Cartoon Graphics in the Communication of Accounting Information for Management Decision Making

Malcolm Smith[*]
Richard Taffler[**]
Lynda White[***]

Abstract

The content and presentation of information impacts on the efficiency of information processing and the effectiveness of management decision making.

The relative merits of schematic faces for the communication of multivariate information have been explored in a number of disciplines, having initially been developed by Chernoff (1971) to facilitate the clustering of geological samples on multiple attributes. Existing studies in the financial environment suggest that they may be superior to conventional methods in both their communication and decision making qualities. Continuing interest in their further development in the management accounting literature is evidenced by Hemmings (1996), Smith and Taffler (1996) and Gifford (1997).

This paper employs an innovative research design to demonstrate the relative usefulness of schematic faces in the failed/non-failed decision making context compared to accounting statements and financial ratios. An optimum assignment of financial variables to facial characteristics is suggested, one which demonstrates the usefulness of schematic faces as a decision tool in the financial management environment. Schematic faces are shown to be processed more quickly and with no loss of accuracy, compared to more traditional means of presenting financial information, findings which have implications for the way in which graphics are employed in the management decision making process.

Keywords

Accounting Communication; Decision Making; Schematic Faces; Financial Reporting.

Introduction

Relatively little attention has been paid by accounting academics to the question of improving the communicative ability of financial statements. Schulz and Booth (1995), for example, review the task efficiency and effectiveness advantages of graphical versus tabular presentations as part of the analytical review procedures in modern audit practice; they recommend (p.128) that future research should consider alternative presentation formats. Accounting data is essentially multivariate and its assessment depends on the simultaneous effect of several variables in different spheres of activity. Complex tabular presentations do not facilitate an integration of the key features of the accounts and a segmented multi-column format may leave an indication of separate aspects of performance rather than an overall assessment. An alternative means of presentation might provide a clearer and more efficient representation, complementing existing methods. The exploration of alternative communication methods, which might be superior to others for certain decision tasks, and the education of the user in the application of these methods, provides a motivation for this study.

Conventional pictorial methods are extremely limited in their application. Traditional graphs and charts work well in

[*] School of Accounting and Information Systems, University of South Australia, Adelaide, South Australia
[**] Cranfield School of Management, Cranfield, UK
[***] Imperial College of Science and Technology, London

The authors acknowledge the helpful comments of participants to the BAA Conference, Manchester and the AAANZ Conference, Cairns on previous versions of this paper.

only two or three dimensions and quickly become over-complicated when multivariate information is employed. Working within three dimensions is extremely advantageous from a communications point of view, but in many practical instances this is rarely possible if more than a superficial overview is to be conveyed. Many alternative pictorial methods have been employed in an attempt to facilitate the communication of information - ranging from the familiar bar and pie charts and pictograms to more obscure forms. The pie-chart, bar chart and trend graph have become familiar and acceptable in the financial report as alternatives to the narrative and numerical form. Pictorial methods, especially those able to represent several dimensions simultaneously in a form that may be perceived holistically, may potentially be useful in this supporting role.

Valentine (1986) views the human face as a series of vectors in multidimensional space with dimensions corresponding to significant features. A matching of significant features with financial performance measures, therefore, provides the possibility of communicating multidimensional financial information in a simple, integrated and readily comprehensible form. This paper explores empirically the usefulness of the schematic face as a communication device, in a particular decision context, compared with more conventional presentation formats.

The paper addresses the relative usefulness of schematic faces, financial ratios and accounting statements as information formats for management decision making, together with the importance of the manner of the assignment of financial variables to facial characteristics in the usefulness of schematic faces.

Literature Review

Chernoff (1971) initiated `faces' whose features can be made to vary in size and shape according to the value of the assigned variable. The familiarity of faces and their ease of recognition and description makes them potentially superior to other pictorial forms of representation.

Laughery, Alexander and Lane (1971) emphasise the importance of certain areas, namely the eye and mouth regions, which are more mobile than others and so convey more information about an individual's mood, assisting us in comprehending what they are trying to communicate. McKelvie (1973), comparing alternative forms of schematic face suggests that the interaction of eyebrow slant and mouth curvature provides an effective force in the communication of meaning in facial expressions. These findings among others in the psychological literature provide evidence employed in the facial constructions in this study.

Empirical evidence on the detection of facial cues and the interpretation of faces suggest that the assignment of appropriate variables to facial components is important. Bruckner (1978), Moriarity (1979) and Stock and Watson (1984) describe studies involving random and author-selected assignments of variables to facial features paying insufficient attention to the relevant psychological literature on feature saliency. Smith and Taffler (1984) reference the appropriate psychological literature but fail to suggest an optimal feature assignment.

Everitt (1978) justifies a random assignment procedure as a means of reducing the problem of subjectivity caused by different observers using different features of the face to judge their similarity, but experimental evidence from Chernoff and Rizvi (1975) suggests that the use of random permutations in the assignment of facial parameters may effect classification task error rates using faces by a factor of as much as 25 per cent. Their results establish that certain facial features carry little significant information under certain conditions but the artificial situation does not permit a clear evaluation of the relative efficiency of different facial features.

The seminal empirical accounting applications fail to incorporate all of the available evidence: Moriarity (1979) does not attempt to manipulate eyebrow slant in the construction of schematic faces; Stock and Watson (1984) in the assignment of facial features, employ a feature direction the inverse of that suggested by the literature, and accord the nose a prominent assignment role contrary to the suggestions of the relevant psychological evidence.

