TELIDON

TELIDON BEHAVIOURAL RESEARCH

6 Content Assessment
A measurement model for videotex

IC

Government of Canada
Department of Communications

Gouvernement du Canada
Ministère des Communications

Cat. no. Co22-55/1985

ISBN 0-662-53677-0

Rec'd free from Behavioural Research on 25/5/85

CONTENT ASSESSMENT

A MEASUREMENT MODEL FOR VIDEOTEX

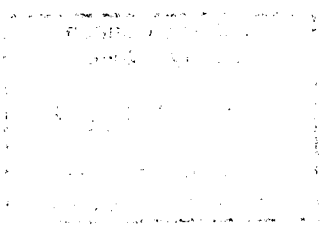
D. Elaine Pressman

Irwin S. Pressman

This project was funded by the Department of Communications
Government of Canada
(No. 36100-3-0353)

Copies of this document are available from:

Information Services
Department of Communications
300 Slater Street
Ottawa, Ontario
K1A 0C8



PREFACE

Content analysis is a technique that has been used to describe the characteristics of videotex pages and programs. In order to be of use to the videotex industry, analysis must go further. It must provide benchmarks by which content developed for public use can be evaluated, both during and after its creation. This implies a blending of content analysis with content assessment.

The videotex industry is still in its infancy. The full potential of this medium has yet to be exploited. Innovative uses of animation for example, are only beginning to emerge in programs. Other new developments may help to revolutionize the type of content developed for this medium.

The best content for videotex will be that which can attract and hold the largest audience for the industry. Such content would be of greatest utility to advertisers and/or other service underwriters. Although it is clear that good page creation requires creativity and a strong visual sense, an objective system of evaluation using available knowledge would be desirable to monitor the product.

Content is being developed today. In Canada alone, 27 content development projects eligible for a total of five million dollars in assistance are being supported in 1984 by the federal government. The Canadian and American private sector is also active in videotex program development. These projects promise to create sophisticated and innovative materials for videotex systems. A comprehensive method for monitoring the merit of this content has been lacking.

The content of a videotex page is the total information received by the viewer. This information is both textual and graphic. It depends upon the sizes and shapes of the letters and their relation, placement and density. It also depends on the colours used, the context in which the message occurs and the capacity of the recipient to decode, interpret and comprehend the message. The ease with which a reader understands the text message is related to the syntactic and semantic complexity of the sentences.

By content assessment we shall mean the evaluation of these features of the message. These aspects of the programs or pages can be defined and monitored to provide specific feedback for the industry or government regarding the value of the content created. The components identified in this document do not represent an exhaustive categorization, but they do reflect major aspects of videotex pages that have been identified as relevant for developing content. They include: linguistic complexity of the text (readability), utilization of graphics, typographical design features, and utilization of colour. The content is viewed as the sum of these features per page or per program.

A thematic description of videotex, whether industry-wide or of a specific program, is a form of macrocontent analysis. An industry-wide analysis could determine the trends of content, but would be beyond the scope of most content analysis relevant to information providers.

The model described in this report and the rating scheme adopted represents a first approximation to the solution of a difficult problem. The criteria proposed are based on documented research that suggests factors to facilitate ease of understanding, retention and recall. Factors related to user appeal are also included.

It is hoped that the approach taken will stimulate discussion and investigations that will lead to refinements of the content analysis system proposed here.

Thanks are due to Dorothy Phillips, Director of Behavioural Research and Evaluation, Department of Communications for her thoughtful support, and to Sjef Frenken, Director, Policy Development, Broadcast Programs Directorate, Canadian Radio-television and Telecommunications Commission for his helpful comments.

Ottawa, May 1984

ABSTRACT

Content analysis is a technique that has been used to describe and/or sort characteristics of the content of various types of media. Videotex has specific characteristics that distinguish it from other media. The purpose of this document is to present an approach to content analysis that would be specifically relevant to videotex and would be useful for the evaluation of videotex pages and programs.

An introduction to a proposed model of assessment for videotex is included in the first chapter. The following four chapters address the analysis of four major areas: graphics; colour; linguistic features, in particular, readability and human interest measures; and visual display features. These four aspects were identified as particularly relevant for videotex content analysis.

Each of these category areas include subcomponents. These are separately identified, defined according to delineated criteria for a rating scale, and then combined to give a general subscore for the category. An overview of the scoring system and the rationale for the technique is provided in chapter 6. A discussion of the strengths and weaknesses of this approach, as well as the possible interpretation of results obtained, are also included in chapter 6. Scoring details and an example of a profile sheet used in the analysis are included in chapters 2 through 6. Concluding remarks are offered in chapter 7.

The approach described in this document is an attempt to provide a methodology for the analysis of observations of videotex pages. Through a delineation of categories and factors, rated according to defined criteria, it is anticipated that the reliability of the analysis will be enhanced.

The score and subscores that result from the proposed scheme of analysis are considered a quantitative description of content. Further studies will have to be undertaken to relate the findings of the methodology to the results of market studies. It is hypothesized that a more precise definition of pages produced on videotex will assist the identification of the most important characteristics of content in the eyes of the consumer.

CONTENTS

1 INTRODUCTION TO THE PROPOSED MODEL FOR CONTENT ASSESSMENT

In order to provide an evaluation of the content of videotex, a model had to be developed to indicate which characteristics of content should be selected for measurement. An examination of relevant literature indicated that while a number of special purpose tools and techniques have been developed over the years for measurement of aspects of content, there was no single, widely recognized or universally accepted methodology for performing an assessment of content.

The assessment of content in the proposed model is broken down into a number of subtests, each of which concentrates on a particular aspect of the analysis. These aspects include: graphics, use of colour, text features, and linguistic features. At this stage of development of the model, ratings will be obtained for each subsection according to prescribed criteria that are based on information generated from documented studies in the relevant literature. Further research will be required to establish the relationship between the ratings based on reported studies and audience perception of videotex content. The goal of this document is to suggest a possible first step towards a model for content assessment of videotex.

