

Doctor-patient communication in rheumatology: studies of visual and verbal perception using educational booklets and other graphic material

J M H MOLL

From the Sheffield Centre for Rheumatic Diseases, Nether Edge Hospital, Sheffield S11 9EL

SUMMARY Patients (n=404) with osteoarthritis and control subjects (n=233) were studied to examine the communicational value of five styles of illustration (cartoon (C), matchstick (M), representational (R), symbolic (S), photographic (P) and two levels of text ('easy', 'hard'), presented as educational booklets about osteoarthritis. Booklet comprehension was tested with a multiple choice questionnaire (MCQ) scored by two raw score and two, more sensitive, weight-of-evidence methods. Further studies assessed perception of image detail, tone, and colour by ranking, rating, latency, and questionnaire methods. A subgroup was tested psychometrically. The main findings were: (a) pictures in booklets enhance communication; (b) perception of pictorial style depends on its vehicle of presentation, cartoons being most effective in booklets, photographs overall; (c) simplifying text does not significantly enhance communication; (d) certain picture-text 'interactions' appear to increase comprehension (e.g. 'hard' text with 'easy' pictures); (e) several 'endogenous' factors are associated with increased comprehension: 'psychological' (e.g., intelligence, memory, reading skill); 'demographic' (e.g., the young, males, higher social grades, higher educational levels); 'disease' (e.g., longer disease duration, previous information about the disease).

Key words: physician-patient relations, medical communication, psychometrics, graphics, instructional handbooks, visual aids, books—illustrated.

The foundation of all learning consists in representing clearly to the senses sensible objects, so that they can be appreciated easily.

COMENIUS, 1592–70

The standard of medical care largely depends on satisfactory communications between staff and patient, not least in rheumatology. This is often of an alarmingly poor standard and needs to be improved.^{1 2} This is compounded by the fact that hospital patients have poor recall.^{3 4} Failure to achieve satisfactory rapport with patients may lead not only to wasted resources but also to harm or even death.⁵

With these concerns to the fore an interest in doctor-patient communication was fostered in the Rheumatism Research Unit, University of Leeds some years ago and led to publications^{6 7} concerned with evaluating and improving the illustrated book-

lets of the Arthritis and Rheumatism Council (ARC).

These studies raised certain questions about the value of pictures in the booklet format, and these, together with a need to look at textual factors in more detail, led to the studies reported here.

Materials and methods

MATERIALS

General plan of studies

The studies were divided into three parts aimed at deriving information about patients' reactions to visual material, both pictures and words, with particular reference to educational booklets.

Part 1. Patients with osteoarthritis (OA) (n=

373, male (M) 140, female (F) 233) were each randomly allocated an experimental booklet from a total range of 12 formats. Two to four weeks later subjects were given a multiple choice knowledge testing questionnaire based on the booklets.

Part 2. A subgroup of the above (n=84, M 24, F 60) was tested further by means of a battery of psychological tests relating to intelligence, personality, memory, reading ability, and compliance, and also certain perceptual tasks involving picture ranking and rating in the context of the experimental booklets.

Part 3. A small group (n=30, M 10, F 20) of patients with mixed rheumatological diagnosis was examined for perception of certain aspects of detail, tone, and colour in the context of the booklets.

The patients were derived from two centres (Sheffield 300; Leeds 73) and the completion rate (patients willing to participate ÷ patients available for study × 100) was 85.7%.

Control subjects

Several control groups were involved:

- (a) to assess booklet/no booklet difference (n=31, M 12, F 19);
- (b) to assess patient/non-patient difference (n=100, M 41, F 59);
- (c) to assess knowledge testing sensitivity of questionnaire (n=41, M 35, F 6).

Questionnaire

This was a multiple choice knowledge testing questionnaire based on material about osteoarthritis covered in the experimental booklets. The questionnaire contained 24 questions with five items in each question (total items=120). It went through several design and pilot stages before printing, with the aim of producing a balanced research tool of maximal reliability and validity.⁸ The total number of 'Yes' items virtually equalled the 'No' items. The number of questions based on illustrated booklet facts virtually equalled the number based on non-illustrated facts. The definitive questionnaire, a stapled four page document printed in 'Times Roman' was designed to look formal but not formidable, and was prefaced by carefully worded instructions including some in rubric.