Smith and Taffler (1996) incorporate the available psychological evidence in conducting an experiment, with users of varying sophistication, to determine the impact on task efficiency and effectiveness of different information processing media in a failed/non-failed decision making environment. They observe that schematic faces are processed more quickly than either financial ratios or accounting statements, and result in significantly fewer Type I (missed failure) errors, with no corresponding increase in the number of Type II (overprediction of failure) errors. However, they do not attempt to vary the assignment of facial features to financial variables.

None of these studies appropriately addresses the problem of subject variability, nor do they consider the effect of differential priors or differential misclassification costs. Although the significance of the dimensionality of the data in influencing discriminatory ability remains in doubt, the need for a clear decision on the assignment of features is well established (e.g., Chernoff and Rizvi 1975; Bruckner 1978), and the psychological literature on saliency should be of great assistance here (e.g. Ekman Friesen and Ellsworth 1972). In these respects and in the provision of an extensive sampling design, employing multiple treatments and varied order of presentation, the present study represents a distinct improvement over its predecessors.

The literature suggests that there will be some differences in classification efficiency employing alternative means of processing. But deficiencies in previous research designs and the chosen feature assignments make the conclusions drawn unsafe. A number of research issues are therefore presented for consideration:

1. The difference in processing times between the faces and conventional means of communicating accounting information.

2. The difference in the number of misclassification errors resulting from the use of faces compared to those from financial ratios or accounting statements.

3. The dependence of the number of misclassification errors resulting from the use of the faces on the assignment of financial variables to facial characteristics.

4. The difference in the costs of misclassification associated with alternative feature assignments.

5. The dependence of the number of misclassification errors on the response time for each of the processing media.

Research Method

A group of 100 MBA Finance students of the City University Business School, London, comprise the respondents. They are required to make failed/non-failed decisions on a group of companies when presented with financial information for those companies represented alternatively in the form of simplified accounting statements, financial ratios and schematic cartoon faces. An evaluation of the solvency of companies, with whom they might potentially do business, is viewed to be an important management decision: incorrectly classifying distressed companies as failures may involve a self-fulfilling prophesy in that they are no longer able to attract the funds necessary to secure their survival; similarly, failing to identify potential future failures will result in avoidable losses being incurred by creditors, investors, employees and the community.

Sample Selection

An initial research sample is chosen comprising 33 pairs of failed and non-failed quoted manufacturing companies with annual accounting year ends between December 1978 and June 1985[1]. These companies are selected and matched on a number of key criteria, in order to provide a set of companies which is large enough to encompass the whole range of performance possibilities, but small enough to allow the generation of efficient experimental designs consistent with the realistic administration of empirical work.

A matching principle is employed to sample pairs of companies with many matching attributes, but differing significantly in aspects of their financial performance. Five essential characteristics of the matched pair are specified:

- the two companies comprise a failed/non-failed combination. Following Taffler (1983), receivership, voluntary liquidation and compulsory liquidation are taken to be evidence of failure;
- the two companies display Z-scores of opposite sign. The failed company is predictable as "distressed" on the basis of its final published annual disclosure prior to failure, recording a negative Z-score based on Taffler (1983), while the non-failed enterprise records a positive sign.
- the two firms are members of a common industrial sector ideally having common product areas.
- the firms are of the same "order" of size; ideally they have equivalent size, as measured by sales turnover, external to the organisation and net of sales taxes.
- the firms have common financial year ends to facilitate comparison and minimise the effects of external economic factors.

A sub-sample of 30 companies (fifteen pairs) is selected from this matched major sample to provide a diverse set of financial profiles, based on only three financial ratios. Chernoff and Rizvi (1975) suggest that experiments of this type examine all of the alternative multiple combinations in view of the likely variation in results from alternative assignments. The specification of only three financial variables allows this process to take place within an efficient experimental design. A linear combination of these three financial ratios representing profitability, liquidity and financial gearing is sufficient to classify correctly all company cases[2]. The resulting test instrument comprises all possible variations of feature assignment and provides the opportunity of varying the base rate for failed companies.

Experimental Design

Accounting statement information is presented in a simplified form, both for content and format to follow the principles advocated by Ehrenberg[3] (1977). A similar procedure is adopted for the financial ratios. Only one years' data, is provided for analysis, for failed companies this being the last disclosure prior to failure. The means and standard deviations of financial ratio variables for manufacturing companies are

[1] The data selected is at least fifteen years old to ensure the survival of those companies deemed 'non-failed' as an assurance that they were not displaying early signs of impending failure at the time of the experiment.

[2] A simple unit weighted model:
$Z = \frac{PBIT}{TA} - \frac{QA}{CL} - \frac{TL}{NW}$ and a decision rule of "Z < 0 is predicted failed" produces two Type II errors, but no Type I errors among the 30 sample cases. A slightly more sophisticated 'simple' model, incorporating the manufacturing industry means (IM) for each ratio:

$$Z = \left(\frac{PBIT}{TA} - IM_{PROF}\right) + \left(\frac{QA}{CL} - IM_{LIQ}\right) - \left(\frac{TL}{NW} - IM_{GEAR}\right)$$

and the same decision rule correctly classifies all of the sample cases.

[3] Ehrenberg (1977) suggests that tables should contain rounded values, columns aligned for ease of comparison, means and totals where possible, some degree of redundancy and plenty of white space.

provided, since they are implicitly employed in the construction of facial profiles. Schematic faces are computer generated using the program detailed in Wang (1978, p.115)[4], as illustrated in Figure 1.

A total of eight experimental tasks is considered:

S: analysis of accounting statements
R: analysis of financial ratios
A,B,C,D,E,F: the six assignments possible of 3 financial variables to 3 facial variables.

Each subject has 5 tasks in all (S, R and 3 of A,B,C,D,E,F in some order) and generates three responses for each:

1. the time taken
2. the number of failed companies correctly detected
3. the number of non-failed companies correctly detected.

