ANNEX TO FINAL REPORT

July, 1977

DOC CONTRACT # OSU5-0072

SCIENTIFIC AUTHORITY

Mr. L. Durr
Department of Communications
16th Floor, Journal Tower North
300 Slater Street
Ottawa, Ontario
K1A 0C8

PREPARED BY

P. D. Guild
Wired City Laboratory
Department of Systems Engineering
& Computing Science
Carleton University
Ottawa, Ontario
and
Department of Experimental
Psychology
University of Oxford
Oxford, England

For engineers who design telecommunications systems, there is continual trade-off between performance capabilities and cost. Dickson and Bowers (1972) have stated: "The bandwidth of any communication network directly determines its capacity to transmit information. To understand the relative effectiveness of various alternatives for communications one must inevitably ask about bandwidth. Generally speaking, the more bandwidth provided for a communications channel, the greater the cost of the system." This notion is also of central concern in human factors research, but here the focus of attention is usually upon efficient social interaction versus cost limitations (Ried, 1973). It seems natural, therefore, that the examination of the bandwidth parameter should be performed using a person perception task.

By the late 1960's, there were several hundred research publications dealing with the technical aspects of bandwidth compression. At the same time, human factors research was largely limited to subjective assessments of picture quality. The last decade has seen an enormous increase in the number of engineering publications on bandwidth compression but there has been only a handful of studies by social scientists investigating the consequences of signal alteration for person perception. Human factors relating to bandwidth compression remain relatively unknown (McManamon, 1975).

Bandwidth may be compressed by reducing the frame rate of video transmission. Systematic altering of the frame rate from full bandwidth, namely, 30 frames per second, to 15, 10, 5, 2 and finally 1 frame per second would create a unique opportunity to study the perception of emotional expression. Such a continuum might form an empirical connection between the majority of studies concerned with recognition of facial emotion which have used static photographs as stimuli and those minority of studies which have depicted stimulus material dynamically, either live or on film (e.g. Thompson and Meltzer, 1964; Osgood, 1966; Ekman and Friesen, 1969; Thayer and Schiff, 1969; Graham, Bitti and Argyle, 1975).

Several methodological strategies have been used in recent years for the study of emotional expression. Frijda (1969) and Stringer (1973) have reviewed the various approaches and find three main orientations: (a) the dimensional approach (b) the categorical or clustering approach and (c) the hierarchical approach.

All three orientations have demonstrated their usefulness. For reasons that will become apparent, the present investigation will consider the dimensional approach.

Perhaps the best known dimensional approach has been that of Schlosberg (1952 and 1954). He identified three dimensions for the analysis of any facial emotion (a) Attention-Rejection (b) High-Low Activity (or Tension-Sleep) and (c) Pleasantness-Unpleasantness.[1] Although there have been numerous attempts to refine the number and nature of dimensions, these three seem to recur (e.g. Abelson and Sermat, 1962; Thorngate, 1971; Stringer, 1973).

In a recent study of bandwidth compression by Lindsey (1974) one of his conclusions showed that as the degree of activity in the video image increases, subjects rated picture quality more critically under conditions of high bandwidth compression. That is, subjects could tolerate a high compression level with greater ease when viewing a low-action "talk show" than when viewing a high-action "dance scene". For the present study, this would imply that the compression level manipulation should exert greatest influence upon those test items classified as High Activity as opposed to those of Low

Activity.

Thompson and Meltzer (1964) studied the accuracy with which different emotions are judged. They found that the positive emotions such as happiness and pleasantness are more easily understood than negative emotions such as unhappiness and unpleasantness. This too has implications for the present study which has dichotomized the Unpleasantness-Pleasantness dimension. One might expect that as the compression level increases, the Unpleasantness index should show greater impact than the Pleasantness index. The Schlosberg circle (shown in Figure 1) tends to support the notion that positive emotions are classified with greater agreement than negative emotions.

<u>The Schlosberg Dimensions of Facial Emotion</u>

The first step in the present investigation was to generate appropriate stimulus material containing the three Schlosberg dimensions. The filmed stimulus material was a series of emotional reactions depicted by actors. The procedure outlines the way in which this stimulus material was constructed. The person perception film was standardized or normalized at full video bandwidth (30fps) by having subjects offer their interpretations of what emotion or circumstance best describes each stimulus segment. No account was taken of what emotion was intended to be portrayed by the actor. No index of absolute accuracy of judgement or validity of classification was attempted. Instead, the person perception test was designed so that we know how each stimulus segment is typically perceived when viewed at full bandwidth. That is to say, we have established the modal interpretation of these emotional displays. The next step was to segregate these stimuli into polar groups which reflect the three Schlosberg dimensions. This was accomplished by way of a content analysis of subjects' open-ended statements and by reviewing the assembled person perception film. This was done by a person who was blind to the present experimental hypotheses.

The Schlosberg Circle and Predicted Ambiguity of Facial Expression

Another reason to build upon the early work of Schlosberg in the present study is that he has investigated the relative "confusability" in classification of particular facial emotion. The Schlosberg circle (see Figure 1) arranges six basic categories of emotion so that segments which are most similar are beside each other. It was found that fewer errors in the judgement of emotion were made across non-adjacent segments than across those that were adjacent. In this study, we expect to find that classification of emotion, when required within a single circle segment (e.g. joy versus love), should be less consistent than classification across adjacent segments (e.g. joy versus surprise). In turn, the latter should be less consistent than classifications across non-adjacent segments (e.g. joy versus sullen). Moreover, since bandwidth compression is generally expected to reduce consistency of judgement, we might expect the non-adjacent classification task to discriminate compression levels least well and within segment classifications to discriminate compression levels best.

The Semantic Differential Ratings of the Actors' Personalities

By the end of the person perception film, subjects will have seen each actor depict nineteen emotional reactions lasting nine seconds each. From this experience, subjects will have formed some impression of each actor's personality. It might be expected that if these actors have been presented in an artificial way (i.e. extreme bandwidth compression), this artificiality might reveal itself in the semantic differential rating made of the actors themselves. The implications of this should be obvious. If the medium portrays an individual in an unflattering manner, a negative reaction may generalize beyond the medium, to the person himself. A similar prediction may be made with respect to the "social distance" component of the semantic differential. The more bandwidth compression there is, the more perceived distance might be expected between the actor and the subject who is doing the rating.

A further prediction can be made with respect to the semantic differential data. It has been stated already that, as we move down the compression continuum, we expect to find an initial increase in the modal concensus of the emotion classification. In theory, this is due to the fact that full bandwidth is sufficiently rich in cues about the emotional reaction being portrayed, that a certain amount of subjective interpretation is possible. At some point further along the continuum (as yet unknown), we might expect that modal judgements will be maximized because those cues necessary for typical perception are retained but, at the same time, there is less room for selective perception of the emotion. The highest degree of bandwidth compression (1 fps) might again yield greater breadth of interpretation. This time, however, it would be the lack of cues preventing concensual interpretations of the emotion. As well as predicting this result for the mean scores on modal perception, it is expected that the semantic differential rating will show this influence (i.e. larger intra-group variability of rating at 30 fps than at the maximally modal compression level and this latter condition showing smaller intra-group variability rating than at 1 fps). It was Barron et al. (1971) who found that an increase in the amount of information describing an individual may lead to an increase in trait attributions made about the individual. More importantly, the inferences were less accurate than in a condition in which a lower amount of information was used to describe the person. With more information, there appeared to be greater latitude for drawing inferences.