Booklets

The experimental books were on the subject of osteoarthritis and the text was based on that of the ARC booklet on that subject. In order to provide a reasonable range of picture/text format five styles of illustration (representational, matchstick, cartoon, symbolic, photographic) and two levels of text ('hard'—standard ARC text; and 'easy'—text sim-

plified by Flesch's formula⁹) were used. In addition, two unillustrated booklets were included to provide a total of 12 formats. All line illustrations were prepared by the author and were aimed at contrasting graphic styles and at comparability of representation of subject matter. Examples are shown in Fig. 1, together with a photographic 'equivalent'. Both the text and the pictures were piloted through several stages to 'weed out' confusions or too difficult material. The booklets were printed by offset lithography on high quality paper in 'Times Roman'. The print was double spaced and set in 11 pt ('Times Bold' 11 pt was used for headings and 'Times Italic' 10 pt for captions) on sheets measuring 7"×4" (standard ARC booklet size). The booklets contained 30 pages (15 leaves).

METHODS

Psychometric and perception tests and tasks

- (a) Reading skill—*Holburn reading scale*.¹⁰
- (b) Intelligence—*Peabody picture vocabulary test*.¹¹
- (c) 'De-Americanised' version of original neuroticism and introversion/extraversion scales. *Eysenck personality questionnaire*.¹²
- (d) *Long term memory test* (simple name and



Fig. 1a



DON'T do this!

Fig. 1b



DON'T do this!

Fig. 1c



DON'T do this!

Fig. 1d



DON'T do this!

Fig. 1e

Fig. 1 Examples of illustrative styles used in the experimental booklets: (a) representational, (b) matchstick, (c) cartoon, (d) symbolic, (e) photographic.

address test containing 10 elements—numbers and words—and none having any relevance for citizens of Sheffield). Specially designed for survey. Scored simply on basis of correct elements remembered (one point for each).

(e) *Compliance matrix*. Specially designed for survey, consisting of questions to do with diet, exercises, and medication. Compliance scored 'no', 'sometimes', and 'yes' (carried arbitrarily weighted scores, 0, 1, 2 respectively). A catch statement carrying penalty marks was incorporated.

(f) *Picture ranking test*. Set of cards containing six scenes: 'osis'; 'diagram'; 'obesity', 'causes', 'lifting', 'tablets'. These scenes covered a wide spectrum of graphic devices and also ranged widely through the booklet. Each scene was represented by five different styles (cartoon, matchstick, photographic, representational, symbolic), giving a total of $6 \times 5 = 30$ pictures. Subjects asked to put cards in order of preference.

(g) *Picture rating task*. Set of 30 cards described in (f) used. Rating carried out with a plain 10 cm line visual analogue scale of the type used by Huskisson.¹³ The distance between the mark placed by the subject on the scale line and the 'very good' pole gave a measure of 'goodness'—pictorial appreciation.

(h) *Inter- and intrainage comparison latencies*. These experiments were based on the assumption that reaction times (latencies) to exposure of picture pairs would mirror preferences; a short reaction time implying little doubt because of a clear cut preference; a long reaction time implying indecision associated with less well defined preferences. Times were measured by Heuer 'Trackmaster' stopwatch reading to 1/10 s.

(i) *Tone/colour tasks*. All subjects previously tested with a shortened version of the Ishihara test.¹⁴

(j) *Tone perception task*. Series of 13 tonal values ranging from black to white with 11 intervening shades of grey derived from a standard graphic source (Letraset). Subjects asked to put discs (arranged in a circle in random order) in order of increasing lightness. Task timed.

(k) *Colour naming task*. Series of 10 coloured discs presented in random order—purple/brown/white/orange/green/grey/red/blue/black/yellow. Subjects asked to name colours and nearest interpretation recorded (e.g., 'bluish green' recorded as green). Task not timed.

(l) *Colour preference task*. Series of 16 coloured discs. Subjects asked to remove discs in the order in which they liked them best. Only first 10 required.

(m) *Colour-pair comparisons*. A series of 10 double colour discs presented to subjects who were

asked (i) which colour of each pair preferred and (ii) which colour pair preferred.

(n) *Colour-triple comparisons*. A series of 10 triple colour discs presented to subjects who were asked (i) which colour in each triple preferred and (ii) which colour triple preferred.

(o) *Picture preference questionnaire*. Designed to cross check with the detail/tone/colour tasks and to assess preferences concerning—humorous v serious pictures, photographs v line drawings, and pictures with captions v pictures without captions.

Questionnaire scoring procedures

(a) *Raw analysis*. Two methods used to obtain raw scores (because of present controversy about raw scoring): *0/+1 system*^{15 16}—one mark awarded for each correct response, no marks for incorrect responses or statements not attempted. *-1/0/+1 system*—one mark awarded for each correct response, no marks for non-attempts, and -1 for each incorrect response.

