Preferred citation: P.A. Daisley. Quantification of subjective quality. In **The Fundamental Properties of Paper Related to its Uses**, *Trans. of the Vth Fund. Res. Symp. Cambridge, 1973*, (F. Bolam, ed.), pp 709–711, FRC, Manchester, 2018. DOI: 10.15376/frc.1973.2.709.

QUANTIFICATION OF SUBJECTIVE QUALITY

P. A. DAISLEY, Wiggins Teape Research & Development Ltd., Beaconsfield, Bucks.

Introduction

MOST attempts to quantify features of paper involve the use or creation of scientific instruments. This contribution claims attention to an entirely different approach in which human judgments can be organised and structured to provide viable quantification that aligns with practical assessment in production and in the market place. Three of the many applications are given as examples.

Paper cleanliness

THE importance of cleanliness of paper for optical character recognition is obvious, but customer decisions to accept or reject are made in practice not merely on the functional suitability at the readers, but also on the visual appearance of cleanliness. The appearance to the buyer, the printer or the user may well have more effect than the reactions of the reader system!

After striving to find an instrumental evaluation and struggling with the tedious dirt size/count methods,[1] one mill adopted the grid assay technique.[2]

A 1 m^2 grid of 100 squares is placed over a large sample of the paper. Using standardised lighting and examination time, on approved observer counts (not the number of visible dirt spots, but the number of squares in which any dirt is evident), the total count on six of these metre square samples from a reel is used as an indicator of cleanliness.

Initially, this total was used to assess the acceptibility of the reel, from association between marketing views and observed results. Later, using accumulated evidence, changes of process and material inputs were readily detected in the grid assay results. Technical and production actions taken reduced the count to so low a level that acceptance/rejection questions are now irrelevant and control is maintained well below the market danger level.

Under the chairmanship of Dr J. D. Peel

Paper personality

NOT all the subjective judgments we need to make can be regarded as a single dimension or restricted to a single dimension. A recent examination of differences in character of papers could not be identified adequately with test properties nor with traditional assessments of handle. Nevertheless, differences were recognisable to the makers and the users. Because a large amount of business was involved—and possibly at risk—an approach had to be found to identify and quantify the difference that the market recognised.

Since judgments of paper are most likely a combination of several different factors, the possibilities were examined against three of the prime senses—sight, touch and sound. An experiment was undertaken using different observers to assess pairs of sample reams and pairs of sheets for the defined features. Efforts were made to keep each assessment independent of other assessments—for example, when judging for sound, the observer could neither see nor feel the paper concerned.

Using the football league system of comparisons of pairs, all possible combinations were evaluated with two points awarded to the winner, one point for a draw, to give a total quantification for each source of paper. The results confirmed that real differences existed and were related to all three senses. Differences between repeat observation, between observers and between reams and sheets were all small compared with the three-dimensional separation found between the sources of paper. From the original experiments, the papers were found to differ almost exactly in the order of market preference.

Small but real differences, induced by changes of making conditions, were reflected in quantified results from paper personality testing though undetected by non-structured judgments. Subsequent experiments confirmed the earlier findings. This encouraged development of a paper personality test method and examination of process/test relationships.

Shade assessments

IN THE manufacture of tinted papers, major sources of broke and complaints were paper off-shade and poor match. Though instrumental colour matching is now a reality, action was necessary to help to promote better visual judgments of shade as well as better control.

By using the 'Rangefinder technique'[2] to generate a rational scale of shade from actual production samples, it was possible for most papers to isolate the dominant dimension or direction of variation. For tints, the intensity of shade was the biggest contribution to variation. For near-white papers, the brightness was the major feature of visual judgment. Hence, different lighting is appropriate for improved discrimination for different papers.

By arranging the series of specimens generated as 'gateposts', with the spaces between graded in consecutive numbers, a yardstick of intensity has been produced. Samples from current manufacture are compared by approved assessors with the shade scale. The resulting measurements are used for control of the process and for decisions on product acceptability. Though arrived at subjectively, these shade numbers can be treated as a visual measurement of quality, except when changes of hue render the scale inappropriate.

Although rigidly organised, this quantification of shade is readily accepted by production personnel to assist control. Further use of results for selection and classification of webs for sheeting helps to make the best use of paper with slight but acceptably different shade.

Reduced broke and negligible complaints for shade confirm that the practice is effective.

Other applications

THESE other applications include—

Level assessment	Process setting techniques
Printability testing	Product grading
Evaluation of mixing	Organoleptic testing
Print set-off measurement	(packaging testing)

Conclusion

PROBLEMS of subjective judgment of quality are common and much can be achieved by using related objective measures. There are, however, many situations for which carefully structured subjective assessments can be used to derive practical, realistic and repeatable measures of quality.

References

1. TAPPI Standard T 437 ts–63, 'Dirt in paper and paperboard'
2. 'Measuring the unmeasurable', P. A. Daisley, Moscow Conference 1971 of EOQC

Discussion

Dr E. L. Back In these two sessions today, with the eye as a most sensitive instrument not only to compare levels, but also variations over large areas and the speed with which variations occur, these variations have been important means for evaluating paper sheets. In the discussion of mechanical properties, it was a little surprising that these variations and variation spectra have not been mentioned in much detail, although important to end uses. This may be because we have small means for evaluating these variations easily enough. Maybe somebody who has worked with mechanical properties extensively would like to comment.

The Chairman As no one has any experience to offer that would enable them to comment on these observations, I would like to say something that has been in my mind while thinking about the appraisal I shall present tomorrow. We have in this session particularly come nearer to the objective of the symposium. We can become obsessed with the instruments that lie between fundamental properties and end uses, to jump from making paper to using it. Have we benefited in this respect?

Mr M. I. MacLaurin How did you select your observers Mr Daisley and make sure that the standards you demonstrated were not varying? How did you manage to maintain the 'calibration' and keep a check on the system? If you are doing a TAPPI dirt assessment, you can check it with a microscope. How do you know from one production shift to the next that the observers are producing similar results?

Mr P. A. Daisley In the initial selection of observers, we carried out a number of experiments using different papers, using different lighting conditions and using different grid sizes in the case of grid assay. We sought to establish whether observer to observer difference was important. We found that about 10 per cent of those we wanted to use as observers were really

not suitable for making this kind of observation. This is what one expects when making judgments based on visual perceptions.

To take care of the crossover between shifts, we have a strict calibration system by which samples taken on shift A are evaluated on shift A. These samples are transferred, without the information, to shift B, who make their assessment. We have been operating this calibration rigorously for many months and I believe that is why the mill has had no trouble with evaluation of paper cleanliness.

The Chairman We have been strongly recommended today to pay heed to the physiological and psychological features of our assessment of the product itself. We have also seen that instruments can be designed to provide information of subjective properties based on an intelligent understanding of these requirements. Having heard further from Mr Daisley, we actually have a commercial success on a rather different approach to the specification of properties. I feel personally that it has been well rewarding and I would like to thank all the speakers for their contributions.