Abstract—This study investigates the effect of global signal characteristics, such as bandwidth and number of events on human sampling behavior. It also examines whether the influence of these characteristics on sampling strategy is affected by local (actually observed) signal features, such as the degree with which a sampled signal value falls short of an event region and the rate of change. In the present study four independent, numerically displayed signals were used; two different bandwidths and two different event numbers were chosen. To take a sample, subjects had to use a mouse. The mouse key responses were used as an index of sampling. The study demonstrates that both bandwidth and number of events equally affected the distribution of samples over signals. In addition, it shows that global signal characteristics determine sampling behavior less prominently when the attentional demands brought about by the local signal features become of greater importance. This indicates that not predictability as governed by global signal characteristics as such, but rather predictability given certain local signal features is a crucial factor in determining sampling behavior.

I. INTRODUCTION

Since the 1960's, a number of analytic models of human sampling behavior in dynamic environments has been developed (see [14], [15] for detailed reviews). Putting together the most characteristic features of these models, Moray [14], [15] proposes a canonical model which can be summarized as follows.

1) The human “internal model” of the dynamic environment and its knowledge of the last observed signal values underlie monitoring behavior.
2) Statistical characteristics of the signals (e.g., the number of critical events per signal) produce uncertainty that humans wish to reduce. Attention allocation will be a function of the relation of the magnitude to the value which must not be exceeded (the “critical area limit”) in tolerance band monitoring.

3) Sampling behavior is sensitive to the dynamic properties of the signals (e.g., the signal bandwidth), the cost and pay-offs associated with the sampling of the different signals, and the probability matrices relating the value of the different intervals.
4) Other signal factors determining sampling behavior are: correlation and causal relations among signals, the relevance of signals and the signal to noise ratio.

Sampling models are useful for the investigation of supervisory control, presently the most important operator task in the process industry. One of the basic activities in supervisory control is tolerance band monitoring [5], [21]. In such a task the operator has to notice whether a process stays within the limits of a tolerance band. It is not necessary to “reconstruct” a process completely.

In an earlier experiment, Bohnen and Leermakers [3] investigated the influence of signal bandwidth in a tolerance band monitoring task. Bandwidths of 0.03, 0.06, and 0.09 Hz were used, the number of events being equal for all signals. By the “number of events” we mean the number of times the signal exceeds the tolerance band during the period of observation. The experiment demonstrated that bandwidth as such provided enough information to discriminate among signals. Notwithstanding the absence of the number of events as a discriminating factor, the sampling frequency tended to increase linearly with an increase in bandwidth. This research was in line with Senders’ models and experiments [18], [19] although a different method for data acquisition was used. Senders used eye movement recordings, whereas Bohnen and Leermakers used the observing response method.

One of the purposes of the present study is to investigate the effects of the global signal characteristics (i.e., bandwidth as well as number of events) on sampling frequency. It is expected that sampling frequency will be a function of both bandwidth and number of events. Pilot experiments suggested domination of bandwidth effects over event effects.

In the present study, subjects have to ask for (to sample) a signal. The duration of each sample is one second. During such a sample the signal state, consisting of the actual value and an indication of the first derivative (rate of change indication) is presented numerically. This way of presentation is quite normal, for example in the process industry where it is common to present signals numerically.
Bohnen and Leermakers [3] also showed an influence of local (actually observed) signal features on sampling behavior: the degree with which a signal value falls short of a limit value (of the tolerance band), and the rate of change. The present study also examines whether the influence of the global signal characteristics (bandwidth and events) on sampling strategy is systematically affected by the sampled local signal features. The question then is, whether sampling behavior primarily is determined by the predictability as governed by global signal characteristics, or by predictability given local signal features.

As already mentioned, the observing response method is used [9]. Displays are blank unless subjects ask for a sample by pressing a mouse key. This means that key responses are used as an index of sampling. The observing response paradigm has proved to be a successful experimental technique for studying the distribution of samples (see e.g., [3], [5], [7], [10], [12]); but it should be noted that Holland [9] and Haga & Moray [8] indicate that the results of an experiment are dependent upon the chosen data acquisition technique, and this should be taken into account when comparing results of experiments using different techniques.

II. METHOD

A. Subjects

Twelve male university students, among 19 and 24 years of age, participated in the experiment. Subjects had no prior experience in monitoring experiments. To achieve an appropriate level of motivation, subjects were told that the subject scoring lowest would receive DFL. 115,-, and that the remuneration of the remaining subjects was to be based on a linear relation between these two rewards.

B. Signals

The signals were made by filtering Gaussian distributed series of numbers with rectangular low-pass filters with cut-off frequencies of 0.03 or 0.09 Hz. All series contained 1800 numbers. Every number in a series represented the value of a signal during one second, which means that the total series duration was 1800 s.