(b) *Weight-of-evidence analysis*. This approach designed to counter the effect of guessing and provide a measure of the accuracy with which questionnaires were answered. Using special tables raw scores (obtained by 0/+1) were converted to 'creds' (the negative logarithm of the probability that the examinee is *not* guessing). In the study creds scores were obtained by the methods of both Bevans¹⁷ and Good.¹⁸

Computometric and statistical methods

The University of Sheffield ICL 1900 Implementation-Release 2 computing system was used. Statistical data were derived from programs based on the statistical package for the social sciences—version 5.01.¹⁹ The printout included sum, mean, mode, standard deviation, variance, standard error of mean, percentage, n value, and 'Anova' (analysis of variance) tables showing sum of squares of freedom, mean square, Snedecor's 'F'²⁰ and Student's 't'²¹ values. Correlations were expressed by Spearman's rank correlation coefficient (r).

Results

STUDIES TO EXAMINE THE EFFECTS OF BOOKLET PICTORIAL AND TEXTUAL FACTORS

Do booklets improve recall?

Table 1 shows that patients exposed to booklets scored significantly higher marks (by means of all four scoring methods) than control subjects not exposed to a booklet.

Do pictures add to words?

The assumption that pictures add to words in

Table 1 *Do booklets improve recall?*

Group	n	Scores			
		0/+1	-1/0/+1	Bevans	Good
OA patients not exposed to booklet	31	54.4 (±2.69)	27.2 (±3.09)	40.3 (±6.57)	35.5 (±6.35)
OA patients exposed to booklet	373	63.5 (±0.75)	39.2 (±1.06)	69.5 (±2.79)	63.7 (±2.69)
Significance of difference (p)		=0.002	=0.001	<0.001	<0.001

Values are mean±SD.

conveying medical information was supported by the following:

(a) Exposure to illustrated booklets led to higher questionnaire scores than exposure to unillustrated booklets, but the difference was not statistically significant.

(b) Exposure to certain types of booklet, such as the cartoon- and matchstick-illustrated booklets (associated with standard text) led to significantly higher questionnaire scores ($p < 0.05$, Students' *t* test) when compared with subjects exposed to unillustrated booklets with comparable text.

(c) Analysis of individual questionnaire item

scores (regardless of whether based on illustrated, unillustrated, or unrepresented facts) showed that subjects who had read an illustrated booklet achieved more scores that were significantly significant than subjects who had read either an unillustrated booklet or no booklet.

(d) The picture preference questionnaire showed that 70% of subjects preferred reading material with rather than without illustrations ($p = 0.028$).

Are some styles of illustrations more effective than others?

Some styles of illustration were found to be more

Table 2 *Studies of factors affecting recall of medical facts*

Nature of enquiry	n	Procedure	Results
Is recall better for some facts than for others?	373	MCO analysed according to item type: (see text) —causes of disease —nature of disease —clinical features —self help features —treatment aspects	Types of facts recalled in the order: 1. Self help features 2. Clinical features 3. Treatment aspects 4. Causes of the disease 5. Nature of the disease (Pictorial style may differentially influence fact recall)
Is recall related to order of presentation of textual material?	373	MCO analysed according to order in which items featured in text	(a) 'Recency' rather than 'primacy' was observed (later facts remembered best) (b) Relationship between order and recall was significant for illustrated facts ($r = 0.4101$, $p = 0.005$), but not for unillustrated facts ($r = 0.1905$, $p = 0.058$)
Is recall related to their degree of textual repetition?	373	MCO analysed according to degree of booklet statement repetition: —not mentioned (distractors) —repeated × 1 —repeated × 3 —repeated × 3 or more	(a) Recall in general not convincingly related to repetition (b) However, a significant difference ($p < 0.05$) was found between the categories 'no mention' and '×1 (or more) mention'
Is recall related to differences in format of the graphic image?	310	MCO analysed according to format of booklet graphics: —without verbal message —with legends —with slogans —with captions —with legends and captions	The two formats containing legends (hand lettered) scored significantly ($p < 0.05$) less than the other three formats