Videotex constitutes a special environment for the assessment of content, as some measurements can be done quite efficiently due to computer control of the pages. The content assessment measure proposed here consists of a number of structured observations according to defined criteria, some of which have been automated.

Need for a content assessment scheme

Commercial videotex page producers have lamented the fact that videotex content produced by the industry is often ineffective. Where appropriate pages have been produced, they often have not reached the audience that should see them. Some content has been too specific, and some not cost effective.

To attract consumers, the full power of videotex creativity is essential. Creative graphics in combination with properly displayed text, appropriate colour usage, and messages of the right linguistic complexity will significantly enhance the success of the technology.

Content analysis methods can be used to specifically measure these areas and to evaluate them against criteria. Each of these content categories is composed of factors that can be identified and rated to provide some quantitative measure of videotex content. The criteria against which these measurements are judged have been selected from previous studies of content. This type of analysis will provide information about the nature of the videotex content that is being produced today. It provides page creators with a self-evaluation system that should improve their product. Assessment/evaluation of the components of the content coupled with feedback of program usage may provide valuable insight into the components of content that are the major contributors to consumer satisfaction.

Applications of content assessment to the videotex industry

Self-monitoring of programs developed for videotex by creators is likely to result in a superior product. Purchasers and advertisers are interested in a videotex format that will "sell". The content design must satisfy the needs of the advertiser. This implies not only capturing the audience but creating the desired response to the message. Careful analysis of the messages may lead, over time, to a more explicit delineation of relevant factors that produce the desired effect in the viewer.

Content assessment of videotex can be used to address the following types of questions:

- Can the content be understood by the general population or an identified subset of the population?
- Are the graphics integrated with the text?
- Are display techniques used that maximize ease of reading and recall?
- Is the content attractive?
- Is the content weighted in terms of text or graphics?
- Are animation or other attention-getting aspects of videotex employed?
- Are graphics used decoratively and/or with communicative intent?

Marketing studies and research into the relevance of these aspects to audience receptiveness and to the acceptability of videotex is needed. There have been studies of print and other broadcasting media, as well as teletext that suggested methods of content analysis of the specific features of pages. These research results were used to provide the prescriptive criteria for the rating scale of the content assessment.

Four content areas were identified for evaluation in the assessment protocol. These particular areas were chosen because of their relevance to content analysis and the existence of factors within each area that could be objectively measured.

Elements of the model

The model for content assessment consists of four content areas: (1) readability, (2) graphics, (3) visual display features, and (4) use of colour. Each of these areas contains several variables that can be measured and assessed against a criterion based on previous research.

The variables introduced will take on the values 0 and 1 only, and will not take intermediate values. They will correspond to the presence or absence (1 or 0 respectively) of an attribute, or the attainment of a predetermined level of merit. For example, if the Flesch Reading Ease scale (Flesch, 1948) indicates that an identified text is readable for persons with a grade 8 education, then this variable would be scored as 1 if it was within the accepted criterion level. Otherwise, it would be scored as 0.

The following chapters discuss each of the content areas in turn describing the specific variables to be measured, the criteria for assessing each variable, and the methods for arriving at a total score for each content area.

2 READABILITY FOR CONTENT ASSESSMENT OF VIDEOTEX

There seems little question that highly readable, clear and concise writing will attract people to videotex. There are different approaches to the measurement of readability. A historical chronicle of these has been undertaken by Klare (1963).

Readability formulas have the limitation of measuring only a small number of variables related to the level of reading difficulty of text. The indexes represent statistical approximations to the linguistic complexity of text. Although these measures have been criticized because they neither analyze the ideas used, nor measure the coherency or organization of the document, some of these indexes have been found to work well in practice.

The formulas do not measure the level of reading difficulty precisely. They do give an indication of grade level required to comprehend a text. This reading difficulty level is only one measure, but it is often important. This is shown by the relationship of readability to reading speed, acceptability, understanding, and learning (Klare, 1975).

As readability has been found to be both important and easily quantified, it is reasonable to include it in a content assessment scheme. Its measurement has been automated in a number of instances. The application of several distinct measures of readability to a specified text is likely to enhance the reliability of the findings.

Flesch Reading Ease Score

One of the easiest and most popular measures of readability is the Flesch Reading Ease Formula (1948). This formula uses the following steps:

- Systematically select three 100-word samples from the material to be rated. If the material is less than 300 words, then collect fewer samples but these samples should be representative. Contractions and hyphenated words count as one word; words are defined as numbers or letters separated by spaces.

- Determine the total syllables in each of the 100 word samples (WL).
- Determine the average number of words per sentence (SL).
- Apply the following reading ease equation --
  RE = 206.835 - .846 WL - 1.015 SL.

The result will be a number between 0 and 100, with maximum difficulty being indicated by the lower figure. A score of 30-50 would be rated as very difficult and would be typically found in a scientific magazine or journal. A score of 50-60 represents a fairly difficult reading level such as that found in a "quality" magazine. A score in the range of 60-70 was defined as "standard" as found in many popular digests. A rating of 70-80 would be considered "fairly easy" and 80-90 would be "easy". These scores would be characteristic of popular fiction. The category of "very easy" reading level would be a rating of 90-100 and is found in publications such as comic books (see Appendix).

The Flesch Reading Ease Score is available in a computerized format using the UNIX operating system in a program called "The Writer's Workbench" (MacDonald, et al, 1982). In addition to readability analyses, the Writer's Workbench can analyze the surface features of text, such as sorting the parts of speech in English text.

Human Interest Score

It has been found that readers attend more to articles about other people than abstract events. The use of eyewitness reports in newspapers attests to this fact. As a result, the use of personal references or personal pronouns has been related to formulas about readability.

"Personal words" have been defined by Flesch (1948) as follows: all nouns with natural gender; all pronouns except neuter pronouns; and the words "people" (used with the plural verb) and "folks".