Signals filtered with the filter from 0.03 Hz contained only frequencies in the range from 0 to 0.03 Hz. This range is defined as the bandwidth of the signal. As a function of time, there were fewer fluctuations for the 0.03 Hz (Slow) than for the 0.09 Hz (Fast) signals (see Fig. 1(a) and (b)).

Four types of signals were made: a signal of 0.03 Hz with 12 events (S12), a signal of 0.03 Hz with 36 events (S36), a signal of 0.09 Hz with 12 events (F12), and a signal of 0.09 Hz with 36 events (F36). The sum of these bandwidths was 0.24 Hz and the total number of events 96. Pilot results and earlier experiments [3] showed that these signals could keep the subjects reasonably active in such a task situation.

The standard deviation of each series was adjusted to obtain the desired number of events (12 or 36). Signal values could vary between about −50.0 and +50.0 for the 12 events signals and between about −75.0 and +75.0 for the 36 events signals.

Two parts of this range were defined as event regions: the region ≤−25.0 and the region ≥+25.0.

An event started if a signal entered such a region and ended if the signal left the region. So, an event is the period of time one of the four signals actually spends in the event region (see Fig. 1). e1 is an event with a duration of 11 s, and e2 is an event with a duration of 4 s. This means that an 12 events signal spends 12 periods of time within an event region and an 36 events signal, 36 periods of time.

When a sample of one of the four signals was asked for, the value of the particular signal was numerically displayed on the screen (see Fig. 2, e.g., +19.0; −0.5; −9.4) and below the signal value (height: 0.25º of arc) the rate of change of the signal was presented (0.14º of arc). This rate of change gave an indication of the difference (Δ) between the actual value of a signal and the signal value one s before the presentation and could consist of a number of plus or minus signs. A plus sign indicated an increase, a minus sign indicated a decrease. Six rate of change indications were possible: ++++(Δ > 2.0), +++(1.0 < Δ ≤ 2.0), +(0.0 ≤ Δ ≤ 1.0), −(−1.0 ≤ Δ < 0.0), ---(−2.0 ≤ Δ < −1.0), ----(Δ < −2.0). For example, if on the 95th s a signal value is +30.0 and on the 94th seconds +26.3 then Δ is 3.7 and the rate of change indication on the 95th second would be +++. For example −−− means that a signal is rapidly going more negative.

For each signal type six series were filtered. The average percentage of time (over the six series of 30 min) each signal type was in a certain signal region was calculated (see Table I). The average duration of an event was 10.7 s for the S12, 12.8 s for the S36, 3.1 s for the F12, and 3.9 s for the F36 signal.
The purpose of the practice trials was to give subjects the opportunity to learn the dynamics of the signals in a more easy way.

To take a sample of one of the four signals, subjects had to ask for it by using a mouse (IBM PS2: model 6450350) to move a cursor into the window of that particular signal, followed by pressing the left mouse key. Pressing the left mouse key led to the one's presentation (so, a fixed time) of the signal value and rate of change. To force serial sampling, sampling was possible for only one signal at a time. For each new sample, subjects first had to move the cursor into a rectangle in the middle of the screen (3.20 by 2.60° of arc), to hamper the development of fixed sampling sequences based on the ease in making mouse movements. This rectangle disappeared when it was touched with the cursor and reappeared when subjects pressed the left mouse key at the moment that the cursor was in one of the four windows. In order to see sequences of values (and rate of changes) for one of the four signals, subjects have to sample the particular signal successively.

When subjects observed an event, they had to detect it as soon as possible by pressing the right mouse key. A detection was correct if it was given from the moment an event started until 5 s after the event ended, otherwise it was defined as incorrect. A correct detection was followed by a pitch of 880 Hz and an incorrect detection by a pitch of 440 Hz. If a signal was sampled during an event and this event had already been detected, an asterisk (*) was presented to the right of the signal value, to reduce the probability of detecting the same event more than once.

A pay-off score system was used to prevent continuous sampling with no or only little reasoning about future signal values. Subjects were paid according to this system (see Section IIA). At the beginning of an experimental trial the score was set to 1400 points. For each sample asked for it was diminished by 1 point, for each correct detection subjects received 15 points, the cost of a missed event was 15 points and of an incorrect detection 5 points. The ultimate number of points was divided by 200. So, the score = \[ \frac{1400 - NS + 15(CE) - 15(ME) - 5(ID)}{200} \]. (NS = number of samples; CE = correct detected events; ME = missed events; ID = incorrect detections). Suppose in one session, a subject asked for 650 samples, detected 70 events correct, missed 26 events and made one incorrect detection. The resultant score would be: \[ \frac{1400 - 650 + 70(15) - 26(15) - 5(1)}{200} = 7.0 \]. Experience derived from pilot experiments was used to establish the structure of this score, it was adjusted to the subjects sampling behavior so that the score used at the actual experiment delivered scores between about 0 and 10. This was considered as an acceptable range because it was in consonance with the grading system used in the Netherlands.