Table 3 Studies of detail, tone, and colour perception*

Area of study	No of patients	Procedure	Results
Image detail	30	Twelve variants of 'lifting scene', each paired with control (see text)	(a) 33.6% failed to discriminate variant from control (b) 'Eccentric', 'reversed', and 'tinted' variants provided most difficulty (c) One third failed to judge identical pair as identical (d) Image position and size created more difficulty than outline and content (e) Inverse relationship ($r = -0.4876$, $p = 0.045$) between correct assessment and reaction time
Tonal shades	30	Thirteen discs, each with a different tonal shade (see text)	(a) Most subjects achieved roughly correct <i>absolute</i> order ($r > 0.8$, $p < 0.002$) (b) Failing this most achieved correct <i>relative</i> order (c) No relationship between task time and accuracy
Colour naming	30	Ten cards, each tinted with a conventional colour. All subjects had normal colour vision (Ishihara's test)	100% accuracy achieved by all subjects
Colour preference: Colour v black and white	30	Picture preference questionnaire: 'Do you prefer reading material with black and white or colour pictures?'	90% preferred colour to black and white pictures ($p < 0.00001$)
Single colours	30	Sixteen coloured cards (see text)	Dark blue and orange were equally the most popular colours, being ranked first preference by 20% of subjects
Colour pairs	30	Ten discs, each tinted on one side with two colours of equal area (see text)	Colour preferred in each pair: green (59.18%); blue (55.85%); yellow (55.83%). Colour pair preferred: brown-yellow (30%); yellow-green (16.7%); yellow-blue, red-yellow, blue-brown (each 10%)
Colour triples	30	Ten discs, each tinted on one side with three colours of equal area (see text)	Colour preferred in each triple: yellow (40.0%); green (37.95%); red (31.85%). Colour triple preferred: brown-yellow-green (23.3%); green-red-yellow (23.3%); yellow-blue-green (20.0%)

* No significant differences between patients (n=30, M10, F20) and a control group of 10 (M3, F7) healthy subjects.

effective than others, and the variation depended on the way in which the pictorial material was presented (i.e., in booklets or on cards).

(a) *The booklet study.* All four scoring methods showed superior scores in the group exposed to cartoons. The 'matchstick' group was second and the 'photographic' group third. (The ranking of the other styles depended on the scoring system used.)

(b) *The ranking study.* In terms of most first places, superior scores were for symbolic drawings, followed in order by cartoons, photographs, representational drawings, and matchstick figures. (In terms of at least last places, matchstick figures scored best, followed in order by photographs, cartoons, representational drawings, and symbolic drawings.)

(c) *The rating study.* Little difference between styles was shown, but in rank order they were: photographic, representational, symbolic, cartoon, matchstick.

(d) *The latency study.* Scoring according to percentage preferences was in the following order: photographic, representational, symbolic, cartoon, matchstick. (Latencies showed no significant difference between styles of illustration.)

(e) *Picture preference questionnaire.* This showed no significant difference ($p > 0.25$) between preferences for funny and serious pictures. A significant difference ($p < 0.01$) supported a preference for photographs, as opposed to line drawings.

(f) *Overall ranking* of styles from methods (a)–(e): 1. photographic, 2. symbolic, 3. cartoon, 4. representational, 5. matchstick.

Table 4 Psychological factors ($n = 10$)

Variable	n	General comment	Relationship with post-test score	
			r^* or t^*	p
Intelligence (Peabody picture vocabulary test)	84	Significant correlation between questionnaire score and level of intelligence	0.364	<0.001
			0.478	<0.001
			0.454	<0.001
			0.452	<0.001
Personality (Eysenck's neuroticism and introversion/extraversion scales)	84	(a) Significant negative correlation between questionnaire score and neuroticism	-0.114	0.150
			-0.178	0.053
			-0.189	0.043
			-0.187	0.044
		(b) No significant correlation between questionnaire score and introversion/extraversion	0.035	0.375
			0.017	0.438
Long term memory (simple test devised for the study)	84	Significant positive correlation between questionnaire and long term memory score	-0.007	0.475
			-0.007	0.475
			0.205	0.031
			0.159	0.075
Reading skill (Holburn reading scale)	84	Significant positive correlation between questionnaire score and reading skill score	0.163	0.069
			0.164	0.068
			0.349	<0.001
			0.472	<0.001
Compliance (compliance matrix devised for the study)	84	No significant correlation between questionnaire score and compliance score	0.423	<0.001
			0.422	<0.001
			0.053	0.315
			-0.009	0.469
Attitudes towards reading material	371	No significant correlation between questionnaire score and 'enjoyable/not enjoyable' scores or 'helpful/not helpful' scores	0.032	0.385
			0.032	0.385
			$t=0.915$	0.361
			$t=0.018$	0.986
			$t=0.018$	0.986
			$t=0.233$	0.816
Spontaneous comments	371	(a) Significantly higher questionnaire scores among subjects who made spontaneous comments (of whatever type)	$t=0.005$	0.996
			$t=0.819$	0.413
			$t=0.344$	0.731
			$t=0.335$	0.738
		(b) Booklets generating most spontaneous comments were C, P, M, S, R and unillustrated stimulated fewer comments	$t=3.295$	0.001
			$t=2.807$	0.005
'Reinforcement' (number of times booklet read)	371	(a) Analysis of variance showed significant trend pointing to increasing scores with increasing reinforcement (peaking at four times read) (b) Correlation study showed relationship to be significant though not strong	$t=2.161$	0.031
			$t=2.173$	0.030
			Not safely analysable	
			0.1413	0.003
			0.1340	0.005
			0.1247	0.008
Booklet exposure—questionnaire test interval	371	Weak, although significant correlations pointed to higher test scores with lower reading test intervals	0.1246	0.008
			-0.156	<0.05
			-0.134	~0.05
			-0.154	<0.05
Questionnaire completion time	306	Slight indication that the shorter the time to complete the questionnaire the better the score	-0.155	<0.05
			0.001	0.491
			-0.145	0.006
			-0.149	0.004
			-0.015	0.005