Flesch also defined his meaning of the "personal sentences" element that was identified to test the conversational quality and story interest of the passage analyzed. Personal sentences

include the following: spoken sentences, marked by quotation marks or otherwise; questions, commands, requests and other sentences directly addressed to the reader; exclamations; and grammatically incomplete sentences whose meaning has to be inferred from the context.

In order to ascertain the Human Interest Score, the following steps are required (Flesch, 1948):

- Select systematically 100 word samples (as for the Flesch Reading Ease Score).
- Determine the number of "personal words" per 100 words (PW).
- Count the number of personal sentences per 100 sentences (PS).
- Apply the following human interest formula -- HI = 3.635 PW + .314 PS.

The Kincaid Formula and the ARI Formula

The Kincaid Formula (Smith and Kincaid, 1970) is an extension of the Automated Readability Index (ARI) introduced by Smith and Senter (1967). The ARI uses average letters per word as the measure of word difficulty and average words per sentence as the measure of sentence difficulty. The ARI data was originally gathered by having the material typed on an electric typewriter that was slightly modified by the installation of three microswitches attached to cumulative counters. One microswitch was activated each time the ball of the typewriter went forward to give a stroke count. Another switch was activated by use of the space bar (word count). The third microswitch was activated by striking an identified key to give a sentence count. The letters per word (LPW) and words per sentence (WPS) were then computed.

The ARI correlated highly with the Flesch Count (1948). A correlation of .90 was found between the ARI and the Flesch for a 1432 word sample (Kincaid et al, 1967). This formula is also available in the Writer's Workbench. It is as follows:

Reading grade = 4.71 LPW + .5 WPS - 21.43

The Kincaid formula was based on Navy training manuals that ranged in difficulty from 5.5 to 16.3 in reading grade level. It is also based on adult materials rather than school textbooks. As such, it has been proposed as the best formula to apply to technical documents (Cherry and Vesterman, 1981) and is available in the Writer's Workbench.

The Kincaid formula is given by:

Reading grade = 11.8 SPW + 0.39 WPS - 15.59

where SPW denotes the average number of syllables per word.

Summary of readability analyses for videotex

Three measures of readability can be obtained (rather than only one) to enhance the reliability of the identified reading level requirements for text. These include the Flesch Reading Ease Score, the Automated Reliability Index, and the Kincaid Formula. The Flesch Human Interest Index can be obtained to provide a supplementary measure.

The creator of the videotex pages should determine the target population for the program being developed, and identify the required reading skill (easy, very easy, standard, difficult) or the grade level requirement of the reader. In the content assessment scheme, this will become the criterion level.

After each of the formulas is applied to the text material, a score of 0 or 1 will be assigned. A 0 score will be obtained if the reading level obtained from the index used exceeds the chosen criterion level identified. A 1 will be obtained for each readability measure if the level of difficulty of the text is lower than that identified as the maximum desirable level. Scoring details and examples for readability ratings are given in Table 1.

TABLE 1

Scoring details and examples

Readability and human interest subtest

Item	Acceptable score 1	Unacceptable score 0
Flesch R.E. Score	rating of standard style of writing or easier; reading ease score $\leq$ 60	reading ease score > 60
ARI Score	grade level $\leq$ 8	grade level > 8
Kincaid	grade level $\leq$ 8	grade level > 8
Human Interest Score	human interest score; > 30 interesting range or better	human interest score $\leq$ 30

Criterion levels of this section may be modified by the text developer according to the target population. The criterion level used in this table was chosen as the "standard" reading ease level as identified by Flesch (1948).

If the Human Interest Score is relevant to the material being prepared, this index may be applied to the material. Interpretation of this score is also available in Table 1. A score of 1 is obtained if the material is classified as being in the interesting category or better (that is, a score above 30). A score of 30 or less would result in a rating of 0 for this factor.

The Readability Content Score (RCS) is a measure of the appropriateness of the level of reading difficulty (as determined by the creator) and the Human Interest Score (when relevant). The RCS is determined as follows:

$$RCS = \frac{\sum_{i=1}^{4} S_i}{4} \times 100$$

Where: S_1 = Flesch Reading Ease

S_2 = Human Interest Score

S_3 = ARI Score

S_4 = Kincaid Score

One could easily modify the score to give a weighted average. It is the content assessor's responsibility to emphasize the aspects that are most relevant to their particular needs.

3 USE OF GRAPHICS IN VIDEOTEX CONTENT

The function of graphics on a page must be clearly understood by the designer. Graphics can substantially contribute to the conveyance of meaning and the facilitation of recall if used appropriately in videotex. All too often in this medium, graphics have been principally used for a decorative function. These graphics may not contribute to the message and, in fact, may detract from it.

Focus attention with meaningful graphics

Graphics can be used to increase the effectiveness of the communicative intent of the message. Use of graphics in this manner can assist the user in gaining the proper integration of the information being communicated. One suggested way to focus attention with graphics is to use references to the graphics in the body of the text such as a notation (see Figure x) at the appropriate location. This technique could assist with the integration of the graphics with text. This would be relevant when detailed informative text pages are required and where graphics could clarify or enhance the information to be conveyed.

If the graphics used are meaningful and appropriately integrated into the text by page design or by direct reference, this item can be scored as 1 on the 0-1 scale. Otherwise, it is scored as 0 in this scoring system. Table 2 describes this variable for scoring.

Use captions to reduce ambiguity

When graphics are used in an illustrative manner, captions should be employed to reduce ambiguity. Personal experiences influence our perceptions and can result in different interpretations of graphics by different individuals. Captions for graphics, which are used in videotex, can reduce the possible number of picture-related erroneous conclusions. Captions have been identified as one of the most important variables to facilitate the correct interpretation of pictures (Gombich, 1972). This effect is likely to be relevant for videotex pages.

If the graphics contain captions that reduce or eliminate ambiguity, this item can be scored 1; otherwise score 0 (Table 2).

Use of realistic drawings with short build-up time

Research aimed at picture preference has shown realistic formats are generally preferred by both adults and children over abstract representations in art (Dwyer, 1972; Myatt and Carter, 1979). In terms of videotex, realistic graphics might be important for consumer satisfaction.