The group of one-hundred respondents is assigned randomly to one of 10 random sub-groups of equal size. The layout for each group of ten subjects is a 5 × 10 row-column design for the eight treatments, with the columns as subjects and the rows as time-periods, as shown in Table 1.

Thus:

- each subject processes the accounting statements;
- each subject processes the financial ratios;
- each subject processes three sets of faces, so that each set of faces is processed five times.

Each of the two treatments (Statements and Ratios) occurs twice in each time period while every other treatment occurs once in each time period. The columns of the array form a balanced incomplete block design, ignoring time-periods, Statements and Ratios. In addition the treatments are allocated so that for any subject, the Statements and the Ratios tests are separated by at least one Faces test.

Each sub-group's experimental materials comprise the same ten companies with, therefore, a common number of failed companies. This design therefore allows processing time and misclassification errors to be considered as a function of:

- the number of failed cases;
- the order of processing;
- the variable-feature assignment;
- the alternative information source employed.

---

[4] Facial features are drawn within a range of values corresponding to three standard deviations either side of a mean. This allows an exact mapping of financial ratio values. For example, consider the mapping of profitability onto the mouth, where the 'mean' position of the mouth is of a straight horizontal line (i.e., zero curvature). The profit ratio can be represented as number of standard deviations away from the industry mean $\left(\text{i.e.,} \frac{\text{Ratio} - \text{Mean}}{\text{Standard Deviation}}\right)$ where a positive value will give an up-turned smiling mouth and a negative value a down-turned mouth, numerically corresponding in curvature with the ratio value. (see Smith and Taffler) 1996, for sample illustrations.

FIGURE ONE: Computer Generated Schematic Faces

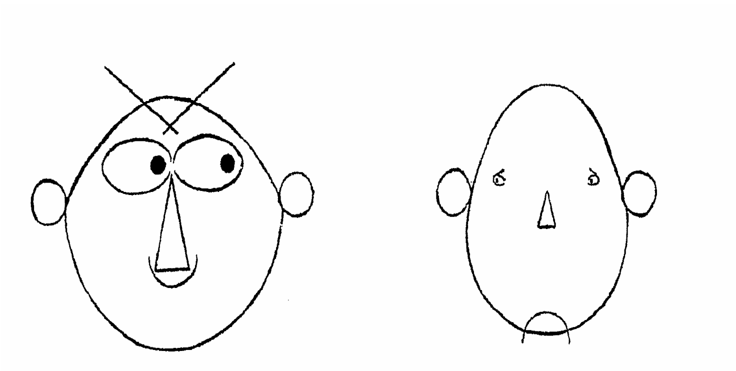

TABLE 1: Experimental Design: Distribution of Tasks

		Subject									
		1	2	3	4	5	6	7	8	9	10
Order of Presentation	1	S	S	C	D	B	R	R	F	E	A
	2	A	B	S	S	D	E	F	R	R	C
	3	R	F	E	C	S	S	B	A	D	R
	4	B	R	R	F	E	C	S	S	A	D
	5	C	D	F	R	R	B	A	E	S	S

TABLE 2: Alternative Assignments of Financial Variables to Facial Features

	ALTERNATIVE ASSIGNMENTS					
	A	B	C	D	E	F
Mouth Curvature Mouth Length	Profitability	Profitability	Liquidity	Liquidity	Financial Risk	Financial Risk
Eye Size Pupil Position	Financial Risk	Liquidity	Profitability	Financial Risk	Profitability	Liquidity
Eyebrow Angle Nose Length	Liquidity	Financial Risk	Financial Risk	Profitability	Liquidity	Profitability

Experimental Procedures

For the six alternative assignments of financial variables to facial characteristics respondents are provided with pictorial guidance as to which feature assignments have been employed in each set of faces with which they are provided. Table 2 details the six alternatives together with feature assignments[5].

In each part of the exercise respondents process the financial information relating to ten companies sequentially with each case presented on a separate sheet of A4 paper[6]. Respondents are informed that the ten companies are different in each section, and random numbering systems ensure that cases cannot be compared meaningfully. Respondents are further informed that any number of their cases can be those of failed companies, but that it is most likely that the set comprises a mixture of failed and non-failed companies. In view of the potential effects of alternative failure probabilities, the design of this experiment incorporates sets of test materials containing, respectively, 1,3,5,7 and 9 failed cases, though the actual base rates are unknown to the respondents.

There was no time limit set for the exercise, though in practice respondents knew that timetabling constraints would restrict the experiment to a two-hour period. During the course of the experiment the respondents were monitored to ensure appropriate completion of the test instrument (in particular the time taken for each section). When close to completing all parts of the exercise, respondents were issued with a 'feedback' sheet to indicate their reaction to the experiment and detail the information cues employed. This 'feedback' sheet was used as a device to relieve peer pressure - by giving early-finishers further tasks to complete - and no formal analysis of these sheets was intended.

Results

When a subject assesses ten different companies, the response-time and the numbers of errors of each type (Type I - "designated healthy when actually failed" or Type II - "designated failed when actually non-failed") are observed. We first analyse these separately and then examine how the response times and errors are related.

Analysis of Response-Times

The average observed response-times (in minutes) for each of the eight treatments (Statements (S), Ratios (R) and the six Feature Assignments (A,B,C,D,E,F)) for the different numbers of failed companies are given in Table 3.

The last line of Table 3 gives the average observed response times for the "Faces" method averaged over the six feature assignments. The average observed response times for Statements and for Ratios are each based on twenty separate observations for each number of failed companies, while those for each individual feature assignment are based on ten observations. For each number of failed companies, each subject unavoidably assesses the same ten companies in each time-period, thus any comparisons between different numbers of failed companies are also comparisons between the five groups of subjects and between the different sets of companies used for the experiment.