Statement of Hypotheses:

Hypothesis 1

It is predicted that as the compression level increases (from 30 fps to 1 fps), there will first be a slight increase and then a dramatic decrease in the modal perception of the emotional displays. This lowering of the modal perception should occur for all of the Schlosberg dimensions, but most apparently in the High Activity index (see Lindsey, 1974) and the Unpleasant index (see Thompson and Meltzer, 1964).

Hypothesis 2

It is predicted that group concensus on the person perception task will be lowest when classification choices are within one segment of the Schlosberg circle and highest when the choices involve non-adjacent segments. Moreover, it is expected that as the compression level increases, there will first be a slight gain and then a dramatic decline in the classification performance on (a) within segment, (b) adjacent segment, and (c) non-adjacent segment choices. The former is the more difficult judgement and, hence, should be influenced most by the compression continuum. The latter should be least influenced, thus being the easiest to judge.

Hypothesis 3

It is predicted that as the compression level increases the semantic differential ratings which reflect "artificiality" and "social distance" in the actor's personality will also increase. At the point at which the perceptions are shown to be most modal, it is expected that there will be less intersubject variability in semantic differential ratings of each of the actors, than at either extreme of the frame rate continuum (i.e. 30 fps and 1 fps).

Procedure:

Subjects:

A total of 122 subjects (equal number of males and females) volunteered to participate in a communications experiment in the Wired City Laboratory, Department of Systems Engineering and Computing Science, Carleton University, Ottawa. They were first year undergraduates, and for their participation they received course credit in introductory psychology.

Preparation of the Stimulus Material

The aim was to produce a series of "emotional interactions" between an actor and an imaginary other person communicating via a teleconference link. The imaginary person was the source of the "emotional reaction" depicted by the actor (e.g. surprise, anger, fear, etc.) and each reaction was then filmed. An actor (male or female) was seated at a teleconference node in the Wired City Laboratory, Carleton University, Ottawa and the only situational or contextual cues provided were those associated with a teleconference interaction. The top half of the actor's body was clearly visible, with the head, shoulders, chest, arms and hands being used to convey emotion. The actor always communicated the emotional reaction in equivalent fashion, holding all things constant except the emotion being depicted. Each reaction was practised several times by the male actor and female actor before the actual filming. When ultimately recorded, each reaction lasted about ten seconds, varying somewhat with the emotion being displayed. They appeared to be clear and unambiguous, and the camera work was precise.

Twelve emotional reactions were depicted twice by each actor and recorded on film. These are:

1.	affection	7.	shame
2.	enjoyment	8.	dislike
3.	happiness	9.	distress
4.	interest	10.	fear
5.	surprise	11.	sadness
6.	"no emotion" or neutrality	12.	anger

The finished film did not have an audio track, however, if an actor wished to verbalize while expressing any of the emotional reactions, he recited the alphabet in order to hold constant any cues gained by lip-reading.

Nine second segments of the expressive scenes were edited together so that each stimuli was introduced with a number and followed by a six second interval which would allow time to classify the emotional reaction depicted. The male actor was shown in the 19 odd-numbered stimuli, the female actor in the 19 even-numbered stimuli; a total of 38 scenes in all. The assembled film was shown to 10 subjects (5 male and 5 female) from the Subject Panel, Department of Experimental Psychology, University of Oxford. They were asked to generate labels for what they saw in each stimulus. For example, stimulus number one was seen as:

1. having an intense discussion
2. arguing with someone

3. being insulted
4. correcting someone who has misinterpreted something
5. making a point in a political argument, etc.

These subjects also classified the 38 stimulus items according to the usual 6-way classification of the Schlosberg circle shown in the inner ring of Figure 1. These results show that the degree of agreement depends on the type of emotion being judged (.85 for Happiness but only .45 for Fear). This is consistent with studies by Ekman _et al_. (1972) and Zuckerman _et al_. (1975).

Based on these results, three classification questionnaires were devised by someone blind to experimental hypotheses of studies that were to follow. Variation I was compiled from the label phrases generated by the subjects. The labels offered in this version may be fully appropriate, partially appropriate, partially inappropriate or fully inappropriate, and subjects in subsequent studies would indicate the extent of their agreement with these labels. Versions II and III offer binary classification choices for each emotional scene. The "modal" or "usual" adjectival descriptor was paired with another descriptor from either within the same Schlosberg circle segment (e.g. irate-displeased), from adjacent segments (e.g. overjoyed-surprised), or from non-adjacent segments (e.g. elated-sullen). A scoring scheme was made for each version so that the modal or typical classifications were weighted to be maximally discriminated above "chance" or "guess-rate". The test of emotion judgement is assessed according to normative perception or typical classification and does not take into account either actors' intended expressions or any absolute accuracy criteria. The studies which employ this technique will be looking for alterations or deviations in the usual classification of the filmed stimuli but not with a supposed reduction of perceptual accuracy.

In order to manipulate the Schlosberg dimensions in the present person perception film, an independent judge classified each stimulus item as belonging to one or the other extreme for each of the dimensions. The result is that each of the 38 stimuli has a three-way classification according to Schlosberg dimensions (see Table 1). As one might expect, within the 38 item test there is asymmetric overlap of these dimensions (e.g. 68% of the items are

Unpleasant and 16% are High Activity, attention and Unpleasant). It seems most appropriate, therefore, to treat each dichotomized dimension as a separate dependent variable in subsequent investigations.

Method

Subjects were randomly assigned to six experimental conditions, which corresponded to the bandwidth compression levels to be investigated. By means of a video signal processor, video tape recordings were produced so that the 38 item person perception film could be shown at 30 (i.e. full bandwidth), 15, 10, 5, 2 or 1 frame(s) per second. The duration of the stimulus presentation and the interstimulus interval remained the same in these six recordings. However, the rate of update of the information varied. Two examples will illustrate the perceptual result. If we consider the 5 fps condition, every sixth frame of full bandwidth video (i.e. 30 fps) was displayed for a fifth of a second so that a total of 45 frames were displayed during the nine second stimulus interval. In the 1 fps condition, every thirtieth frame of video was displayed for one second each so that a total of nine frames were displayed during the nine second stimulus interval.

In all conditions, subjects were seated alone at a node of the interactive video network. Each was given one of the three versions of the response questionnaire, which included instructions for the subject and a semantic differential with which to rate the actors. A video tape of one of the six experimental conditions was then played to the subject who responded by classifying the stimulus items. Upon completion of this task, the subject gave his "impressions of the personality" of each actor through semantic differential ratings. The order in which the actors were rated was counter-balanced. The experimental session was then terminated.