Although detailed drawings enhance realism, much complexity in drawings has been found to be a negative factor and should be avoided (Wohlwill, 1975). Picture complexity, for example, has been found to result in decreased interest and examination time. Excessive complexity in videotex pictures can also create a problem of excessive build-up time. This was particularly problematic in early videotex programs. Examples of this are city skylines in weather reports; tractors in some farm-related videotex; and some designs in restaurant advertisements. The determination of the appropriate level of graphic complexity in videotex is a difficult task for videotex producers.

If graphics used are identifiable drawings and appear within a reasonable build-up time, this variable is scored as 1. If graphics are ambiguous or needlessly detailed, score 0.

Use animation or graphics with a dynamic setting

Research has shown that people look longer and respond more readily to pictures that show dynamic action (Travers and Alvarado, 1970). The animation capability of videotex can capitalize on this factor.

When animation is not used, a graphic representation of a dynamic situation, rather than one of a static situation, has been shown to enhance the receiver response. For example, a person represented as actively chopping wood should generate more response than a picture of an immobile person standing beside a tree with an axe.

If graphics used have movement or are representations of action, score 1, otherwise score 0.

Place graphics on the page to maximize function and use size of graphics to maximize retention

Graphic representations can serve a useful function when placed either before or after a reading passage. When used before a passage, they can provide an orientation to the text as an overview or an introduction. When a relevant graphic representation is used after a text, it may increase retention of the text by providing a review of the message imparted. Small relevant graphics used at the end of the text may assist consumers to remember advertised products.

The relative size of objects being represented may also be important in retention (Kosslyn, 1975). It was reported in this study that more details of a mental image of an object are remembered when the object is imagined next to a small object. As a result, juxtaposition of graphic representations may be important aids to retention.

If the purpose of a graphic on a page is to encourage recall in the user in videotex advertising, juxtaposing the object of interest with a smaller object may facilitate the retention of the advertiser's message.

TABLE 2

Scoring details and examples

Graphics subtest

Item	Acceptable score 1	Acceptable score 0
Focus attention with graphics	use of reference to graphics in the text; meaningful graphics (not solely decorative)	decorative or unrelated graphics
Reduce ambiguity	use of graphics; clear relationship of graphics to content to be conveyed	any possible ambiguity; rationale for use of graphic not evident
Realistic drawings	bright, vivid graphics with limited detail	abstract graphics; very detailed graphic
Animation or dynamic graphics	graphics have movement or are representations of action	no movement and no action in graphic representation
Appropriate size and placement of graphics	juxtaposed graphics of differing sizes; or use of graphic before text to orient; or use of graphic at end of text	no identifiable juxtaposition or placement of graphics for orientation or closure

Summary of graphic use in videotex

In order to obtain a preliminary rating of the graphic content, each sampled videotex page included in the content analysis will be rated 0 or 1 for each of the five factors described. The Graphic Content Score (GCS) is determined as follows:

$$GCS = \frac{\sum_{i=1}^{5} Q_i}{5} \times 100$$

Where: Q_1 = The use of graphics to focus the viewer's attention and enhance the information in the text.

Q_2 = The use of captions to reduce ambiguity of the message.

Q_3 = The use of realistic drawings limited in complexity.

Q4 = The utilization of animation or dynamic representations of graphics.

Q_5 = The utilization of appropriate size and placement of graphic images to enhance retention of the message.

The scoring of these variables is described in Table 2.

In determining the Graphic Content Score, n = the number of components that are scored in the application. In this case, n = 5. Those factors that are not relevant to a specific content analysis should be omitted. The value of n would change accordingly.

Similarly, new factors could be introduced. This is intended to provide one tentative prescriptive approach to the analysis of graphics. Certainly there are other measurements, such as the frequency, size, location, and colours of the graphics that may be introduced in subsequent analyses. This scoring system is meant to be a first step.

4 VISUAL DISPLAY FEATURES ON VIDEOTEX

Human behavioural response to the visual characterics of papertext and teletext has been reported. These findings are relevant to videotex displays and can be used to develop guidelines for page creators.

The following aspects of visual displays can be evaluated for content analysis of videotex: chunking cues, typographical cues, spacing cues, line length, and text justification.

Segmentation and indentation cues

Errors in retrieving information in written text have been substantially reduced when segmentation of text is used (Hartley, Burnhill and Davies, 1979). Indentation and segmentation of text for example, have been found to produce a reading rate improvement ranging from 14 to 18 per cent, and should be relevant to the videotex experience.

Segmentation and indentation display the information in units that are easily encoded by the reader. This has been found to be an important factor in the retrieving and retention of information. The placement of segments of text can also emphasize the relative importance of topics. Care should be taken to ensure that segmentation and placement are appropriately controlled in videotex page creation. See Table 3 for scoring guide.

Typographical cues

Typographical cues such as special fonts, bold face or underlining can be used in text displays. Although emphasis can be achieved through the use of typographical cues, multiple cueing can confuse the reader (Glynn and DiVesta, 1979).

Typographical cueing can result in increased comprehension of text. Comprehension can be enhanced by as much as 16 per cent when important material is double-typed for bold face (Dearborn, et al, 1951; Frase and Schwartz, 1979).

Ease of reading using capital letters as compared to lower case letters has recently been studied. Hartley (1978) found that long strings of capital letters are less readable then lower case letters.

Spacing cues

Printed text that includes generous spacing has been rated as easier and more interesting than text that has a more solid appearance (Smith and McCombs, 1971). Videotex pages with well-spaced text are similarly more likely to provide an attractive display for the consumer.

Spacing should be carefully planned, and each videotex page with text should be self-contained. The end of text on a page should correspond to a natural and logical break in the information. This constrains the type of information that is likely to be most appropriate for videotex and suggests that the information conveyed should be condensed.

It has been suggested (Reynolds, 1982) that headings be placed at the left of the screen. Centred headings may disrupt the normal pattern of reading. Studies that investigated optimal letter size and spacing have been undertaken (Treurniet, 1980).