[5] Four facial features (mouth, eyes, nose and eyebrows) are varied, but as Table 2 indicates, this involves variations in six characteristics.

[6] Appendix 1(a), (b), (c) respectively illustrate the scope of the data provided for some of the same companies (differently numbered) for accounting statements, financial ratios and one feature assignment of the schematic faces.

TABLE 3: Average Observed Response-Times in Minutes to Process Ten Companies

	Number of Failed Companies					
Treatment	1	3	5	7	9	Average
S	12.3	19.9	23.7	11.9	14.6	16.5
R	8.1	7.2	9.6	9.0	7.7	8.3
A	4.7	3.9	7.1	6.8	4.1	5.3
B	3.7	5.9	4.6	5.2	6.9	5.3
C	4.5	5.1	5.0	4.6	4.3	4.7
D	6.0	6.1	6.1	5.7	4.8	5.7
E	5.7	5.1	6.1	5.7	5.3	5.6
F	4.2	4.2	5.6	5.8	4.2	4.8
Faces	4.8	5.1	5.8	5.6	4.9	5.2

Table 3 suggests that (a) there is little difference in the response times for the different feature assignments and (b) the response-times for Statements are higher than those for Ratios, which in turn are higher than those for Faces. The longest response-times, for Statements, Ratios and Faces, are when there are five failed companies presented to the subjects.

Examination of the response-times suggests a log-transformation and for each separate number of failed companies a normal theory homoscedastic linear model of the form

$$E(\log Y_{ijk}) = \mu + \pi_I + \rho_j + \tau_k \quad \textbf{(Model 1)}$$

is fitted, where Y_{ijk} is the response-time for time-period i, subject j and treatment k and μ, π_i, ρ_j, τ_k represent the effects of the overall mean, the time-period, the subject and the treatment, respectively. A plot of residuals against expected normal order statistics produces a very satisfactory model fit, with the exception of three outliers which are removed from this part of the analysis.

The sum of squares for treatments (corrected for all other terms) is decomposed into three orthogonal components:

1. one degree of freedom for the comparison between the two traditional methods and the Faces methods;
2. one degree of freedom for the comparison between the Statements and the Ratios methods; and
3. five degrees of freedom for comparisons between the six alternative feature assignments.

The Mean Squares for these components and for the Residual term are shown below in Table 4.

For each number of failed companies, the appropriate F-tests shows that the differences between the effects of the eight treatments is significant at the 1% level, while there is no evidence of any differences between the effects of the six Feature Assignments. The difference between the effects of the traditional methods and the Faces method is significant at the 1% level, as was the difference between the effects of the Statements and the Ratios methods. This is true irrespective of the number of failed companies.

TABLE 4: Mean Squares for Treatments and Residual for Different Numbers of Failed Companies, based on Log Response-Times

	Number of Failed Companies				
	1	3	5	7	9
Treatments (7 d.f.)	2.3	3.1	4.4	2.0	2.1
(S+R) vs Faces (1 d.f.)	14.5	13.7	22.7	12.1	10.8
S vs R (1 d.f.)	1.5	7.2	7.5	1.0	3.8
Faces (5 d.f.)	0.0	0.2	0.2	0.1	0.0
Residual (d.f. in brackets)	0.16 (69)	0.17 (68)	0.14 (69)	0.12 (67)	0.14 (69)

The residual mean squares for the different numbers of failed companies is very similar allowing the fit of the following linear model for the log response-times of all 100 subjects:

$$E(\log Y_{ijkl}) = \mu + \pi_i + \tau_k + \phi_l + \rho_{l(j)} \qquad \textbf{(Model 2)}$$

where Y_{ijkl} is the response-time for the j th subject to receive failed companies, in time-period i and treatment k; ϕ_l is the effect of l failed companies, $\rho_{l(j)}$ is the effect of the j th subject to receive l failed companies and π_i and τ_k are the effects of the time-period and the treatment respectively.

The terms for time-periods, number of failed companies and for subjects within numbers of failed companies (corrected in each case for the fitting of the other terms) are all significant, but by far the most significant effect is that for Treatments. Another model based on the log response-times, fitted to investigate a possible interaction between Treatments and the number of failed companies, did not supply any evidence for such an interaction.

All the above analyses lead to the conclusion that subjects take longer to process companies using Accounting Statements than they do to process them using Ratios and that Ratios take longer than the Faces method, there being no significant differences between the alternative Feature Assignments. We may, therefore, conclude that faces may be processed significantly faster than by conventional means.

The residual mean square (0.147) from Model 2 is used to derive the standard errors of the average observed response times of Table 3, which are shown in Table 5.

It appears that the response-times for the Faces method are subject to less variability than those of the other, traditional, methods.

For each number of failed companies, the appropriate F-tests shows that the differences between the effects of the eight treatments is significant at the 1% level, while there is no evidence of any differences between the effects of the six Feature Assignments. The difference between the effects of the traditional methods (statements and ratios) and the Faces method is significant at the 1% level, as is the difference between the effects of

TABLE 5: Standard Errors for the Mean Response Times for Statements, Ratios and Faces

	Number of Failed Companies					
Treatment	1	3	5	7	9	S.E. of Average Response-Time
Statements	1.1	1.7	2.0	1.0	1.3	0.6
Ratios	0.7	0.6	0.8	0.8	0.7	0.3
Faces	0.2	0.3	0.3	0.3	0.2	0.1

the Statements and the Ratios methods. This is true irrespective of the number of failed companies.