Results:

The modal perception of the emotional displays decreased as the bandwidth compression level increased. This was consistent with Hypothesis 1. Within each of the Schlosberg dimensions, it was always one of the dichotomized extremes which significantly differentiated the compression variable: Unpleasant, Attention and High Activity (see Table 2). Recall that the Unpleasant and High Activity indexes were predicted to be influenced more than their dimensional counterparts. Figure 2 shows that the mean scores on all dimensions fit a general trend of decreasing modal perception with increased bandwidth compression. Although the mean scores at 30 fps were slightly exceeded by mean scores at 15 fps, the difference was not reliable and was insufficient support for the notion that initial compression levels would enhance judgement consensus. On the other hand, there was a clear cut-off between the high and low consensus: 30 fps and 15 fps conditions were always significantly different from the 10 fps condition. In general, the compression levels that were 10 fps and beyond, yielded less modal perceptions of the emotional displays. Also, one should note that the 10 fps compression level produced strikingly lower scores which were even below those of the 5 fps and 2 fps conditions.

Table 3 shows the degree of overlap of the Schlosberg dimensions. Numbers above the diagonal indicate the degree of test item communality in the dichotomized dimensions. Note that the Unpleasant and Attention indexes were relatively independent of each other (32% overlap) but

Attention and High Activity indexes were relatively dependent (74% overlap). It is not surprising that, with the predicted influence upon the High Activity index, the Attention index should also be significantly affected as well. Below the diagonal, estimated shared variance between dimensional components is shown. These are based on the respective coefficients of determination calculated within subjects, without regard for the frame rate variable. The shared variances, while apparently quite high, cannot be meaningfully compared to each other because of the varying test item communality. For this reason, a communality ratio has been computed to show the statistical dependency, while taking into account the shared item dependency. Now it can been seen that the dependency between the High Activity and Unpleasant indexes is comparatively high, but between Attention and High Activity, dependency is comparatively low. A final point should be made with respect to Table 3. On the diagonal, one may note that a modest amount of variance was shared within the Schlosberg dimensional components. While the correlation coefficients were always positive, they were not particularly large (.15, .06 and .24). This would suggest that in the process of dichotomizing the Schlosberg dimensions, discrete indexes were created.

Having demonstrated both the logical and statistical overlap in the Attention, Unpleasant and High Activity indexes, it seems useful to represent in three-dimensional space the classification differences associated with the various frame rates (see Figure 3). Depicted in this way, it is obvious that the 30 fps and 15 fps conditions were set apart from the others and that the 10 fps condition interrupted the otherwise linear relationship in the frame rate variable.

Predictions with respect to Hypothesis 2 were partially supported. Figure 4 and Table 4 summarize strong evidence that the group consensus on the person perception task was lowest when the binary classification involved choices that were within one segment of the Schlosberg circle. It was expected that, since this judgement would be most difficult, it would best discriminate between the six frame rates. In fact, it discriminated <u>least</u> well because, even at the fullest bandwidth conditions, there was little agreement with the modal classification of the emotional displays.[2] This was true for both version II and III of the questionnaires.

Judgement choices which involved adjacent segments of the Schlosberg circle typically demonstrated a higher degree of consensual classification and the differences between this classification and the former within segment classification were significant and as predicted in nine of twelve instances, considering both questionnaires II and III. Again considering both questionnaires, performance on the adjacent classification task produced some differentiation within the six bandwidth compression levels. With version II of the questionnaire, it was simply that the 2 fps condition was different from the 1 fps condition. With version III, although the 2 fps condition was lower than might be expected, the general trend was as predicted. Moreover, it is in this analysis that the only significant finding was made which would support the notion that the full bandwidth condition would not yield the most modal judgements of the emotional displays. The 30 fps group scored significantly lower than the 15 fps group on the adjacent segment classification task in version III of the questionnaire. Elsewhere, when comparing the group means for these two groups, one can see trends in the same direction.

Hypothesis 2 also predicted that the non-adjacent segment classification would be the easiest to perform. The test of this assumption showed a significant effect with version II of the questionnaire and a marginally significant effect with version III. But for two inexplicable deviations, the groups who viewed the video tapes at a substantially reduced frame rate failed to achieve the same degree of normative consensus in their perceptions of the emotional displays as those groups with high frame rate. Groups with lower normative consensus did not show greater uncertainty in their judgements, nor did they show any greater degree of dissatisfaction with the classification choices offered for the labelling of the emotional displays. Table 5 shows the mean percentage of "uncertain" and "neither appropriate" responses given at each frame rate. These indexes produced clearly non-significant differences, suggesting that groups in the compressed bandwidth conditions were not overly perplexed by their person perception task. Rather, they perceived the emotional displays in different ways when assessed against the full bandwidth normative criteria.

In Hypothesis 3, it was predicted that an increase in the bandwidth compression would cause an increase in the perceived "artificiality" and "social distance" of the actors' personalities, as measured by the semantic differential. It was expected that degradation of the medium would generalize to the perception of the persons depicted in the medium. Table 6 shows that this predicted result was clearly non-significant. The actors were typically perceived as being just beyond the neutral mid-point of the artificiality component, or as slightly "trustworthy, genuine, natural and reliable". The actors were also seen to be moderately low on the social distance component of the semantic differential, regardless of the frame rate at which the video tape was viewed.

Another difference, which was predicted to be found in the semantic differential data, concerned the variability of impressions formed by the actors. Compared with that bandwidth compression level determined to yield most modal perceptions of the actors' emotional displays (i.e. 15 fps), both extremes of the compression continuum were predicted to show significantly greater intersubject variability of impression. This prediction was tested by computing group variance scores for each of the 21 items of the semantic differential. This was done separately to the ratings of the male and female actors for the 30 fps, 15 fps and 1 fps conditions. Paired t-tests were performed on these group variance scores with the expected result that the 15 fps group showed significantly less intersubject variability of perception of both actors than either of the other two groups (see Figure 5 and Table 7). This is taken to be support for the notion that moderate bandwidth compression can lead to a normalization of impressions formed during social perception.

Discussion:

We have seen strong evidence that the interpretation of emotional expressions can be altered by varying the rate of transmission of information to the human receiver. Psychologists concerned with comparability of results based on static and dynamic images may be correct in having some misgivings. As the video bandwidth was compressed by reducing frame rate, there was a widening in the range of possible interpretation for specific emotional displays. A greater variance in the impressions of the actors' personalities at the reduced bandwidth was also apparent.

A distinct cut-off point was identified for the effective communication of nonverbal cues; this occurred between 15 fps and 10 fps. In fact, the 10 fps rate produced unexpectedly low consensus of interpretation. Perhaps this sample rate created a perceptual flicker in the video display which was annoying or distracting. Further investigation of this point is required. Follow-up studies might investigate techniques to allow continuously varying interval selection rather than fixed interval selection. For example, a pseudo-random selection might be made with the mean interval set to one tenth of a second, but varying between one fifteenth and one fifth of a second. This technique, while maintaining the same information level, would avoid a possible perceptual annoyance associated with a particular frame rate. Another possibility may be to soften the transition from frame to frame by using a signal mixer to dissolve images. This might reduce abrupt changes that could distort the continuity of a sequence of behaviour. The ultimate bandwidth compression technique would undoubtedly contain all of these characteristics and, in addition, make the selection of frames contingent upon the amount of activity change occurring between frames. With this type of intelligent selection criterion, only novel or non-redundant information would need to be transmitted.