Reynolds (1982) considered the optimal number of lines in a paragraph on a screen. This work indicated that paragraphs should be between three and six lines in length, separated by a line space. This space is to prevent the text from forming a solid block on the screen.

Line length

The optimal line length has been investigated by Frase, et al (1981). It has been reported that the optimal line length changes as text increases in difficulty. The results of the studies in this area suggest that the "best" line lengths increase from 44 to 50 to 56 characters as reading difficulty moves from easy (Grade 4) to medium difficulty (Grade 10) to very difficult (Grade 18) respectively.

If the page creator intends to produce text that has a readability value related to the lower grade level, a relatively short line length might well be considered to complement the easy reading level. At present, there are 40 characters per line in the current Canadian videotex system. This is consistent with producing text for easy reading levels.

Text justification

Right justification of lines by increasing spaces between words by hyphenation, or by changing the spaces between letters in a uniform fashion, has been found to have little effect on reading speed or comprehension as compared to unjustified text for average readers. For less able readers, unjustified text has been found to be more advantageous (Hartley and Mill, 1973).

If a low grade level of readablity is required for text produced for videotex, unjustified text rather than justified text should be considered.

Summary of text display features

In order to produce attractive and easily readable videotex pages, the page creator must recognize the strengths and limitations of the medium. Efficient and effective videotex page displays can result in the following: increased speed of reading, increased understanding by the reader, improved focus on the main message of the page, facilitation of retention and review, and increased ease and enjoyment of reading.

TABLE 3

Scoring details and examples

Visual display features

Item	Acceptable score 1	Unacceptable score 0
Segmentation and indentation cues	segmentation used to chunk information and facilitate	no segmentation to facilitate encoding reading
Typographical cues	use of special fonts, bold face or underlining to increase comprehension of text	no bold face or other typographical cues used to improve comprehension of text
Spacing	space within text and at end of text corresponds to logical break in information	violation of either of these features in the text
Line length and text justification	less that 49 characters per line; and no right justification	49 or more characters per line, or right justification

The factors associated with the visual display analysis as described earlier can be rated on a 0 - 1 dichotomous scale according to the acceptability criteria defined in Table 3. The Visual Display Content Score (VDCS) is calculated as follows:

$$VDCS = \frac{\sum_{i=1}^{4} V_i}{4} \times 100$$

Where: V_1 = chunking cues

V_2 = typographical cues

V_3 = spacing features

V_4 = line length and text justification

As other visual display features become identified as important analysis features, the n will increase from the 4 characteristics identified here.

5 THE USE OF COLOUR IN VIDEOTEX CONTENT

Colours differ considerably in their physical and psychological characteristics. These differences can affect legibility on the videotex screen and the visual acuity of the user. The colours provide important information to the viewer, both conscious and unconscious.

Characters are easier to identify on a videotex screen if they contrast with the background. Numerous metrics have been developed for expressing colour difference. Recent studies (Carter and Carter, 1981; Sallio and de Legge, 1982) have suggested that colour use in visual displays, such as videotex should be evaluated.

Colour use in text for legibility, conspicuousness and visual acuity

Legibility and conspicuousness of foreground characters have been found to increase as contrast values increase (Carter and Carter, 1981). A recent study investigating the utilization of contrasting colours on a videotex screen has suggested that particular colour contrast combinations aid legibility. Although it is evident that light characters on a light background are not ideal for legibility, the results indicated that dark characters on a light background were often more effective for legibility than the opposite (Sallio and de Legge, 1982). A few exceptions included cyan characters on a blue background or yellow characters on a black background. The only exception to the contradication of using light characters on a light background is green text on a white background.

It has been suggested that when large quantitites of information are displayed on a dark screen, green is likely to be one of the most legible colours. Green has adequate luminance compared to white (59%), and it is one of the colours to which the eye is most sensitive and for which acuity is the greatest (Ton, 1969). This colour is also produced by a single electron gun.

Other colours that are likely to be highly legible on a dark screen are white (100% luminance), yellow (89% luminance), and cyan (70% luminance). Although red and blue have low luminance values (30% and 11% luminance respectively) and the eye is also

relatively insensitive to these colours (Barmack and Sinaiko, 1966; Vanderkolk and Herman, 1975), the use of red or blue text on a light screen has produced text classified as desirable to read (Sallio and de Legge, 1982) on videotex.

Different wave lengths are refracted by different amounts as they pass through the eye. This means that only one colour at a time can be in focus on the retina. If yellow or white are in focus, then blue will be focused in front of the retina and red behind it. This results in blurred images for out of focus colours and causes differently coloured images to appear to be in different planes. As a result, when red and blue are used together in close proximity on a videotex page, the effect may be visually disturbing, and potentially fatiguing as the eye focuses first on one and then the other wave length.

Visual acuity, or the ability of the eye to resolve fine detail, varies with the wave length of light. Visual acuity is highest for white and green light. Our resolving power for blue is one-fifth of that for white, and for red it is only about one-third of that for white (Graham, 1965).

Although some recent evidence suggests that dark characters on a light background may be more effective for legibility than the reverse, a strong colour contrast between foreground and background is essential. Colour groupings for videotex into dark and light (luminance dependent) appear in Table 5. A score of 1 is obtained if colour contrast between characters and background is maintained to ensure adequate conspicuousness and legibility. If colour contrast is not maintained, score 0.

Use of colour for emphasis and retention

Colour can be used in videotex for emphasis in the same way that typographic variations are used in print. It may be assumed that the most visually dominant colours are the brightest colours. Therefore, on a dark background, white or yellow might be used for headings with cyan or green for text, in order to ensure that the headings are bold.

As this use of colour is related to enhanced layout, emphasis by colour can be expected to produce better organization and structure to improve the comprehension of the text in accordance with the principles of structured writing (Horn, 1982).