The above analyses lead to the conclusion that subjects take longer to process companies using Accounting Statements than they do to process them using Ratios and that Ratios take longer than the Faces method, there being no significant differences between the alternative Feature Assignments. We may, therefore, conclude that faces may be processed significantly faster than by conventional means.

The standard errors of the average observed response times of Table 3, suggest that the response-times for the Faces method are also subject to less variability than those of the other, traditional, methods.

Analysis of Errors

Each subject, in a particular time-period, is asked to assess ten different companies. In each case the response is binary (failed or non-failed). There are two types of error which a subject can make:

1. Type I: declaring a failed company to be non-failed.
2. Type II: declaring a non-failed company to be failed.

Let M_1 and M_2 be the numbers of errors of these two types. These will clearly depend on the subject, the time-period and the treatment. If f (= 1, 3, 5, 7 or 9) is the number of failed companies presented to the subject, then M_1 and M_2 have upper limits of f and 10-f respectively, and $M = M_1 + M_2$ is the total number of errors made by the subject, in one particular time-period. The average proportions of errors, of each type, for each treatment and for each number of failed companies is shown in Table 6 parts (a), (b) and (c). The last line of each table shows the average observed proportions of errors when the results for all the Feature Assignments are pooled.

The proportion of Type I errors appears to rise when the number of failed companies reaches 7 or 9, whereas for Type II errors, the proportion of errors dips slightly when there is a high proportion of failed companies. When we combine the two types of error we find that the proportion of total errors rises when we have a high proportion of failed companies.

None of the relevant results in Table 6 indicate any differences between the different Feature Assignments for either Type I or Type II errors.

TABLE 6: Average Observed Proportions of Errors for Each Treatment and Each Number of Failed Companies

(a) Type I Errors ("non-failed when failed")

	Number of Failed Companies				
Treatment	1	3	5	7	9
Statements	0.10	0.18	0.17	0.34	0.44
Ratios	0.15	0.13	0.27	0.37	0.52
A	0.10	0.07	0.10	0.41	0.43
B	0.20	0.10	0.14	0.40	0.44
C	0.00	0.20	0.14	0.23	0.42
D	0.20	0.23	0.06	0.30	0.47
E	0.30	0.00	0.18	0.40	0.34
F	0.50	0.03	0.22	0.41	0.48
Faces	0.22	0.11	0.14	0.36	0.43

(b) Type II Errors ("failed when non-failed")

	Number of Failed Companies				
Treatment	1	3	5	7	9
Statements	0.24	0.26	0.19	0.03	0.00
Ratios	0.19	0.21	0.22	0.12	0.10
A	0.22	0.19	0.14	0.20	0.50
B	0.11	0.23	0.20	0.10	0.20
C	0.28	0.19	0.19	0.07	0.00
D	0.17	0.09	0.10	0.10	0.00
E	0.17	0.23	0.17	0.10	0.10
F	0.22	0.10	0.12	0.10	0.20
Faces	0.16	0.17	0.15	0.11	0.17

(c) All Errors

	Number of Failed Companies					
Treatment	1	3	5	7	9	Average
Statements	0.22	0.24	0.18	0.25	0.40	0.25
Ratios	0.17	0.19	0.22	0.30	0.48	0.28
A	0.20	0.15	0.13	0.35	0.44	0.24
B	0.10	0.19	0.18	0.31	0.42	0.25
C	0.25	0.19	0.16	0.18	0.38	0.22
D	0.20	0.13	0.09	0.24	0.42	0.20
E	0.18	0.16	0.18	0.31	0.32	0.23
F	0.20	0.08	0.17	0.32	0.45	0.24
Faces	0.17	0.15	0.15	0.29	0.41	0.23

In order to model the errors made by the subjects, it is assumed in the first instance that a subject's assessments of the ten companies are independent and that each company is equally difficult for the subject to assess. With these assumptions, M_1 and M_2 have Binomial distributions with indices f and 10-f respectively. Generalised Linear Models (see McCullagh and Nelder 1983, Chapter 2), each with a logistic link function and Binomial errors, are fitted for each type of error (I and II) and for each number of failed companies. It is assumed that for each number of failed companies, and for a particular type of error,

$$\log\frac{p_{ijk}}{1-p_{ijk}} = \mu + \pi_i + \rho_j + \tau_k \qquad \textbf{(Model 3)}$$

where p_{ijk} is the probability of an error (of that particular type) and μ, π_i and ρ_j are parameters for the effects of the overall mean, time-period, subject and treatment, respectively.

These models are fitted using GLIM 3.77. For each model, a measure of goodness of fit, the deviance, is produced. For any particular effect in a model (e.g. treatments, subjects) we can calculate the deviance with and without this effect. Under the null hypothesis that the effect is non-existent, the change in deviance so produced has asymptotically (as the total number of observations tends to infinity) a chi-squared distribution with degrees of freedom equal to those for the effect in question. McCullagh and Nelder (1983, p.28) suggest that this chi-squared approximation is often unreliable and particularly so if the expected numbers of errors is less than one. In this experiment, this is the case for Type I errors when f = 1, 3 or 5 and for Type II errors when f = 5, 7 or 9. It is useful to look at the changes in deviance due to different effects in these models and Table 7 gives the results for (a) Type I errors and (b) Type II errors for each number of failed companies. Similar models are also fitted to the proportions of total errors with the results shown in Table 7(c). In each case, the changes in deviance due to:

- the comparison between the two traditional methods and the Faces methods,
- the comparison between the Statements and the Ratios methods, and
- comparisons between the six alternative Feature Assignments

are given, together with the residual deviance obtained by fitting the model with all terms included.