The attempt to manipulate the degree of ambiguity in stimulus items met with only partial success in the present study. It was conclusive that classifications involving emotional labels within one segment of the Schlosberg circle demonstrated lowest consensus of perception. Conversely, classifications involving emotional labels from non-adjacent segments demonstrated highest consensus of perception. Both of these findings were expected. However, it was anticipated that as the compression level increased, the scores on the more ambiguous task should decrease. This did not occur because all of the experimental groups found the within-segment classification task to be difficult. Discrimination of the frame rate conditions was found only with the adjacent and non-adjacent segment classification tasks. In these cases, the trend was as expected; the fuller bandwidth compression conditions (e.g. 30 fps and 15 fps) scored higher than the reduced bandwidth compression conditions (e.g. 2 fps and 1 fps).

The present study can again testify to the usefulness of the three dimensions of emotion identified by Schlosberg (1952 and 1954). Previous research (Thompson and Meltzer, 1964; Ekman et al., 1972; Zuckerman et al., 1975) advised that classification of unpleasant emotions were most difficult to perform. This led to the prediction that within the more complete bandwidth conditions the task would be performed better than in the less complete bandwidth conditions. The Unpleasant index did, in fact, discriminate the frame rate groups. Little should be made of the results that scores on the Unpleasant index, in most cases, actually exceeded the scores on the Pleasant index because no attempt was made to make the indexes meaningfully comparable.

Previous research into video bandwidth compression by Lindsey (1974) found that extreme compressed images were least acceptable when the video material contained a high degree of activity. The results of this study support such findings. High activity items in the test received greater consensual classification when there was little or no compression rather than when a large amount of bandwidth compression was involved. One can imagine that a superior compression technique would be one which could adapt to the content of video material to be transmitted.

The semantic differential ratings of the actors' personalities produced results that are of interest to both system designers and social psychologists. Firstly, there was not even a hint that the rather severely compromised medium which presented the actors was responsible for increasing perceived artificiality or social distance within the personalities of the actors. The medium and the messenger seemed to be assessed independently of each other. Secondly, there were more homogeneous impressions formed of the actors' personalities when the person perception film was viewed at 15 fps than at either of the extremes of the continuum, namely, 30 fps and 1 fps. Initial analysis, in the present study, identified the 15 fps condition as the maximally modal perception condition. That is to say, the 15 fps condition produced the highest mean scores on the person perception task when assessed against the criterion of typical perception.

This criterion was standardized at full video bandwidth. Consider that, not only did the 15 fps group show greatest intra-group agreement on what emotions were displayed, this group also formed the most consistent impressions of the actors' personalities. This result was predicted on the basis that a slightly diminished channel of communication would still contain the most salient cues from which to make accurate and consistent social perceptions yet would not contain the excessive subtle cues that might permit highly subjective and personal interpretations. If a very high fidelity channel of communication is used to make social perceptions (e.g. face-to-face or high quality colour film or full bandwidth video), it may be that the number of verbal cues available are overwhelming. In fact, the cues may be too numerous for all to be taken into account and too inconsistent to be reconciled into a single agreed-upon interpretation (Thayer and Schiff, 1969). It has been shown that increased information about an individual can generate irrelevant trait inferences about that person (Burron et al., 1971).

At the 1 fps end of the video bandwidth continuum, the point may be reached where too few social cues are present for consistent and accurate perceptions to be made. Argyle and McHenry (1971) found that a brief exposure to a photograph of an individual wearing spectacles led to more superficial inferences about high intelligence than when on-going behaviour of the same individual is presented on video tape. Undoubtedly, we are capable of drawing inferences based on extremely limited information, but doing so may result in judgement errors. Rosenthal et al., (1974) attempted to measure the accuracy threshold for nonverbal perception by systematically reducing the duration of a filmed stimulus display. Emotional scenes were reduced from two second duration to 1.13, .38, .13 and .04 seconds. One of their conclusions was as follows:

> "Apparently, people are able to assess nonverbal cues correctly in very short periods of time. Interestingly, those subjects who were more accurate at very brief exposure times (3/8th sec. or less) judged their interpersonal relationships as less satisfactory than those who were less accurate. That result suggests that our superfast nonverbal readers may be paying a price for their sensitivity; perhaps in social situations, they 'know too much'."
>
> (Rosenthal et al.,1974, p.3)

Many questions remain concerning the optimization of broadband video systems for person-to-person communication. Engineers and social psychologists alike might be teased by the results of the present study. It appears that, for certain social tasks, one half of the normal frame rate may be redundant; this should interest those concerned with cost-efficiency. Moreover, the human factors researcher should be intrigued by the prospect of communication experiences which facilitate high communality of social perception.

Footnotes

1. Schlosberg (1952 and 1954) found that facial expression of emotion could be classified according to three dimensions. These are:

   (a) Attention-Rejection
   (b) High-Low Activity (or Tension-Sleep)
   (c) Pleasantness-Unpleasantness

   (a) Attention-Rejection. "Attention is exemplified by surprise, in which all receptors are maximally open to stimulation. Rejection is the best term we have found for the other end of this axis; it is shown most clearly in contempt and disgust, in which eyes and nostrils appear to be actively shutting out stimulation." (Schlosberg, 1952, p.230)

   (b) High-Low Activity. "The term activate means a bit more than to make active; the dictionary tells us that it also means to make reactive. Activation would seem to be a very good name for what emotion does to us; the angry man overreacts to stimulation. Strong emotion thus represents one end of a continuum of activation; the other end, the condition of minimum activation, is found in the sleeping man who doesn't respond to stimulation." (Schlosberg, 1954, p.82) " ... general level of activation, has its low end in sleep, its middle ranges in alert attention, and its high end in the strong emotions". (Schlosberg, 1954, p.87)

(c) Pleasantness-Unpleasantness. " ... pleasantness-unpleasantness ... needs no further explanation." (Schlosberg, 1952, p.230) The enduring viability of these dimensions has been demonstrated in numerous studies which followed (e.g. Abelson and Sermat, 1962; Thorngate, 1971; Stringer, 1973).