Use of foreground and background colours may also affect the memorization and recall of characters on the videotex screen. Recall of messages on videotex was facilitated by red characters on a cyan background, blue characters on a yellow background and black characters on a magenta or cyan background (Sallio and de Legge, 1982). Although such results require further substantiation before being included in the rating scheme, the optimum use of colour to facilitate retention is likely to become an important factor in videotex advertising.

Appropriate distribution and subjective utilization of colour

The distribution of colours should be related to the structures of the information. The different luminances of colours on the videotex page may appear to make some items of information more important than others, may imply hierarchical relationships where none exist, or, where relationships do exist, they may be misrepresented.

The use of colour may also assist in conveying the message on the videotex screen to the viewer. Certain colour combinations have been identified for conveying informative messages (rules for games, legal information, press releases) and other colour combinations have been identified when the message is meant to project positive feelings (love letter, child's message, a vacation in the sun) or negative feelings (bad news). Suggested examples of the use of colour to provide agreement with the message is as follows: good news (blue text on green background, cyan on white, or white on blue); bad news (red, magenta, cyan or yellow characters on a black background). These results, although preliminary, suggest interesting directions for future analysis.

Illogical use of colour can impair the clarity of the message. Care should be exercised to avoid such misapplication of colour.

TABLE 4

Scoring details and examples

Colour subtest

Item	Acceptable score 0	Unacceptable score 0
Legibility	sufficient use of colour contrast in foreground and background (see Table 5 for details)	sufficient distinction of foreground and background (see Table 5 for details)
Emphasis	use of colour to emphasize important content on page	no identifiable use of colour to provide emphasis
Distribution of colour and subjective utilization	use of colour distribution to clarify organization of content or hier-archical relationships, and to establish agreement with message	no use of colour to clarify organization of content or to be consistent with content itself
Colour variety	appropriate colour variety with no more than four main colours for text in foreground	jarring use of 5 or more competing colours for text in foreground

TABLE 5

Colour contrast use on videotex

according to luminance values

Black Blue Red Magenta	dark colour range
Green Cyan Yellow White	light colour range

Interpretation

Use of colour contrast between foreground characters and background are expected to enhance the conspicuousness and legibility of the message.

Colours in the luminance range of black, blue, red and magenta are classified in the dark range (low luminance) and colours in the luminance range of green, cyan, yellow and white are considered in the light range (high luminance).

Scoring

For purposes of scoring, if both target and background colours have luminance values in either the dark or light range score 0. The one exception will be the use of green characters on a white background (score 1). If foreground and background colours are from different ranges score 1.

Colour variety in text

The use of more than three or four colours on a CRT has been shown to fragment rather than assist in the integration of information (Cahill and Carter, 1976). This suggests that for videotex applications, it may be advisable to limit the number of competing colours of foreground characters in text on the screen to a maximum of four.

Additional colours may be used creatively in the graphics on the page, but care should be exercised to ensure that the integration of the page is not lost through inappropriate and illogical colour usage.

Summary of colour content assessment for videotex

Each of the colour factors in the analysis per page will be given a 0-1 rating. The definitions of an acceptable (1) or unacceptable (0) rating for each variable are found in Table 4 and Table 5. The Colour Content Analysis Score (CCS) will be determined as follows:

$$CCS = \frac{\sum_{i=1}^{4} R_i}{4} \times 100$$

Where: R_1 = The use of colours to maximize legibility and visual acuity on videotex pages.

R_2 = The use of colour for text emphasis.

R_3 = The appropriate distribution of colour on a videotex page.

R_4 = The appropriate use of colour variety on a videotex page.

6 OVERVIEW OF THE CONTENT ASSESSMENT SCORING TECHNIQUE (CAST)

Over the past several decades, attention has been directed towards an evaluation of the content characteristics of various types of media. The emergence of videotex has produced demands for a systematic, standardized method for evaluating the content produced.

The content of the videotex page includes the graphics, linguistic characteristics, colour and visual display features. Although these aspects of content can be measured individually, their weighted sum might more accurately represent the total effect of the content on the user.

Uses of the CAST

The Content Assessment Scoring Technique proposed in this document has the following uses:

- to begin to use a systematic approach to content analysis of videotex;
- to determine the specific strengths and weaknesses of content developed on the basis of the areas identified;
- to document changes in content developed as a consequence of analysis;
- to begin to accumulate data that may assist in the clarification of other essential features of content assessment of videotex.

Use of a systematic approach

A major purpose of this document is to identify a possible approach that the industry might use to evaluate and monitor content. In its current state, the CAST does not include the complete set of necessary and sufficient features for a "complete" content assessment. The appropriate total set of relevant features should become more evident as the knowledge base of videotex is enhanced. This approach is meant to initiate a discussion regarding the feasibility of using a more objective systematic, analytical, reliable and consistent approach to evaluating content of videotex programs.

For many years, professionals interested in identifying the salient features of content have relied upon their subjective judgement. Many of these evaluations focused on only one aspect of the content. As a result, the evaluations were either unreliable or did not reveal a comprehensive picture of the nature of the content. The procedures that were used were often not stable or reliable enough to provide sufficient objective information regarding the content. This approach is an effort to provide an objective, efficient measure of videotex content that avoids the pitfalls above.

Determination of the strengths and weaknesses of content

The CAST is not a standardized test but a descriptive technique for evaluation of videotex pages. Its results may be useful for the differential description of videotex content, that is, the determination of strengths and weaknesses in specific areas of content. The CAST measures various features of content including visual display features, readability, human interest, and colour and graphic characteristics. Its results reveal a profile of the specific attributes and potential limitations of the context evaluated. Such information should be useful for the videotex industry as it may provide an orientation for the revision and improvement of content.

Documentation of changes in content

The third use of an approach such as the CAST is to provide a means of evaluating the changes or development of content for videotex over time. This may later provide the field of videotex with a historical perspective on content development. Furthermore, when changes are made to content subsequent to this type of analysis, it can be documented and monitored.

Research

The use of this type of approach to content analysis will be useful for the future development of standardized evaluation tools. Correlations between the features of selected videotex pages and consumer ratings of acceptability may provide valuable insight into the most salient features to aim for in content development.