* r =Spearman's rank correlation coefficient (all values not prefixed by t); t =Student's test value.

Table 5 Correlations between psychological attributes

	Reading ability	Long term memory	Intelligence	Neuroticism	Introversion/ extraversion	Compliance
Reading ability	—	r=0.1696 p=0.062	r=0.5666 p=0.001	r=-0.1790 p=0.052	r=0.0790 p=0.0237	r=0.0178 p=0.436
Long term memory	—	—	r=0.1743 p=0.056	r=0.0082 p=0.471	r=0.1325 p=0.115	r=0.0016 p=0.494
Intelligence	—	—	—	r=-0.2245 p=0.020	r=-0.0620 p=0.288	r=-0.0580 p=0.300
Neuroticism	—	—	—	—	r=-0.3107 p=0.002	r=-0.1235 p=0.132
Introversion/ extraversion	—	—	—	—	—	r=0.1993 p=0.035
Compliance	—	—	—	—	—	—

Do captions alter preferences for pictures?

The differences between the series with and without captions (some of them statistically significant) were thought to reflect experimental 'set' rather than true perceptual differences.

Does improving textual readability enhance recall?

In general there were no statistical differences between the 'hard' and 'easy' groups. However, the results suggested an interaction between 'hard' text and 'easy' pictures, and vice versa.

STUDIES OF FACTORS AFFECTING RECALL OF MEDICAL FACTS

The results of certain inquiries into booklet factors and their effect on recall are summarised in Table 2.

STUDIES TO EXAMINE PERCEPTION OF DETAIL, TONE, AND COLOUR

The results of these studies are shown in Table 3.

STUDIES OF 'ENDOGENOUS' FACTORS

Analysis of certain groups of endogenous factors of direct or indirect relevance to educational booklet comprehension is subdivided and tabulated as follows: psychological factors (Table 4), correlations between psychological factors (Table 5), physiological factors (Table 6), personal factors (Table 7).

QUESTIONNAIRE AND SCORING FACTORS

The questionnaire in pilot and post-test studies showed reasonable reliability, validity, and discriminatory power (Table 8). The weight-of-evidence methods proved more sensitive than the raw scoring methods, but there was a high degree of

correlation (r=0.8015 to 0.9999, p=0.001) between all four methods. The -1/0/+1 method had an intermediate place between the 0/+1 raw score method and the weight-of-evidence methods.

Discussion

The material reported here comprises a summary of a series of studies. For convenience of discussion they will be divided accordingly into three sections: the effects of booklet pictorial and textual factors; the perception of detail, tone, and colours; and the 'endogenous' factors in rheumatological communication.

STUDIES TO EXAMINE THE EFFECTS OF BOOKLET PICTORIAL AND TEXTUAL FACTORS.

Overall conclusions from these studies are:

(a) Instructional booklets, whether written in 'easy' or 'hard' style, or whether illustrated or unillustrated, improve recall of medical facts in a rheumatological context.

(b) Illustrated booklets are more effective than

Table 6 Physiological factors (n=3)

Variable	n	Relation with post-test score		
		F	p	Test
Eye dominance	30	1.178-0.009	> 0.05	F*
Handedness	30	0.222-0.001	> 0.05	F
Colour blindness	30	Not measurable		—

*Snedecor's F test.