2. The zero point on the mean score scale indicates "chance" or "guess rates".

Table 1

Classification of Test Items
According to Three Schlosberg Dimensions

Item Number	Attention-Rejection		Pleasant-Unpleasant		High-Low Activity	
	A	R	P	U	H	L
1	X			X	X	
2	X		X		X	
3		X		X		X
4		X		X	X	
5	X		X		X	
6		X		X		X
7		X		X	X	
8		X		X	X	
9		X		X		X
10		X		X	X	
11	X			X	X	
12		X		X		X
13		X		X		X
14	X		X		X	
15	X			X	X	
16		X		X		X
17	X			X	X	
18	X		X		X	
19		X		X		X
20		X		X		X
21	X		X			X
22		X		X	X	
23		X		X		X
24	X			X		X
25		X		X		X
26	X		X			X
27		X	X			X
28		X		X		X
29	X		X		X	
30		X		X		X
31	X			X	X	
32	X		X		X	
33	X		X			X
34		X		X	X	
35	X			X	X	
36	X		X		X	
37	X		X		X	
38		X		X		X

(Odd numbered items by Male actor; even by Female actor)

Table 2

Frame Rate		Means and Standard Deviations: 30	15	10	5	2	1	$\hat{\omega}^2$	Anova Summary Table: Total SS	Treat. SS	Error SS	Treat. MS	Error MS	F Ratio
Pl.	$\bar{X}$	1.10	1.13	.93	.99	1.00	.95	.00	17.28	.59	16.69	.12	.16	.72
	S	.38	.39	.44	.44	.38	.39							
Unpl.	$\bar{X}$	1.24	1.28	1.02	1.05	1.04	.91	.12	11.15	1.80	9.35	.36	.09	3.92**
	S	.28	.19	.32	.38	.33	.29							
Diff.	$\bar{X}$	.14	.10	.09	.11	.04	-.04		196.86	10.42	186.44	1.74	1.57	1.11
	S	.48	.51	.44	.53	.39	.57							
Rej.	$\bar{X}$	1.15	1.17	.98	1.08	1.07	.90	.03	11.97	.94	11.03	.19	.11	1.74
	S	.24	.30	.30	.36	.37	.38							
Att.	$\bar{X}$	1.31	1.31	.99	1.06	1.07	.94	.09	17.94	2.29	15.65	.46	.15	2.99*
	S	.50	.37	.44	.33	.37	.31							
Diff.	$\bar{X}$	.15	.14	.01	-.03	-.12	.03		27.14	.95	26.19	.19	.26	.74
	S	.54	.54	.46	.44	.46	.58							
Low	$\bar{X}$	1.16	1.20	1.00	1.11	1.07	.91	.03	13.76	1.06	12.70	.21	.12	1.71
	S	.32	.35	.32	.33	.43	.35							
High	$\bar{X}$	1.22	1.26	.97	1.03	.98	.93	.12	11.25	1.82	9.43	.36	.09	3.94**
	S	.33	.20	.43	.27	.31	.23							
Diff.	$\bar{X}$	.05	.06	-.04	-.08	-.10	.05		19.13	.47	18.66	.09	.18	.51
	S	.48	.40	.48	.26	.46	.44							

df = 5/102 $^*p < .05$ $^{**}p < .01$

Multiple Comparisons

Unpleasant: 30 ≠ 10,2,1.   15 ≠ 10,5,2,1.
Attention: 30 ≠ 10,1.   15 ≠ 10,1.
High Activity: 30 ≠ 10,2,1.   15 ≠ 10,5,2,1.

Table 3

Overlap of Dimension Test Items
and Proportion of Variance Shared
(Expressed as a Percentage)

	Attention	Unpleasant	High Act.
Attention	Rejection 2	32	74
Unpleasant	24	Pleasant <1	52
High Activity	46	50	Low Act. 6

Above Diagonal: test item communality
Diagonal: variance shared within dimensional components
Below Diagonal: variance shared between dimensional components

	Unpleasant-Attention	High Act.-Unpleasant	Attention-High Act.
R	$24/32 = .75$	$50/52 = .96$	$46/74 = .62$

Communality Ratio:

$$R = \frac{\%\ \text{shared variance (estimated)}}{\%\ \text{common items (actual)}}$$

Table 4

Multiple Comparisons for Schlosberg
Circle Analysis

Questionnaire II   SPF 6.3 Design, N = 60

a. Within one segment (F = 0.43, NSD)
b. Adjacent segments (F = 6.50, p < .01, 2 ≠ 1 fps)
c. Non-adjacent segments (F = 2.25, marginal)

30	15	10	5	2	1
a ≠ b	a ≠ b	a ≠ b	a ≠ b	NSD	NSD
b ≠ c	b ≠ c	b ≠ c	b ≠ c	NSD	b ≠ c
a ≠ c	a ≠ c	a ≠ c	a ≠ c	NSD	a ≠ c

Questionnaire III   SPF 6.3 Design, N = 36

a. Within one segment (F = 1.00, NSD)
b. Adjacent segments (F = 2.38, p < .05; 30 ≠ 15; 5 ≠ 2; 2 ≠ 1)
c. Non-adjacent segments (F = 4.26, p < .01; 15 ≠ 10; 5 ≠ 2; 2 ≠ 1)

30	15	10	5	2	1
a ≠ b	a ≠ b	NSD	a ≠ b	a ≠ b	a ≠ b
b ≠ c	NSD	NSD	NSD	b ≠ c	b ≠ c
a ≠ c	a ≠ c	NSD	a ≠ c	a ≠ c	a ≠ c

Table 5

Frame Rate	30	15	10	5	2	1
Percentage of "uncertain" responses (i.e. ???)	5	5	6	7	4	7
Percentage of "neither appropriate" responses* (i.e. n/a)	21	23	19	17	19	24
Percentage of responses remaining	74	72	75	76	77	69

* The n/a response was expected to occur approximately 22% of the time.

Table 6

Semantic Differential Scale Components Expected to Discriminate Compression Levels

I. "Artificiality" Scale: (1 = artificial 7 = geniune)

Encoder		30	15	10	5	2	1
Male	X̄	4.64	4.80	4.59	4.83	4.84	4.86
	S	1.12	1.08	1.05	1.13	1.16	1.31
Female	X̄	4.98	4.27	4.66	4.65	4.75	4.27
	S	1.05	1.18	.93	1.11	1.19	1.07

Anova Summary Table

	Source	SS	df	MS	F	
Male	Treatment	1.52	5	.30	.23	(NSD)
	Error	172.53	132	1.31		
	Total	174.05	137			
Female	Treatment	8.92	5	1.78	1.50	(NSD)
	Error	157.01	132	1.19		
	Total	165.93	137			

II. "Social Distance" Scale: (1 = low SD 7 = high SD)

Encoder		30	15	10	5	2	1
Male	X̄	3.08	2.87	3.15	3.04	2.92	2.89
	S	1.01	.98	.95	.82	.90	.96
Female	X̄	3.35	3.44	3.32	3.50	3.47	3.78
	S	1.18	.85	.92	.91	.96	1.10

Anova Summary Table

	Source	SS	df	MS	F	
Male	Treatment	1.50	5	.30	.34	(NSD)
	Error	116.28	132	.88		
	Total	117.78	137			
Female	Treatment	3.14	5	.63	.64	(NSD)
	Error	130.44	132	.99		
	Total	133.57	137			

Table 7

Overall Mean Score on Semantic Differential Items

Frame Rate	30	15	10	5	2	1
For Male Actor	2.72	2.50	2.77	2.67	2.51	2.53
For Female Actor	2.78	3.04	2.86	3.00	2.89	3.27