The variables used to measure sampling behavior were:

1) The total number of samples, for a general description over the sessions.

2) The inter-sample-interval (ISI) that is, the time in seconds between a sample for a certain signal and the first following sample for the same signal. ISI's were measured from each previous sampling value because it was assumed that the timing of the following sample for

<table>
<thead>
<tr>
<th>Signal region</th>
<th>Signal Value (SV)</th>
<th>S12</th>
<th>S36</th>
<th>F12</th>
<th>F36</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>0 ≤ SV &lt; 5.0</td>
<td>28.48%</td>
<td>17.41%</td>
<td>36.39%</td>
<td>27.89%</td>
</tr>
<tr>
<td>II</td>
<td>5.0 ≤ SV &lt; 10.0</td>
<td>28.65%</td>
<td>17.92%</td>
<td>29.52%</td>
<td>25.00%</td>
</tr>
<tr>
<td>III</td>
<td>10.0 ≤ SV &lt; 15.0</td>
<td>19.83%</td>
<td>17.49%</td>
<td>18.13%</td>
<td>18.34%</td>
</tr>
<tr>
<td>IV</td>
<td>15.0 ≤ SV &lt; 20.0</td>
<td>12.37%</td>
<td>12.96%</td>
<td>9.93%</td>
<td>12.98%</td>
</tr>
<tr>
<td>V</td>
<td>20.0 ≤ SV &lt; 25.0</td>
<td>7.57%</td>
<td>11.66%</td>
<td>3.98%</td>
<td>7.98%</td>
</tr>
<tr>
<td>VI</td>
<td>SV ≥ 25.0</td>
<td>7.16%</td>
<td>25.66%</td>
<td>2.07%</td>
<td>7.81%</td>
</tr>
</tbody>
</table>

**TABLE II**

The average percentage of time for a certain magnitude of the rate of change was calculated for each signal type and summarized over separate minus and plus signs (see Table II).

For the practice trials, extra signal series were made with features similar to those of the experimental trials.

**C. Task**

Subjects were seated in front of a table in a dimly-lit, sound-attenuating room, facing a video screen from a distance of about 95 cm. The experiment was carried out with the aid of an IBM PS2 Model 50 Z computer and monitor for the presentation of the signal samples, asked for by subjects. Each experimental trial the observing response method was used.

The average percentage of time for a certain magnitude of the rate of change for each signal type was calculated (see Table II).

For each new sample, subjects first had to move the cursor into a rectangle in the middle of the screen (3.20 by 2.60° of arc), to hamper the development of fixed sampling sequences based on the ease in making mouse movements. This rectangle disappeared when it was touched with the cursor and reappeared when subjects pressed the left mouse key at the moment that the cursor was in one of the four windows. In order to see sequences of values (and rate of changes) for one of the four signals, subjects have to sample the particular signal successively.

When subjects observed an event, they had to detect it as soon as possible by pressing the right mouse key. A detection was correct if it was given from the moment an event started until 5 s after the event ended, otherwise it was defined as incorrect. A correct detection was followed by a pitch of 880 Hz and an incorrect detection by a pitch of 440 Hz. If a signal was sampled during an event and this event had already been detected, an asterisk (*) was presented to the right of the signal value, to reduce the probability of detecting the same event more than once.

A pay-off score system was used to prevent continuous sampling with no or only little reasoning about future signal values. Subjects were paid according to this system (see Section IIA). At the beginning of an experimental trial the score was set to 1400 points. For each sample asked for it was diminished by 1 point, for each correct detection subjects received 15 points, the cost of a missed event was 15 points and of an incorrect detection 5 points. The ultimate number of points was divided by 200. So, the score = \[ \frac{1400 - NS + 15(CE) - 15(ME) - 5(ID)}{200} \]. (NS = number of samples; CE = correct detected events; ME = missed events; ID = incorrect detections). Suppose in one session, a subject asked for 650 samples, detected 70 events correct, missed 26 events and made one incorrect detection. The resultant score would be: \[ \frac{1400 - 650 + 70(15) - 26(15) - 5(1)}{200} = 7.0 \]. Experience derived from pilot experiments was used to establish the structure of this score, it was adjusted to the subjects sampling behavior so that the score used at the actual experiment delivered scores between about 0 and 10. This was considered as an acceptable range because it was in consonance with the grading system used in the Netherlands.

The variables used to measure sampling behavior were:

1) The total number of samples, for a general description over the sessions.