Theoretical frame of reference

The basis for this analysis system is derived from the various aspects of literature related to content analysis that could be approached with some objectivity. The contributions of many researchers to the technology of text, linguistic features of language, and graphic features were combined.

The linguistic subtest is measured by three readability formulas and a human interest index. The visual features are tested by colour, graphic, and visual display subtests. The content analysis of these subtests involve the use of factors of each of these features.

The graphics subtest consists of an evaluation of the factors of: attention focus, reduction of ambiguity, use of dynamic graphics, place of graphics on the page, and size characteristics of the graphics. These features relate to human interest features of the display. The colour subtest consists of the factors relating to the following: appropriate distribution of colour, colour variety on a page, use of colour for emphasis, and use of colour for optimum legibility. The visual display subtest includes the following factors: chunking cues, typographical cues, spacing cues, line length, and text justification.

In summary, the model generated four subtests, each of which included four or five factors that can be measured. (See detailed discussion in Chapters 2 -5).

Scores and their interpretation

The CAST yields four types of scores: raw scores, subtest percentiles, a composite score, and a weighted score.

Raw Scores

Raw scores are simply the number of items scored 1 on each subtest. Since the number of items and their possible weightings differ in each subtest, the raw scores for each subtest section are not comparable to each other.

Percentiles

Percentiles represent a value on a scale of 100. Percentiles can be obtained for each subtest by determining the number of variables scored as 1 divided by the total number of variables in the subtest and multiplied by 100. Percentile values that are consistently low might suggest that the measures on a particular subtest are not being used appropriately.

Composite Score

The Content Assessment Composite Score is determined by obtaining the sum of the four subtests: the Graphic Content Score (GCS), the Colour Content Score (CCS), the Readability Content Score (RCS), and the Visual Display Score (VDS). The composite content score is an approximation of the "total" content value. The total score is derived by averaging the subtest scores.

Weighted Scores

A weighted score can be obtained by selecting weights Wj, $1 \leq j \leq 4$, where $W1 + W2 + W3 + W4 = 1$. These weights would be empirically or subjectively derived. The weighted score is given by:

$$CAST\ (weighted) = W1\ (GCS) + W2\ (CCS) + W3\ (RCS) + W4\ (VDS).$$

This will yield a score between 0 and 100. The composite score occurs when the Wi are all equal to 0.25.

Summary sheet and profile chart

A summary sheet provides a record of the videotex pages sampled, the raw scores obtained for the subtests the percentiles obtained, actual readability levels obtained, and the composite content score.

FIGURE 1
CAST summary form

Content Analysis Scoring Technique for videotex

Program or Page name _____

Date _____

Evaluator _____

Page creator _____

Composite CAST score _____

Section I RECORD OF SCORES

subtests	raw scores	percentile
I Readability subtest		
II Graphic subtest		
III Visual display subtest		
IV Colour subtest		

Section II CAST PROFILE

	readability	graphic	visual display	colour
Percentiles				
100	.	.	.	.
90	.	.	.	.
80	.	.	.	.
70	.	.	.	.
60	.	.	.	.
50	.	.	.	.
40	.	.	.	.
30	.	.	.	.
20	.	.	.	.
10	.	.	.	.
0	.	.	.	.

In addition to numerical scores, the summary sheet includes a profile chart that may be used to graph the percentiles. Figure 1 is an example of a summary sheet that is used to document the results of a videotex page.

The analysis of a large videotex program containing many pages can be carried out to obtain a Comprehensive Program Assessment Score. In order to determine this comprehensive program score, subtest percentile scores for each page are computed and then averaged. If there is a very large number of pages in a videotex program, and they are similar in content format, a random sampling of such pages may be conducted to select pages for analysis. If pages are significantly different in format, that is, the program consists of text only pages as well as graphics only pages, a stratified sampling methodology should be used to provide a random sampling of the different types of videotex pages used in the program for the analysis.

Specific subtest scores will only be computed for videotex pages for which the analysis is relevant. For example, a graphic subtest score will not be computed for videotex pages that are exclusively text. For each subtest score computed per videotex page, the percentiles are accumulated and a mean percentile score calculated for each of the four subtests. The standard deviations will indicate the variability in scores obtained among the format related videotex pages within a program. The Comprehensive Program Assessment Score is computed by multiplying the sum of the four percentile means by 0.25.

This comprehensive score can provide a reflection of the consistency with which a program meets the identified criteria for each of the subtest categories. The comprehensive score can also provide the industry with an indication of how well a videotex program meets their objectives or designated criteria.

7 CONCLUSION

Much content analysis in the past has been the application of a sorting or counting methodology to text, audio, or audio and visual material. This type of analysis is not only tedious, but does not provide benchmarks by which the videotex industry can assess the production of its program or page content.

This document has focused on parameters of videotex that can be used for evaluation. The established criteria for scoring were based on documented research. The content assessment approach proposed here represents a departure from a solely descriptive sorting task and adds the dimension of content evaluation to the analysis. It is expected that as more research in this field is reported, the components, criteria and approach to the analysis introduced here, will require modification.

The attempt to synthesize the expanse of research available to produce a criterion-based assessment necessitated a review of the relevant areas of artificial intelligence, technology of text, psycholinguistics, linguistics, human factors, engineering and psychology.

It is hoped that this approach will encourage the development of more refined tools to objectively assess the content, not only of videotex, but other areas as well.

REFERENCES

Barmack, J.E., and Sinaiko, H.W. 1966. **Human factors problems in computer generated graphic displays.** Arlington, VA: Institute for Defense Analysis.

Carter, E.C., and Carter, R.C. 1981. Colour and conspicuousness. Vol. 71(6). **Journal of Optical Society of America.** 723-729.

Cahill, M.C., and Carter, R.C. 1976. Colour code size for searching displays of different density. Vol. 18(3). **Human Factors.** 273-280.

Cherry, L. 1982. Writing Tools. **IEEE Transactions on Communications.** 100-105.