Table 7 Personal factors (n=10)

Variable	n	General comment	Relationship with post-test score		Test
			F	p	
<i>Demographic</i>					
Age	372	Significantly higher scores in younger age groups (particularly < 35 years)	7.774 8.318 6.626 6.613	<0.001 <0.001 <0.001 <0.001	F*
Sex	373	Significantly higher scores among men	13.953 13.821 7.971 7.952	<0.001 ≪0.001 <0.001 <0.001	F
Social class	366	Significantly higher scores in higher social grades (particularly social class II [‡])	2.382 3.751 3.730 3.687	> 0.05 <0.025 <0.025 <0.025	F
<i>Educational</i>					
Educational level	370	Significantly higher scores in the higher educational level patients (particularly among those with tertiary education)	5.354 8.203 9.013 8.798	<0.01 <0.001 <0.001 <0.001	F
School leaving age	370	Significantly higher scores in patients of higher school leaving age (particularly among those not leaving until age 18)	4.780 6.196 5.775 5.649	<0.01 <0.01 <0.01 <0.01	F
<i>Disease</i>					
Presence	131	No significant difference between groups with and without disease	-0.03 -0.59 -1.03 -1.05	0.979 0.561 0.308 0.303	t [†]
Duration	84	Significant trend suggesting higher scores in disease of longer duration	-0.899 -1.289 2.319 2.294	<0.05 <0.05 0.05 < p < 0.10 0.05 < p < 0.10	F
Family history	373	No significant correlation with family history	-1.06 -1.37 -1.46 -1.50	0.291 0.170 0.146 0.134	t
Previous information	373	Significantly higher scores in patients with previous information about the disease	-2.84 -3.11 -3.36 -3.37	0.005 0.002 0.001 0.001	t
Treatment centre	373	Significantly higher scores in the Sheffield patients compared with the Leeds patients	-5.38 -2.96 -2.47 -2.55	<0.001 0.004 0.015 0.012	t

* Snedecor's F test.

† Student's t test.

‡ No subjects in social class I.

Table 8 Questionnaire analysis

Type of analysis	Aim	Result/comment
Pilot studies	To reduce mean score to an acceptable level by increasing plausibility of distractors (false statements)	1st—84.1% 2nd—69.0% 3rd—62.0% 4th—59.6% Satisfactory reduction (towards 'ideal' 50%)
Reliability	To test ability of questionnaire to achieve reproducible results by internal checks and test-retest analysis	Reliability confirmed by both approaches. Test-retest difference 10.6%
Validity	To test whether questionnaire measures what it is supposed to measure	Reasonable gradation of performance in groups with varying rheumatological knowledge and experience confirmed validity: Consultant rheumatologists 91% General practitioners 83% Medical students 79% Qualified nurses 76% Student nurses 78% Dental students 65% Patients exposed to booklet 64% Patients not exposed to booklet 54% Non-medical healthy controls 55%
Discriminatory power	To check that questions and items are not too easy, not too hard, and not too 'neutral' due to bad wording	Anderson's method ¹⁵ : correlations between individual questions (as opposed to items) > r=0.3, the 'good' discrimination level Fieller's modification (personal communication): general pattern satisfactory, many questions achieving high discrimination values (Snedecor F values with p < 0.001)

unillustrated booklets. This finding is supported by the work of Booker,²² who also showed that pictures are important for the speed of learning, whereas print is important for accuracy. Another study of bearing on the present work is that of Hamilton,²³ who found that flat, moving, or still pictures score better than purely textual, oral, 'audio', and three dimensional presentations of educational material.

(c) The preferred style of illustration varied according to the design of the presentation. However, overall, photographs emerged as the most generally favoured style, followed in order by symbolic drawings, cartoons, representational drawings, and matchstick images. Cartoons scored particularly well when presented as booklet illustrations. Preference for photographs fits with the thinking vogue of the era in which we live—an era heavily influenced by a respect for and an appreciation of machine rather than man made articles. Photographs also saturate our environment both in static and 'movie' forms, and books, magazines, newspapers, television, and cinema largely depend on this familiar mode for their effects. For a visually relatively unsophisticated population such 'environmental programming' may be influential. The emergence of the cartoon as the preferred style in the

context of booklet presentation can be interpreted in various ways: a reflection of the general popularity of cartoons in everyday life; an effect of assuaging anxiety in 'trivialising' the significance of rheumatic ailments; a 'warming' effect in promoting relaxation in the doctor-patient relationship. It could also be argued that cartoons might work in the opposite direction in generating irritation among patients who interpret cartoons as being inappropriate in a field that is far from humorous.

(d) Improving textual readability did not lead to improvement of recall. Several previous studies have been concerned with improving the readability of texts (e.g. employee handbooks and pamphlets—Davis and Hopkins²⁴; Farr²⁵; Carlucci²⁶; children's textbooks—Faison²⁷; psychology texts—Ogdon²⁸; Anderson²⁹), but few have been concerned with evaluating the results of such 'improvements'. A study of relevance here is that of Ley *et al*³⁰ who examined comprehensibility of x-ray leaflets using Flesch's principles, as in this study. These workers found increased recall in one type of leaflet (containing cholecystogram details) but not with another (leaflet on barium studies).