TABLE 7: Deviances for Binomial Models for Each Number of Failed Companies

(a) Type I Errors

	Number of Failed Companies				
	1	3	5	7	9
(S+R) vs Faces (1 d.f.)	1.08	2.03	6.84	0.02	2.18
S vs R (1 d.f.)	0.45	0.93	3.92	0.30	2.75
Faces (5 d.f.)	10.25	12.83	9.31	8.60	1.51
Residual (69 d.f.)	45.10	37.92	66.00	79.65	73.24

(b) Type II Errors

	Number of Failed Companies				
	1	3	5	7	9
(S+R) vs Faces (1 d.f.)	0.15	5.76	10.23	2.48	3.82
S vs R (1 d.f.)	1.50	1.72	2.11	2.20	1.35
Faces (5 d.f.)	1.30	5.34	4.23	8.51	*
Residual (69 d.f.)	70.38	103.39	74.63	36.42	16.38

* No convergence using GLIM

(c) All Errors

	Number of Failed Companies				
	1	3	5	7	9
(S+R) vs Faces (1 d.f.)	0.01	6.93	13.25	0.23	1.04
S vs R (1 d.f.)	1.00	2.20	5.14	1.08	3.18
Faces (5 d.f.)	1.25	4.32	11.79	9.88	2.04
Residual (69 d.f.)	61.21	87.68	64.31	56.85	62.21

When Model 3 is fitted to the observations for a particular number of failed companies, we can assess the goodness of fit from the residual deviance, which should, if the model fits well, be from a chi-squared distribution with 69 d.f. The upper 1% point of this distribution is 99.2 and the upper 5% point is 89.4, so the only set of data providing a significant departure from the model is that for 3 failed companies and Type II errors.

While we should beware of using the raw deviances to assess significance, or otherwise, of certain model effects, Table 7 suggests that among the results for Type I errors (f = 7 or 9) and Type II errors (f = 1 or 3), the cases where the expected frequency is greater than or equal to one, the only deviance which exceeds the upper 5% point of the appropriate chi-squared distribution is that for "traditional vs new methods" for Type II errors and f = 3 (where deviance at 5.76 is greater than 3.84, the upper 5% point of χ^2_1). None of the relevant results in Table 7 indicate any differences between the different Feature Assignments for either Type I or Type II errors. Amongst the deviances in Table 7(c) for all errors, those which exceed the upper 5% point of the appropriate chi-squared distribution are for "traditional vs new methods" for f = 3 and for all the results (except the residual deviance) for f = 5. The results for all errors need to be viewed with some caution as we have seen (Table 6) that Type I and Type II errors have different probabilities, especially for higher values of f, the number of failed companies, indicating that M may not have a Binomial distribution.

A Generalised Linear model for the data from all subjects in which

$$\log\frac{p_{ijkl}}{1-p_{ijkl}} = \mu + \pi_i + \tau_k + \phi_l + \rho_{l(j)}$$

(Model 4)

where p_{ijkl} is the probability of an error in time-period i when the j th subject to receive failed companies has treatment k, and μ, π_i, τ_k, ϕ_l and $\rho_{l(j)}$ are parameters for the overall mean, the time-period, the treatment, the number of failed companies and the subject, respectively, is fitted for each error Type (I and II) and for all errors. The corresponding deviances are shown in Table 8 below.

TABLE 8: Deviances for Binomial Models for All Subjects

	Type I Errors	Type II Errors	All Errors
(S+R) vs Faces (1 d.f.)	4.18	5.13	7.86
S vs R (1 d.f.)	3.92	0.02	1.89
Faces (5 d.f.)	3.24	5.92	3.34
Residual (389 d.f.)	360.95	387.69	393.34

Since the residual deviances for this last model are very close to their degrees of freedom (389), the fit of Model 3 to the full set of data appears to be satisfactory.

There is some suspicion that a subject might somehow detect that he/she was being presented with the same ten companies in each time-period and would use this information to assess the companies. However, the binomial model error patterns considered here fit so well that they seem to go some way to allaying this suspicion.

The calculation of goodness of fit of a binomial model of decision making from Table 8 shows that the deviances for "traditional vs new methods" for Type I errors, Type II errors and for all errors exceed 3.84, the upper 5% point of chi-squared on one degree of freedom indicating that the traditional methods may be less accurate than the Faces methods. In addition, the deviance for "Statements vs Ratios" for Type I errors also exceeds 3.84. All of these results suggest a significant superiority for the schematic faces as a processing medium. There is, however, no evidence of any differences between the Feature Assignments.

Differential Misclassification Costs

Since Type I errors are more serious than Type II errors and Altman, Haldeman and Narayanan (1977) and Zmijewski (1984) suggest that a Type I misclassification error is of the order of 40 times more serious than a Type II misclassification error, we consider the following "scores", which reflect a different weighting attributable to different types of error.

Where M_1 = No. Type I Errors, M_2 = No. Type II Errors, f = No. Failed Companies (out of 10), the score is a weighted combination of the observed proportions of the two types of error and places 40 times as much weight on Type I errors as it does on Type II errors. Each score is scaled to lie between 0 and 1 with a low score representing an accurate response and a high score an inaccurate response. The average scores for different values of f and for different treatments are displayed in Table 9.

For each number of failed companies, the best (i.e., lowest) "faces" result is starred (*).