Cherry, L., and Vesterman, W. 1980. Writing Tools - The STYLE and DICTION programs. **Computing Science Technical Report No. 91.** Murray Hill, N.J: Bell Laboratories.

Dearborn, W.F., Johnston, D.W., and Carmichael, L. 1951. Improving the readability of typewritten manuscripts. Vol. 37(10). **Proceedings of the National Academy of Sciences.** 670-672.

Dwyer, F.M. 1972. **A guide for improving visualized instruction.** State College, PA: Learning Services.

Flesch, R. 1948. A new readability yardstick. Vol. 32. **Journal of Applied Psychology.** 221-233.

Frase, L.T. April, 1980. Writer's Workbench: Computer supports for writing and text design. Paper presented at the annual meeting of the American Educational Research Association, Boston.

Frase, L.T., and Schwartz, B.J. 1979. Typographical cues that facilitate comprehension. Vol. 71(2). **Journal of Ed. Psychology.** 196-206.

Frase, L.T., MacDonald, N.H., and Keenan, S.A. 1981. Intuition, algorithms, and a science of text design. **Designing Usable Text.** Edited by T. Duffy, and R. Waller. New York: Academic Press.

Glynn, S.N., and DiVesta, F.J. 1979. Control of prose via instructional and typographic cues. Vol. 71. **Journal of Ed. Psychology.** 595-603.

Graham, C.H. 1965. Discriminations that depend on wave lengths. **Vision and Visual Perception.** Edited by C.H. Graham. New York: John Wiley.

Gombrich, E.H. 1972. The visual image. **Scientific American.** 227-82-96.

Harley, J. 1978. **Designing Instructional Text.** London: Kogan Page; New York: Nichols.

Hartley, J., Burnhill, P., and Davies, L. 1979. The effects of line-length and paragraph denotation on the retrieval of information from prose text. Vol. 12(2). **Visible Language.** 183-184.

Hartley, J., and Mills, R. 1973. Unjustified Experiments in Typographical Research and Instructional Design. Vol. 4. **British Journal Ed. Tech.** 120-131.

Horn, R.E. 1982. Structured Writing and Text Design. **The Technology of Text.** Edited by D.H. Jonassen. Englewood Cliffs, New Jersey: Educational Technology Publications.

Kincaid, J.P., Yasutake, J.Y., and Geiselhart, R. 1967. **Use of the Automated Readability Index to assess comprehensibility of Air Force technical orders** (SEG-TR-67-47). Wright-Patterson AFB, Ohio: Aeronautical Systems Division.

Klare, G.R. 1974-75. Assessing readability. Vol. 10. **Reading Research Quarterly.** 62-102.

Klare, G.R. 1963. **The Measurement of Readability.** Ames, Iowa: Iowa State University Press.

Kosslyn, S.M. 1975. Information representation in visual images. Vol. 7. **Cognitive Psychology.** 341-370.

MacDonald, N.H., Frase, L.T., Ginrich, P.S., and Keenan, S.A. 1982. The Writer's Workbench: Computer Aids for text analysis. Vol. 30(1). **IEEE Trans. Comm.** 105-110.

Myatt, B., and Carter, J.M. 1979. Picture preferences of children and young adults. Vol. 27. **Educational Communications and Technology Journal.** 45-53.

Reynolds, L. 1982. Display problems for teletext. **The Technology of Text.** Edited by D.H. Jonassen. Englewood Cliffs, New Jersey: Educational Technology Publications.

Sallio, P. and de Legge, J. October 1982. Quelques aspects opérationnels concernant l'utilisation des contrastes colorés dans le cas d'images purement textuelles sur écran vidéotex. Rennes, France.

Smith, E.A., and Kincaid, P. 1970. Derivation and validation of the automated readability index for use with technical materials. Vol. 12. **Human Factors.** 457-464.

Smith, E.A., and Senter, R.J. 1967. **Automated Readability Index** (AMRL-TR-66-22), Wright Patterson AFB, Ohio: Aerospace Medical Division.

Smith, J.M., and McCombs, M.E. 1971. The graphics of prose. Vol. 48. **Journalism Quarterly.** 134-136.

Ton, W.H. 1969. Optimal visual characteristics for large screen displays. Vol. 6(4). **Information Display.** 48-52.

Travers, R.M.W., and Alvarado, V. 1970. The design of pictures for teaching children in the elementary school. Vol. 18. **AV Communication Review.** 47-64.

Treurniet, W.C. 1979. Spacing of characters on a television display. **Processing of Visual Language II.** Edited by P. Kolers, M.E. Wrolstad, and H. Bouma. New York: Plenum Press.

Vanderkolk, R.J., and Herman, J.H. July 1975. **Dot matrix display symbology study.** TRC Report T76-2172.

Wohlwill, J.F. 1975. Children's responses to meaningful pictures varying in diversity: Exploration time vs. preference. Vol. 20. **Journal of Experimental Child Psychology.** 341-355.

APPENDIX

APPENDIX

Pattern of "Reading Ease" Scores (Flesch, 1948)

"Reading Ease" Score	Description of Style	Typical Magazine	Syllables per 100 Words	Average Sentence Length of Words
0 to 30	Very difficult	Scientific	192 or more	29 or more
30 to 50	Difficult	Academic	167	25
50 to 60	Fairly difficult	Quality	155	21
60 to 70	Standard	Digests	147	17
70 to 80	Fairly easy	Slick-fiction	139	14
80 to 90	Easy	Pulp-fiction	131	11
90 to 100	Very easy	Comics	123 or less	8 or less

Pattern of "Human Interest" Scores (Flesch, 1948)

"Human Interest" Score	Description of Style	Typical Magazine	Percentage of Personal Words	Percentage of Personal Sentences
0 to 10	Dull	Scientific	2 or less	0
10 to 20	Mildly interesing	Trade	4	5
20 to 40	Interesting	Digests	7	15
40 to 60	Highly interesting	New Yorker	11	32
60 to 100	Dramatic	Fiction	17 or more	58 or more