which subjects were required to identify nomenclature and functions of a projectile hoist. This subject matter was taught by a recorded lecture illustrated by slides. There were two types of pictorial stimuli used in producing the slides: (1) black and white photographs, (2) out-line drawings made by tracing the photographs. Therefore the two teaching methods, which were to be compared in effectiveness, differed only with respect to iconicity, version (2) being less iconic.

There were three highly reliable forms of test, differing only with respect to the visual cues that were provided, because they all utilised the same set of items:

1 The actual equipment test, which served at the same time as a criterion measure.
2 The test using photographic slides.
3 The test using outline drawing slides.

The subjects used were naval recruits. As a result of the data collected during the experiment the hypothesis given above could be retained: 'There existed a significant interaction between teaching methods and testing methods.' According to stimulus generalisation theory, this point is of great methodological importance:

> If the experimenter had used the photograph test, he would have concluded that the photograph teaching method was more effective. But if he had used the outline drawing test, he would have concluded that the out-line drawing teaching method was more effective. (Lefkowith, 1955)

Following on from this work, Dwyer designed four individual criterial measures, which were administered in the following order: drawing, identifi-

cation, terminology, comprehension. After the presentation of the instructional materials each student was permitted to take as much time as he required to complete one criterial measure before proceeding to the next.

Drawing Test: The drawing test (18 items) provided the students with a list of specific terms corresponding to the parts of the heart discussed in the instructional unit. The students were required to draw a representative diagram of the heart and place numbers of the listed parts in their respective positions.
For this test the emphasis on the positioning of the verbal symbols with respect to their concrete referents, not on the quality of the drawing. This test evaluated student learning of specific locations of the parts of the heart. Knowledge of positions, recall of patterns, structures, or settings of parts within an entity could be included in a test of this nature.

Identification Test: This multiple-choice test (20 items) required students to identify the numbered parts on a detailed drawing of a heart. Each part of the heart that had been discussed in the instruction was numbered on the drawing and appeared in a list on the answer sheet. Students were required to look at the detailed drawing and select the appropriate identifying response in each multiple-choice item. The objective of this test was to measure the student's ability to identify parts of the heart from information received in the instruction.

Terminology Test: Initially, this test (20 items) consisted of a series of fill-in questions in which one answer only was correct. For later studies it was changed into a multiple-choice format to facilitate scoring. This test evaluated student knowledge of referents for specific symbols.

Comprehension Test: The comprehension test consisted of 20 multiple-choice items. Given the location of certain parts of the heart at a particular moment of its functioning, the student was asked to locate the position of other specified parts of the heart. This test required that the student have a thorough understanding of the heart, its parts, its internal functioning, and the simultaneous processes occurring during the systolic and diastolic phases.

> The comprehension test was designed to measure a type of learning that occurs when the individual understands what is being communicated and can use the information to explain some other phenomenon.
>
> **Total Criterial Test:** The items contained in the four individual criterial tests were combined into a 78-item total criterial test. The purpose of this evaluation was to measure the student's total understanding of the information.

It should be noted that in this investigation the type of learning measured by the different tests is the important aspect. The learning objectives have generalizability in that they may be appropriate in all content areas. For example, (a) the objectives measured by the drawing test could be applied to students in a geography class who are learning the location of the fifty US states, or to students in a mechanics course learning the relative locations of the various parts of a carburettor; (b) tests similar to the identification test could be used in any course in which the student is required to know the various parts of objects (for example, in a physiology course to indicate the parts of the human body, or in a botany course to name the parts of different types of plants); (c) the objectives of the terminology test — to evaluate student knowledge of specific verbal symbols — is appropriate to any content area; and (d) tests similar to the comprehension test can be used in any subject area in which the objective is to measure the student's understanding of complex procedures and processes.

For the original experiment Dwyer used 108 students for the four conditions (Oral; Simple line; Detailed shaded; Realistic Photograph) and he concluded that:

> The results of this study indicated that when students viewed their respective instructional presentations for equal amounts of time, the simple line drawing presentation was significantly more effective in facilitating achievement than was the oral presentation without visuals on the drawing, identification, and total criterial tests. The oral presentation without visuals of the heart was found to be as effective as each of the visually complemented treatments on both the terminology and comprehension tests.

Dwyer also concluded that, contrary to previously stated theories of visual communication (eg. Carpenter, 1953), the more realistic illustrations were found to be the least effective in complementing oral instruction.

Later versions of the experimental study utilised five different presentations monochrome presentations and colour illustrations were also introduced at a later stage in the series of experiments.

Figure 3.8: Presentation Most Effective in Facilitating Achievement on Each Criterial Measure.

Criterial Measures	Instructional Treatments
Drawing Test	Simple Line Drawing
Identification Test	Simple Line Drawing
Terminology Test	Oral Presentation
Comprehension Test	Oral Presentation
Total Criterial Test	Simple Line Drawing

The purpose of this and other experiments was to test the hypothesis that an increase in realistic detail in visual illustrations increases the probability that learning will occur. Eventually nine complete slide sequences, possessing differing degrees of realistic detail, were produced so that variations in visual stimuli could be assessed in terms of their ability to facility student achievement on five criterion measures. In a further experiment two hundred and seventy-six ninth grade students were randomly assigned by class to one of the nine treatment groups. Each student received a pretest, participated in his respective instructional presentation, and received four individual criterion measures. The results seemed to indicate that increasing the amounts of realistic detail in visual illustrations does not necessary lead to greater learning. Apparently the increase of realistic detail presents excessive stimuli that interfere with the transmission of the intended information thereby reducing the effectiveness of the more realistic visuals as learning media.