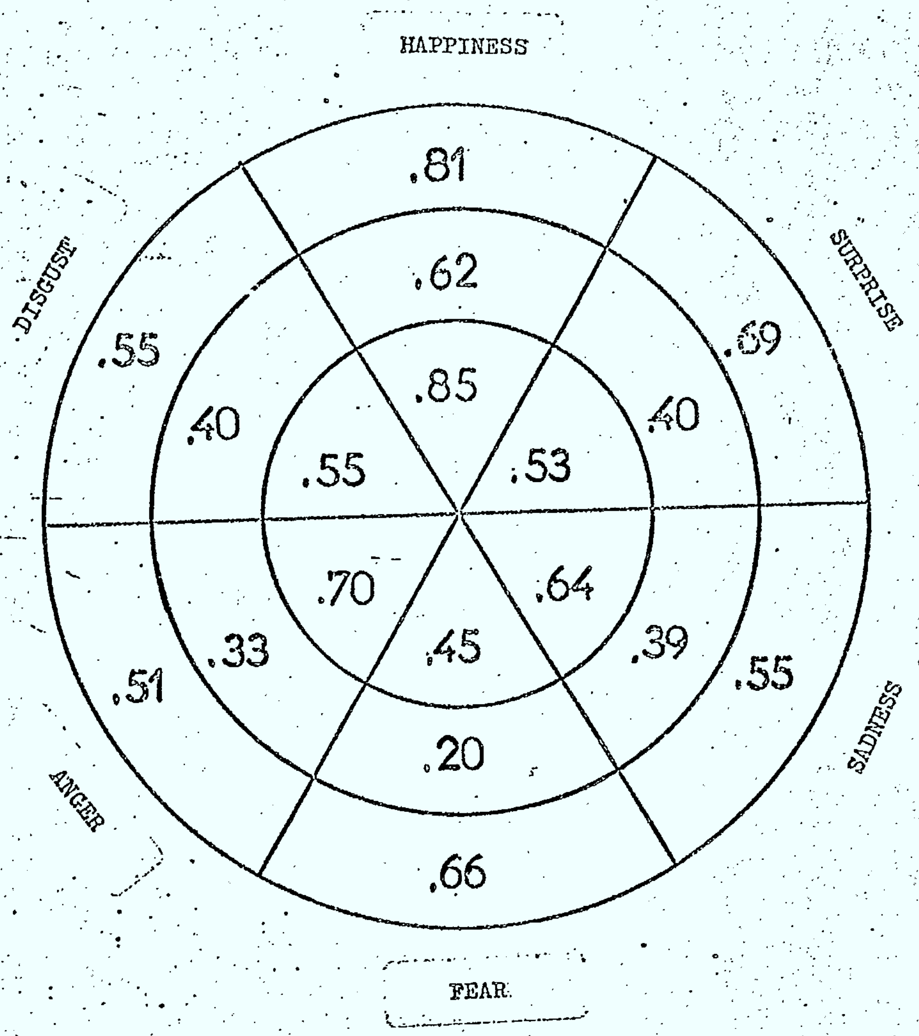
(1 = positive and 7 = negative)

Overall Variance Score on Semantic Differential Items

Frame Rate	30	15	10	5	2	1
For Male Actor	2.24	1.70	2.17	2.06	1.79	2.01
For Female Actor	2.14	1.78	1.83	2.09	1.99	2.26

Figure 1

Mean Percentage Accuracy of
Sending and Receiving Emotion

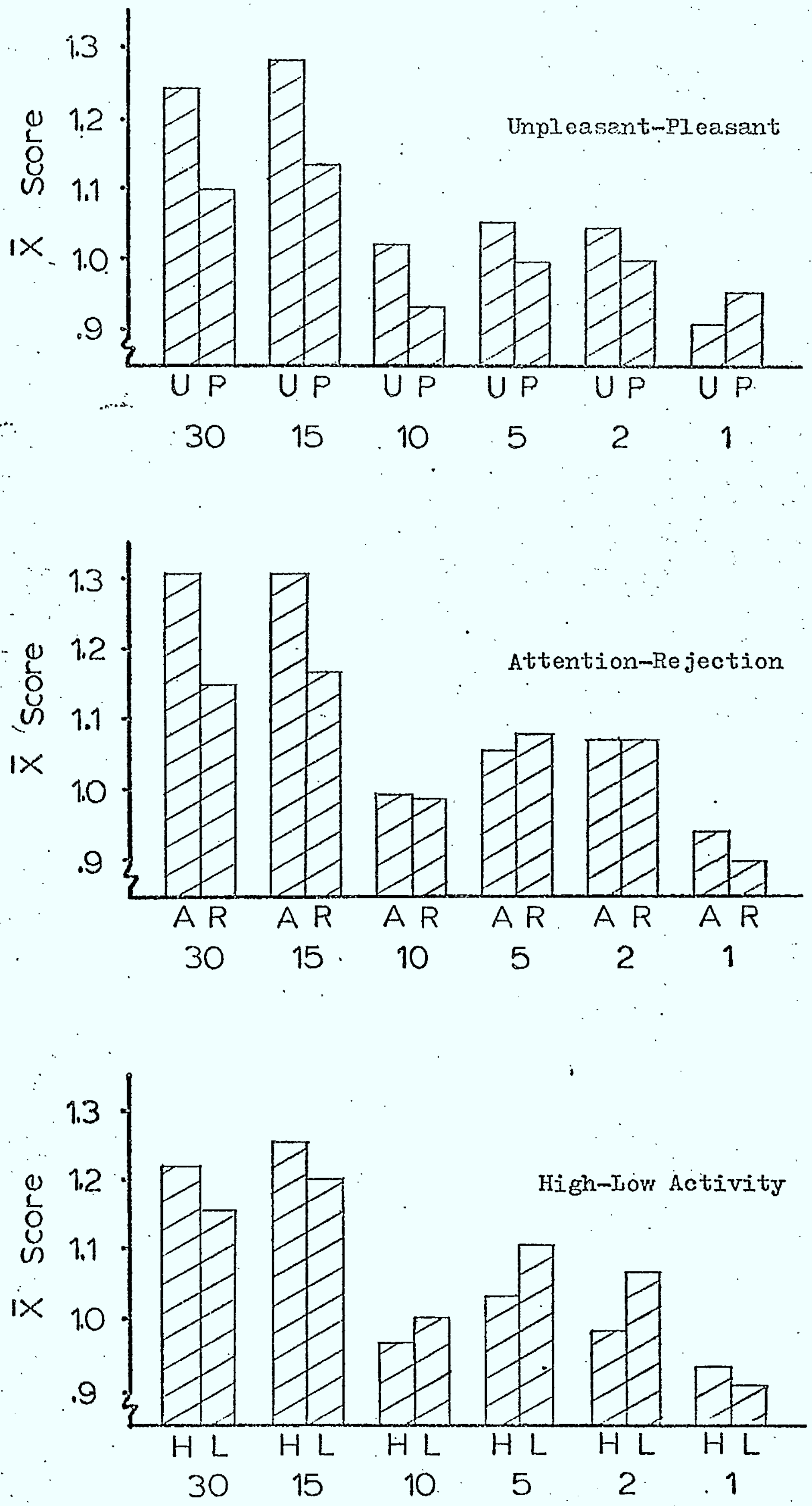


Outer ring: Summary of results of nine studies on accuracy sending by Ekman et al, 1972.

Centre ring: Results of study on accuracy sending by Zuckerman et al, 1975.