TABLE 9: Weighted Scores Reflecting Different Misclassification Costs

Treatment	Failed Companies				
	1	3	5	7	9
S	0.12	0.19	0.17	0.34	0.44
R	0.15	0.14	0.27	0.37	0.52
A	0.11	0.07	0.10	0.41	0.43
B	0.19	0.11	0.14	0.40	0.44
C	0.03*	0.20	0.14	0.23*	0.42
D	0.19	0.22	0.06*	0.30	0.46
E	0.28	0.01*	0.18	0.40	0.34*
F	0.43	0.04	0.22	0.41	0.48

Table 9 suggests that where the incidence of failure is at its lowest, most closely corresponding with economic reality, relative misclassification costs of Type I and Type II errors show Faces with Feature Assignment C to generate the optimum score. However, the variability of outcomes supports the contention of Smith and Taffler (1996) that the holistic facial representation is more important than the specific feature assignment.

Relationship Between Response-Times and Accuracy

The accuracy of a subject may well be influenced by the time taken to respond, so it is instructive to examine how the response-times and the total number of errors (M) are related for different treatments. Table 10 shows the average observed response-times for different values of M and for Statements, Ratios and Faces. Numbers in brackets give the numbers of observations over which an average is taken and a dash denotes no observations.

TABLE 10: Average Observed Response-Times (minutes) for Different Numbers of Errors

Treatment	M = Total Number of Errors								
	0	1	2	3	4	5	6	7	8
Statements	9.9 (12)	15.9 (17)	19.2 (28)	18.7 (16)	18.3 (12)	15.1 (10)	12.7 (3)	6.0 (2)	-
Ratios	5.4 (12)	9.5 (22)	8.0 (14)	8.6 (18)	10.0 (12)	9.1 (11)	7.5 (8)	4.3 (3)	-
Faces (mean)	3.6 (34)	5.3 (89)	6.0 (57)	5.7 (47)	5.3 (37)	5.3 (24)	4.4 (7)	3.0 (3)	3.5 (2)

There seems to be some indication for the Statements method that those who respond quickly are either very accurate, or very inaccurate, while those who take longer to respond have intermediate levels of accuracy, being neither very accurate nor very inaccurate. This effect is rather less marked for Ratios and Facial Feature Assignments. Overall, although there is sufficient evidence to justify a non-linear relationship between response time and the incidence of errors for Statements, there is little evidence of a relationship for either Ratios or Faces.

However, it is likely that some cases would have been more difficult to process than others, and that individual respondents would have had differing information requirements and different cognitive processing styles. These issues need to be addressed in future research.

Conclusions

Evidence is provided that schematic faces are processed more quickly than either of the more traditional methods of information presentation, with no loss of accuracy. The weighted error scores (Table 9) reflecting different misclassification costs, demonstrate the superiority of facial representation; the average response times (Table 3) demonstrate the marked time differences. Repeated misclassifications are consistent with overemphasis on the profit figure. Where the facial profiles produce misclassifications not apparent with the other processing media, this is consistent with undue overemphasis on the mouth as a facial characteristic. Where profitability is assigned to the mouth such overemphasis is further amplified.

Evidence is generated towards the provision of an optimum assignment of financial variables to facial characteristics consistent with the psychological evidence on the saliency of features. Preference for assignment C is ultimately dependent on the cost of a wrong decision in circumstances where the incidence of failure is closest to economic reality. However, the results (Table 9) suggest that the facial representation itself is more important than any specific feature assignment.

Feedback from subjects suggests that knowledge of the actual feature assignments is unimportant, most decisions being made independent of a reference guide. This finding is consistent with that of Moriarity (1979), suggesting that the facial profile is processed as a compromise between features within a facial context, without the necessity of referring to financial meaning. The use of faces, rather than the choice of a particular feature assignment, appears to be of paramount importance.

Feedback also suggests that very little use is made of the means and particularly of the standard deviations of the financial ratios provided. This may be attributable to a lack of statistical education or a lack of facility with matters statistical. Either way a potentially powerful piece of information with which to gauge relative financial performance is largely ignored. This may contribute to faces outperforming financial ratios as an information source, since the former implicitly include mean and standard deviation measures in their construction ensuring that these items of information cannot be ignored entirely.

A research design is employed which overcomes the deficiencies of earlier studies in this area, notably the use of multiple treatments to address the problem of subject variability and the consideration of the effects of differential priors and differential misclassification costs. The results demonstrate the usefulness of schematic faces as a decision tool in the financial environment with the potential to have a significant impact on information processing in management decision making.

The outcomes of this work prompt a further discussion of those aspects which would facilitate improvements in the use of schematic faces as a communication and information processing medium. Further research would address an increased number of variables, and associated increases in levels of processing complexity, even though this increases the number of alternative combinations so much that only a sample may be contemplated; alternative pictorial representations might be introduced to provide a further standard for comparisons; the individual differences between users might be recognised by explicitly including gender, culture, cognitive processing style and personality variation in the variables to be tested, especially if one of our ultimate objectives is the tailoring of management decision making systems to the needs of the specific user. For practical purposes we might see schematic faces as a complementary data representation to accompany traditional datasets; further work must consider the impact of complementary information, appealing to different senses, on decision efficiency and effectiveness.

References

Altman, E.I., Haldeman, R.G. and Narayanan, P. (1977), "Zeta Analysis: A New Model to Identify Bankruptcy Risk of Corporations", *Journal of Banking and Finance,* (June), pp. 29-54.

Bruckner, L.A. (1978), "On Chernoff Faces" in P.C.C.Wang (Ed.) *Graphical Representation of Multivariate Data*, New York, Academic Press, pp. 93-122.

Chernoff, H. (1971), *The Use of Faces to Represent Points in n - Dimensional Space Graphically*, Technical Report No.71, Department of Statistics, Stanford University, California.

Chernoff, H. and Rizvi, H.M. (1975), "Effect on Classification Error of Random Permutations of Features in Representing Multivariate Data by Faces", *Journal of the American Statistical Association,* 70, pp. 548-54.