(e) The apparent interplay between textual difficulty and style of textual illustration is interesting

and previously undocumented. Further work is needed to examine this effect in more detail. Interpretation of this picture/text interaction could relate to perceptive arousal. For example, when the style of illustration is relatively 'easy' from the point of view of interpretability and familiarity a 'hard' text may be necessary to maintain an optimal level of arousal.

(f) Analysis of the relative recall of booklet facts according to their medical significance, order of presentation, and textual repetition provided data that may be useful in the general planning of future educational booklets. The fact that recall is better for some medical facts (particularly 'self help' and 'clinical features') has been found by others.^{3 4 31 32} The finding that 'primacy' rather than 'recency' is related to improved recall of medical facts (particularly when the facts are illustrated) is at variance with a study by Ley and Spelman¹ and with one by the present author.⁶ Either results can be 'justified' psychologically—'primacy' being related to the impact of first impressions, 'recency' to memory delay with time. The finding that repetition of facts (especially with higher orders of repetition) was not convincingly related to recall militates against general assumptions and evidence in the field of memory and learning in animal and human (non-medical) settings. However, this area needs further study as it could be argued that facts mentioned only once were mentioned only once because they were 'easier' to understand than facts mentioned twice or more.

STUDIES TO EXAMINE PERCEPTION OF DETAIL, TONE, AND COLOUR

The visual perception of patients differed in no significant way from that of healthy controls and varied according to the nature of the task. Thus reasonable competence was shown in correct identification of tonal shades and colours, but some difficulty was found in differentiating between image details (particularly positional and size differences). The studies based on colour showed preferences that could be of relevance to future booklet design (e.g., for covers and textual illustrations). Dark blue and, slightly less so, orange were high preference colours when used in isolation. Yellow, on the other hand, though not popular when viewed in isolation, was probably useful in creating an arousal effect when presented with other colours.

STUDIES OF 'ENDOGENOUS' FACTORS IN RHEUMATOLOGICAL COMMUNICATION

Recall of booklet material was found to be significantly correlated with several psychological and personal factors. In the design and evaluation of

future booklets, importance should be directed to these influences. For example, patients scoring high on intelligence, long term memory, and reading skill are likely to score well on booklet comprehension. Improved scores are also more likely the more times a booklet is read and the shorter the interval between booklet exposure and testing. On the other hand, a decreased performance can be expected in the neurotic and in those who take longer to complete their questionnaire. No correlation was found between booklet performance and compliance rating.

Allowances should also be made for certain 'personal' factors (demographic, educational, and disease factors). Thus higher scores on booklet comprehension can be expected in younger patients, in males, and in those of higher social grade. Superior results are also more likely in those of higher educational level and of higher school leaving age. Certain disease factors are also positively and significantly correlated with booklet comprehension. These include longer disease duration and previous information about the disease.

Other workers have drawn attention to the importance of some of these endogenous factors, though in general little work has been done in this sphere, except in that relating to age. Thus long term memory³³ and intellectual function³⁴ decrease with age; although Birren and Morrison³⁵ stressed that the effect of educational level supersedes that of age, in test performance.

In conclusion, these studies have shown that illustrated educational booklets in the field of rheumatology exert a significant communicational effect. The data reported therefore provide objective support for the current policy by various rheumatology organisations (e.g. ARC, American Rheumatism Association) of using instructional booklets in this field.

The studies were generously supported by a research grant from the Arthritis and Rheumatism Council and led to a PhD (University of Leeds). I am especially grateful to Mrs P Drake for data collection and for technical help; to Drs H G Bevans and N R J Fieller for statistical advice; to Mr P R F Clarke, Dr P Ley, Professor P B Warr, and Mrs S McLaren for psychometric advice; to Mrs Robbie Hopkins for collecting Leeds data; to the Medical Illustration Unit, Northern General Hospital, Sheffield; and to Professors V Wright and H Gwynne Jones for helpful discussions.

References

- 1 Ley P, Spelman M S. *Communicating with the patient*. London: Staples Press, 1967.
- 2 Egbert L D, Battit G E, Welch C E, Bartlett M K. Reduction of postoperative pain by encouragement and instruction of patients. A study of doctor-patient rapport. *N Engl J Med* 1964; 270: 825-7.
- 3 Ley P, Spelman M S. Communications in an out-patient setting. *Br J Soc Clin Psychol* 1965; 4: 114-6.