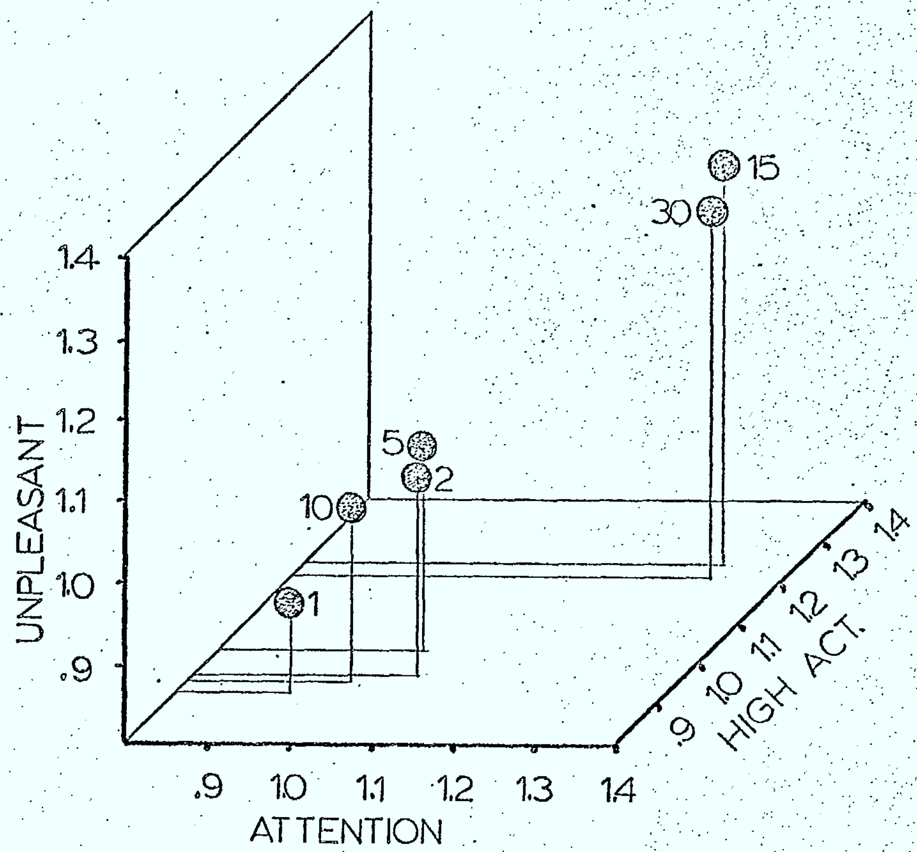
Inner ring: Results of present study of consistency receiving.

Figure 2



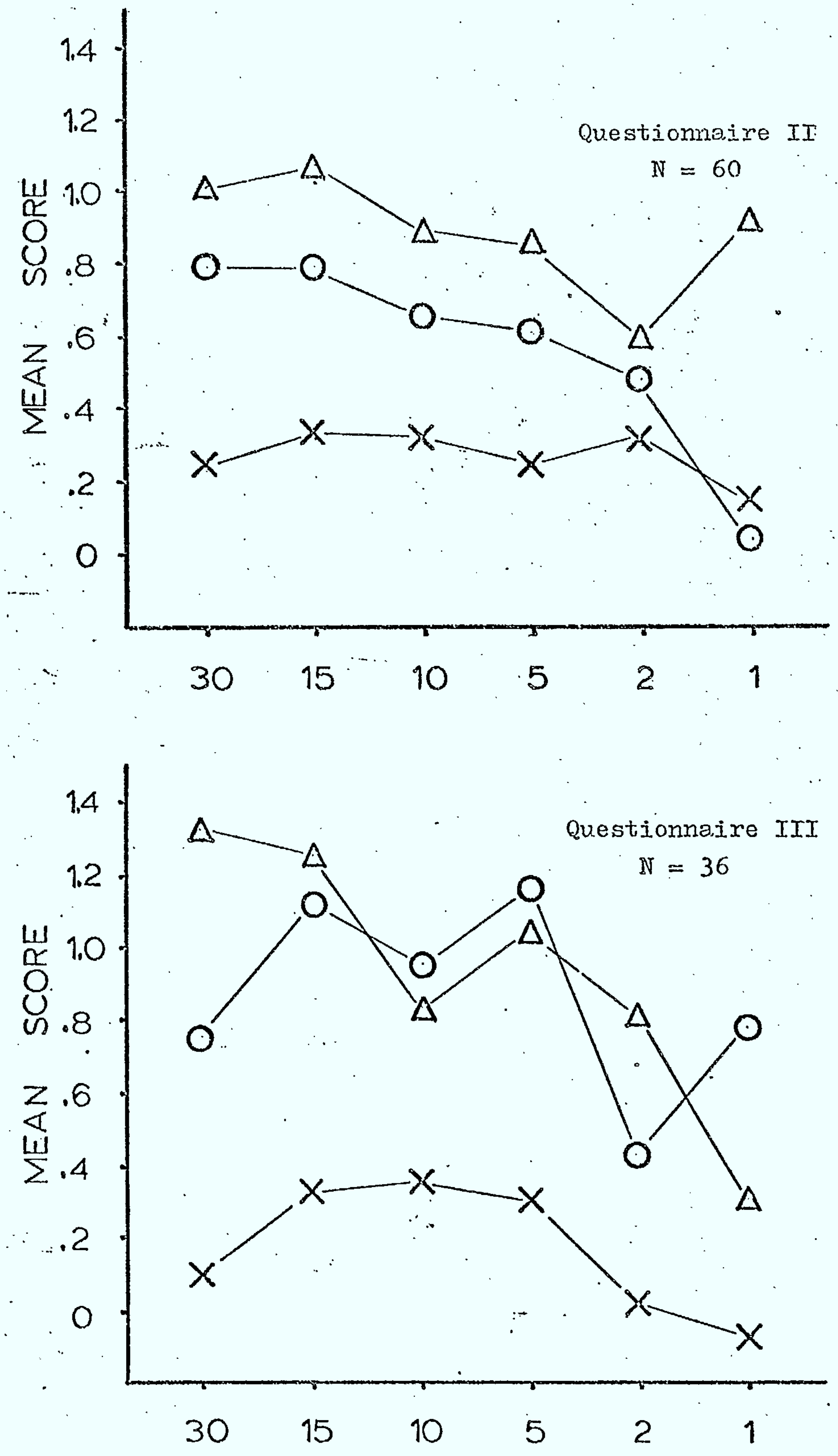
Three Schlosberg Dimensions
Viewed at Six Frame Rates