Dwyer went on to conclude that:

1. The use of visuals to complement oral instruction is an effective way to improve student achievement of specific learning objectives; however, it must also he remembered that for some learning objectives oral instruction alone is as effective as visually complemented

instruction.

2. The same types of visuals are not equally effective at different grade levels in facilitating student achievement of identical educational objectives.
3. For specific objectives and for students in certain grade levels, colour appears to be an important instructional variable for improving student achievement.
4. In general, the differential effects attributed to the different visualised treatments on the immediate retention tests disappeared on the delayed retention tests.
5. The realism continuum for visual illustrations is not a reliable predictor of learning efficiency for high school students. The results of these studies substantiate the hypothesis that reality may be edited for instructional purposes.

REFERENCES

Brown, H.E. (1928) Motion picture or film-slide. Science and Maths, 28, 517-526.

Carpenter, C.R. (1953) 'A theoretical orientation for instructional film research', Audio Visual Communication Review, 1, 38-52.

Carson, D. (1948) The American way of life as portrayed in filmstrips - an experiment in visual education. Research Publication no.2; Scottish Educational Film Association: Glasgow.

Cox, S. (1976) 'Colour in learning from film and tv: a survey of the research with some indications for future investigations', Educational Broadcasting International, June.

Dwyer, F.M. (1967) 'Adapting visual illustrations for effective learning', Harvard Educational Review, 37, 250-263.

Freeman, F.N. (1924) Visual Education. Chicago: University of Chicago Press.

Halbert, M.G. (1943) An Experimental Study of Children's Understanding of Instructional Materials. Kentucky: University of Kentucky.

Hartman, F.R. (1961) 'Recognition learning under multiple channel presentation and testing conditions', Audio Visual Communication Review, 9, 24-43.

Houser, R.L., Houser, E.J. and Van Mondfrans, A.P. (1970) 'Learning a motion and a non-motion concept by motion picture versus slide presentation', Audio Visual Communication Review, 18(4), 425-430.

Kanner, J.H. (1968) The instructional effectiveness of colour television: a review of the evidence. Stanford University: Stanford, California. ed 015 675.

Kanner, J.H. and Rosenstein, A.J. (1960) 'Television in army training: colour versus black and white', Audio Visual Communication Review, 8, 243-252.

Laner, S. (1954) 'The impact of visual aids showing a manipulative task', The Quarterly Journal of Experimental Psychology, 6(3), 95-106.

Lefkowith, G. (1955) The validity of pictorial tests and their interaction with audio visual teaching methods. Technical report no. sdc 269-7-49 Instructional Film Program. Naval Training Devices Centre-Office of Naval Research:Port Washington.

May, M.A. and Lumsdaine, A.A. (1958) Learning from films. Yale University Press: New Haven.

McClusky, F.D. (1924) 'Comparison of different methods of visual instruction.' In F.N. Freeman (ed) Visual education. Chicago: University of Chicago Press, 83-166.

Mellinger, B.E. (1932) Children's Interest in Pictures. New York: Bureau of Publications, Teachers College, Columbia University.

Miller, W.A. (1938) 'Reading with and without pictures', Elementary School Journal, 38, 676-682.

Morris, C.W. (1946) Signs language and behaviour. New York: Prentice Hall.

Roshal, S.M. (1949) Effects of learner representation in film mediated perceptual-motor learning. Instructional Film Program. Naval Training Devices Centre. Office Of Naval Research: Port Washington-New York..

Rudisill, M. (1952) 'Children's preferences for colour versus other qualities in illustrations', Elementary School Journal, 52, 444-451.

Silverman, R.E. (1958) The comparative effectiveness of animated and static transparencies. Technical report no. navtradevcen 78-1. Naval Training Devices Centre. Office of Naval Research:Port Washington-New York.

Spangenberg, R.W. (1973) 'The motion variable in procedural learning', Audio Visual communication Review, 21, 419-435.

Spaulding, S. (1955) 'Research on pictorial illustration', Audio Visual Communication Review, 3, 35-45.

Twyford, L.C., Davis, J.S. and Seitz, C.P. (1954) Evaluation of the film: military police support in emergencies (Riot Control). Technical Report no. sdc 269-7-52. Instructional Film Program.. Naval Training Devices Centre; Office of Naval Research:Port Washington-New York.

Van Mondfrans, A.P and Houser, R.L. (1970) 'Selecting media to present basic concepts', Educational Technology, 11(4), 40-43.

Vander Meer, A.W. (1954) 'Colour vs black and white in instructional films', Audio Visual Communication Review, 2, 121-134.

Wells, R.F., Van Mondfrans, A.P., Postlethwait, S.N. and Butler, D.C (1973) 'Effectiveness of three visual media and two study formats in teaching concepts involving time space and motion', Audio Visual Communication Review, 21(2), 233-241.