- 4 Joyce C R B, Caple G, Mason M, Reynolds E, Mathews J A. Quantitative study of doctor-patient communication. *Q J Med* 1969; **38**: 183-94.
- 5 Medical News. *Br Med J* 1978; **ii**: 1796.
- 6 Moll J M H, Wright V. Evaluation of the Arthritis and Rheumatism Council handbook on gout. An objective study of doctor-patient communication. *Ann Rheum Dis* 1972; **31**: 405-11.
- 7 Moll J M H, Wright V, Jeffrey M R, Goode J D, Humberstone P M. The cartoon in doctor-patient communication. Further study of the Arthritis and Rheumatism Council handbook on gout. *Ann Rheum Dis* 1977; **36**: 225-31.
- 8 Oppenheim A N. *Questionnaire design and attitude measurement*. London: Heinemann, 1966.
- 9 Flesch R F. *How to test readability*. New York: Harper and Row, 1951.
- 10 Watts A F. *The Holburn reading scale*. London: Harrap, 1948: 3-8.
- 11 Dunn L M. *Expanded manual for the Peabody picture vocabulary test*. Minnesota: American Guidance Service, 1965.
- 12 Eysenck H J, Eysenck S B G. *Manual of the Eysenck personality questionnaire*. London: Hodder and Stoughton, 1965.
- 13 Huskisson E C. Measurement of pain. *Lancet* 1974; **ii**: 1127-38.
- 14 Ishihara S. *The series of plates designed as a test for colour-blindness*. 38 plates ed. Tokyo: Kanehara Shuppan, 1969.
- 15 Anderson J. *The multiple choice questionnaire in medicine*. Tunbridge Wells: Pitman, 1976.
- 16 Fleming P R, Sanderson P H, Stokes J F, Walton H J. *Examinations in medicine*. Edinburgh: Churchill Livingstone, 1976.
- 17 Bevans H G. Confidence (probability) scoring for the standard progressive matrices and the advanced matrices. *British Psychological Society, Annual Conference*. Swansea, 1966.
- 18 Good I J. Weighing evidence. In: *Probability and the weighting of evidence*. London: Griffin, 1950: 62-75.
- 19 Nie N, Bent D, Hull C. *Statistical package for the social sciences*. New York: McGraw-Hill, 1970.
- 20 Snedecor G W. *Statistical methods*. 4th ed. Ames, Iowa: Iowa State College Press, 1946.
- 21 Siegel S. *Nonparametric statistics for the behavioural sciences*. New York: McGraw-Hill, 1956.
- 22 Booker H R. Relative comprehensibility of pictorial information and printed words in proceduralized instructions. *Hum Factors* 1975; **17**: 266-77.
- 23 Hamilton E A. *Graphic design for the computer age*. New York: Rheinhold, 1970: 147.
- 24 Davis K, Hopkins J O. Readability of employee handbooks. *Personnel Psychol* 1950; **3**: 317-26.
- 25 Farr J N. Readability and interest values in an employee handbook. *J Appl Psychol* 1950; **34**: 16-21.
- 26 Carlucci C. How readable are employee handbooks? *Personnel Psychol* 1951; **4**: 383-95.
- 27 Faison E W J. Readability of children's textbooks. *J Educ Psychol* 1951; **42**: 43-51.
- 28 Ogdon D P. Flesch counts of eight current texts for introductory psychology. *Am Psychol* 1954; **9**: 143-4.
- 29 Anderson W. Readability of readers. *Am Psychol* 1956; **11**: 147-8.
- 30 Ley P, Goldman M M, Bradshaw P W, Kincey J A, Walker C M. The comprehensibility of some x-ray leaflets. *J Inst Health Educ* 1972; **10**: 47-55.
- 31 Grennan D M, Taylor S, Palmer D G. Doctor-patient communication in patients with arthritis. *NZ Med J* 1978; **88**: 431-4.
- 32 Anderson J L, Dodman S, Kopelman M, Fleming A. Patient information recall in a rheumatology clinic. *Rheumatol Rehabil* 1979; **18**: 18-22.
- 33 Gilbert J G, Levee R F. Patterns of declining memory. *J Gerontol* 1971; **26**: 70-5.
- 34 Bromley D B. *The psychology of human ageing*. Pelican 2nd ed. Harmondsworth: Penguin, 1974.
- 35 Birren J E, Morrison D F. Analysis of the WAIS subtests in relation to age and education. *J Gerontol* 1961; **16**: 363-9.