Figure 3



Bandwidth Compression Levels Discriminated
in Three Dimensional Space

Figure 4

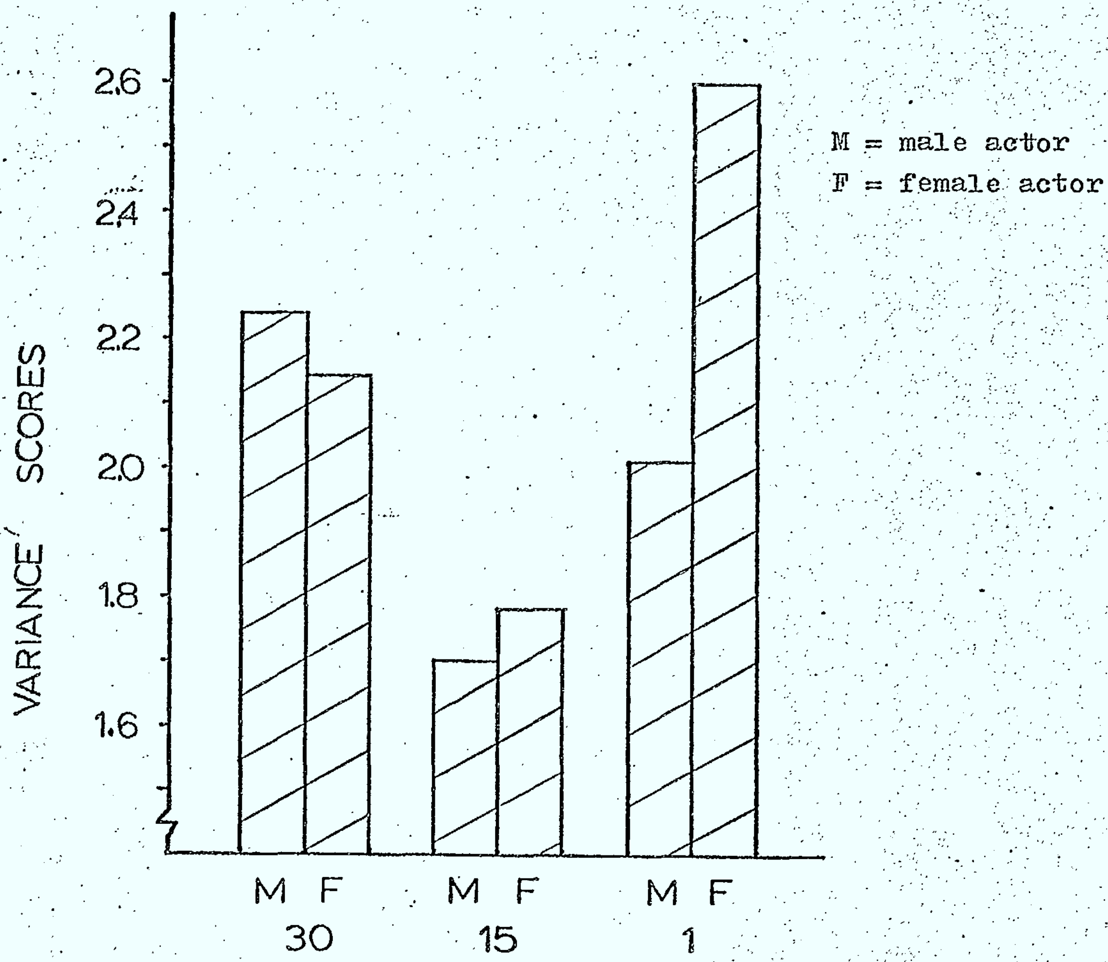


Segments of Schlosberg Circle:

X within one   O adjacent   Δ non-adjacent

(0 = chance or guess rate)

Figure 5



Mean Variable Scores on the Semantic Differential Ratings of Actors' Personalities: Comparisons between the Maximally Modal Perception Condition (15 fps) with the Conditions at the Extremes of the Frame Rate Continuum (30 fps and 1 fps)

References

Abelson, R.P. and Sermat, V. Multidimensional scaling of facial expressions. Journal of Experimental Psychology, 1962, 63, 546-554.

Argyle, M. and McHenry, R. Do spectacles really affect judgements of intelligence? British Journal of Social and Clinical Psychology, 1971, 10, 27-29.

Burron, B.F., Carlson, K.A., Getty, G.R. and Jackson, D.N. The effects of informational characteristics on the perception of real and hypothetical target persons. Psychonomic Science, 1971, 23 (2), 145-147.

Dickson, E. and Bowers, R. The Videotelephone: A new era in telecommunications. Report to the National Science Foundation from Cornell University, New York.

Ekman, P. and Friesen, W.V. Non-verbal leakage and clues to deception. Psychiatry, 1969, 32, 88-105.

Ekman, P., Friesen, W.V. and Ellsworth, P. Emotion in the Human Face: Guidelines for Research and an Integration of Findings. New York: Pergammon Press, 1972.

Frijda, N.H. Recognition of emotion. In L. Berkowitz (ed.) Advances in Experimental Social Psychology (4). New York: Academic Press, 1969.

Graham, J.A., Bitti, P.R. and Argyle, M. A cross-cultural study of the communication of emotion by facial and gestural cues. Journal of Human Movement Studies, 1975, 1, 68-77.

Kirk, R.E. Experimental Design: Procedures for the Behavioural Sciences. Belmont, Calif.: Brooks/Cole Publishing Company, 1968.

Lindsey, G. A method for subjective evaluation of video-compressed television images. Preliminary draft, Institute for Communication Research, Stanford University, California, August 28, 1974.

McManamon, P. Technical implications of teleconference service. I.E.E.E. Transactions on Communications, 1975, COM-23 (1), 30-38.

Osgood, C.E. Dimensionality of the semantic space for communication via facial expressions. Scandinavian Journal of Psychology, 1966, 7, 1-30.

Ried, A.A.L. Channel versus system innovation in person/

person telecommunications. Human Factors, 1973, 15 (5), 449-457.

Rosenthal, R., Archer, D., Koivumaki, J.H., DiMatteo, M.R. and Rogers, P.L. Assessing sensitivity to nonverbal communication: The PONS test. Division 8 Newsletter, Division of Personality and Social Psychology, January, 1974.

Schlosberg, H. The description of facial expressions in terms of two dimensions. Journal of Experimental Psychology, 1952, 44, 229-237.

Schlosberg, H. Three dimensions of emotion. Psychological Review, 1954, 61, 81-88.

Stringer, P.H. Do dimensions have face validity? In: M. von Cranach and I.Vine (eds.) Social Communication and Movement. London: Academic Press, 1973.

Thayer, S. and Schiff, W. Stimulus factors in observer judgement of social interaction: Facial expression and motion pattern. American Journal of Psychology, 1969, 82 (1).

Thompson, D.F. and Meltzer, L. Communication of emotional intent by facial expression. Journal of Abnormal and Social Psychology, 1964, 68, 129-135.

Thorngate, W. Facial cues and emotion inferences. Technical Report n. 71-2, Social Psychology Laboratory, Department of Psychology, University of Alberta, Edmonton, August, 1971.

Zuckerman, M., Lipets, M.S., Koivumaki, J.H. and Rosenthal, R. Encoding and decoding nonverbal cues of emotion. Journal of Personality and Social Psychology, 1975, 32 (6), 1068-1076.

APPENDIX

THE APPENDIX OF THIS ANNEX IS APPENDIX IV OF "FINAL REPORT", DECEMBER 1976, DOC CONTRACT # OSU5-0072.

RESEARCH INTO MORE EFFICIENT USE OF STANDARD TELEVISION AND TELEPHONE CHANNELS IN EDUCATIONAL COMMUNICATION SYSTEMS: ANNEX TO FINAL REPORT. -- Guild, P.D.