Ekman, P., Friesen, W.V. and Ellsworth, P. (1972), *Emotion in the Human Face*, Pergamon Press, New York.

Ehrenberg, A.S.C. (1977), "Rudiments of Numeracy", *Journal of the Royal Statistical Society* (Series A), 140, part 3, pp. 277-297.

Everitt, B.S. (1978), *Graphical Techniques for Multivariate Data*, Heinemann, London.

Gifford, D. Jr. (1997), "Facing up to the Numbers", *Harvard Business Review*, May/June, pp. 10-11.

Hemmings, J. (1996), "Smile, please", *Management Accounting (UK)*, 74(6), p. 62.

Laughery, K.R., Alexander, J.F. and Lane, A.B. (1971), "Recognition of Human Faces: Effects of Target Exposure Time, Target Position, Pose Position and Type of Photograph", *Journal of Applied Psychology,* 51, pp. 477-83.

McCullagh, P. and Nelder, J.A. (1983), *Generalized Linear Models*, Chapman and Hall, Cambridge.

McKelvie, S.J. (1973), "The Meaningfulness and meaning of Schematic Faces", *Perceptions and Psychophysics,* 14(2), pp. 343-348.

Moriarity, S. (1979), "Communicating Financial Information through Multidimensional Graphics", *Journal of Accounting Research,* 17, April, pp. 205-224.

Schulz, A.K-D, and Booth, P. (1995), "The Effects of Presentation Format on the Effectiveness and Efficiency of Auditors' Analytical Review Judgments", *Accounting and Finance*, 35(1), pp. 107-132.

Smith, M. and Taffler, R.J. (1984), "Improving the Communication Function of Published Accounting Statements", *Accounting and Business Research,* 54, Spring, pp. 139-146.

Smith, M. and Taffler, R.J. (1996), "Improving the Communication of Accounting Information through Cartoon Graphics", *Accounting, Auditing and Accountability Journal*, 9(2), pp. 68-85.

Stock, D. and Watson, C.J. (1984), "Human Judgement Accuracy, Multidimensional Graphics and Human versus Models", *Journal of Accounting Research,* 22(1), Spring, pp. 192-206.

Taffler, R.J. (1983), "The Assessment of Company Solvency and Performance using a Statistical Model: A Comparative UK

Based Study", *Accounting and Business Research,* 15(52), Autumn, pp. 295-308.

Valentine,T . (1986), "Encoding Processes in Face Recognition", unpublished PhD thesis, University of Nottingham.

Wang, P.C.C. (ed.) (1978), *Graphical Representation of Multivariate Data*, Academic Press, New York.

Zmijewski, M.E. (1984), "Methodological Issues Related to the Estimation of Financial Distress Prediction Models", *Journal of Accounting Research*, Vol.22 (Supplement), pp. 59-82.

APPENDIX 1(a): Experimental Data - Financial Statement Analysis

	1	8	9	12	25
PROFIT AND LOSS ACCOUNT	(£000)	(£000)	(£000)	(£000)	(£000)
Sales Turnover	32,212	29,865	20,219	4,973	18,240
Trading Profit	1,980	2,925	1,806	576	1,936
Depreciation	470	512	229	277	626
Pre-Interest Profit	1,510	2,413	1,577	299	1,310
Interest	123	350	207	-91	143
Pre-Tax Profit	1,387	2,063	1,370	390	1,213
Tax	105	418	492	184	201
Profit after Tax	1,282	1,645	878	206	1,012
Dividends	209	239	401	378	248
Extraordinary Items	-239	0	5	0	-22
Retained Profits	834	1,395	482	-172	742
Earnings per Share	51.1p	20.7p	4.2p	0.73p	9.99p
BALANCE SHEET					
Stocks	6,120	6,963	4,854	990	4,350
Debtors	5,859	5,653	5,232	1,414	4,856
Cash	991	463	212	425	466
Other	237	0	0	1,159	0
Current Assets	13,207	13,079	10,298	3,988	9,672
Creditors	7,609	6,155	3,857	693	1,765
Overdraft and Short-Term Loans	796	280	1,417	0	2,616
Tax	325	300	449	281	636
Dividends	0	124	252	216	248
Other	0	0	0	0	0
Current Liabilities	8,730	6,859	5,975	1,190	5,586
Net Current Assets	4,477	6,220	4,323	2,798	4,086
Tangible Fixed Assets	6,268	6,703	3,215	1,604	4,756
Intangibles	0	0	192	0	0
Investments	0	0	0	266	76
	10,745	12,923	7,730	4,668	8,918
Share Capital and Reserves	10,182	10,346	6,155	4,501	8,703
Long Term Loans	0	2,376	583	0	54
Minority Interests	0	22	54	0	159
Provisions and Deferred Tax	563	179	938	167	2
Government Grants	0	0	0	0	0
	10,745	12,923	7,730	4,668	8,918

APPENDIX 1(b): Experimental Data - Financial Ratio Analysis

CASE	PROFITABILITY PBIT/TA	LEVERAGE TL/NW	LIQUIDITY QA/CL
2	-0.077	1.461	0.475
11	-0.042	0.987	0.582
13	-0.151	3.010	0.477
19	0.156	2.626	0.406
20	0.096	1.068	0.180
28	0.114	0.841	1.128
32	0.056	1.114	0.574
40	0.090	0.636	0.953
42	0.023	2.168	0.357
43	0.062	3.589	0.373
Mean	0.07	1.12	0.82
Standard Deviation	0.06	1.28	0.11

APPENDIX 1(c): Schematic Faces Analysis: Sample Materials

(Cases 9, 30, 38, 51 for Assignment D)