2) The inter-sample-interval (ISI) that is, the time in seconds between a sample for a certain signal and the first following sample for the same signal. ISI's were measured from each previous sampling value because it was assumed that the timing of the following sample for
the same signal is related to the last sampled value of that signal. For example (see Fig. 1(a)), in the case of a signal value at the moment of sampling of $-17.6$ (s1) and given the next sample (s3) for the same signal (S12) $26$ s afterwards, an IS1 of $26$ s was associated with the signal value of $-17.6$. Within the $26$ s between the first (s1) and second (s3) sample for the S12 signal, a sample for the other signals could be asked for too (see e.g., s2 in Fig. 1(b)).

The variables used to measure the accuracy of performance were: the percentage of events for each signal that was detected correct (PCD), the number of incorrect detections and the score (calculated from the cost of sampling, the value of detections etc.).

D. Procedure

A within-subjects design was used with signal bandwidth (0.03 and 0.09 Hz) and number of events (12 and 36) as the main independent variables, affecting the sampling strategy of the subjects. Due to practical considerations only six combinations of signal series were chosen. One subject never received the same combination in consecutive sessions and two simultaneously tested subjects always received different combinations.

For the 12 subjects, the distribution of the different signal types (S12, S36, F12, and F36) over the windows on the screen (see Fig. 2) was determined with the help of a 'quasi' random selection from the 24 possible combinations of distributions. However, for each individual subject a particular signal type appeared during all trials in the same window.

Subjects participated on three consecutive days. On the first day performance was recorded in four sessions (the first was used to familiarize subjects with the task), while on the following two days three sessions were done. This resulted in a total of ten sessions for each subject. The first quarter of each session included the practice trial and the following half an h the experimental trial. On the first day sessions were run from 8.30 am until 5.30 pm, the following days sessions ended at 2.30 pm. Each day two subjects participated; the first subject started at 8.30 am and during this period the other subject rested. This schedule was continued with periods of one h (the session time plus 15 min for saving data etc.) till both subjects had completed all sessions.

Before the preliminary session of the first day subjects received instruction during which it was pointed out that the signals could differ in predictability and number of events. Information concerning the number of events was not given. They were told that after each practice and experimental trial, they would obtain knowledge of results about the percentage missed events. After each experimental trial this percentage was given together with the number of samples for all signals and the score that could be seen as an indication for the level of performance.

E. Methodological Considerations Concerning Data Analyzes

The ISI was necessary to determine whether sampling behavior was contingent on the observed local signal features (i.e., to examine the relation between sampling and actually observed signal values). Since subjects were more or less free to determine the moments of sampling, and because of the different percentages of time a signal type was in a certain signal region (see Table I) and the difference in the distribution of the rate of change indications over different signals (see Table II), no fully balanced design was possible. To correct for these unbalanced nature, unweighted marginal ISI means [13] are presented in the figures.

For the ISI, two analyses of variance (ANOVA's) for unbalanced designs were conducted with the help of the SAS version 6.03 GLM procedure, using the type III estimable function [17]. The reason for two ANOVA's was the lack of computational resources. An ANOVA considering the complete model could not be conducted [17, pp. 611-612]. In general only lower-order interactions were considered, meaning that effects of higher-order interactions were included in the mean square error.

The signals were more or less symmetrical around zero and sampling behavior seemed to be identical for negative and positive signal values. Therefore for the ANOVA's on the ISI's negative and positive regions were combined, for example, the samples in the region $\leq -25.0$ were combined to those in the region $\geq + 25.0$ etc. The range of the signal was then divided into six signal regions (see Table I). For a signal with a value of $+23.4$, for example, a rate of change indication of $+++$ means that it was (rapidly) moving toward the nearest event region limit, whereas for a signal with a value of $-24.1$, this indication means that it was (rapidly) moving away from the nearest event region limit. For ISI analyzes, therefore the rate of change indication was divided into two factors:

1) A trend, for the direction of the changes, with two levels (\$\uparrow\$, indicating a change in the direction of the nearest event region limit, and \$\downarrow\$, indicating a change away from this limit).

2) A trend magnitude, for the size of the changes, with three levels (\$b\$ for the \$--\$ and \$++\$ symbols, \$bb\$ for the \$--\$ and \$++\$ symbols, and \$bbb\$ for the \$--\$ and \$++\$ symbols), indicating small, intermediate and large absolute rate of changes. For example, the signal value of $+23.4$ with the rate of change indication of $+++$ was divided into a \$\uparrow\$ trend and a \$bbb\$ trend magnitude and the value of $-24.1$ with the same rate indication was divided into a \$\downarrow\$ trend and a \$bbb\$ trend magnitude.

III. RESULTS

A. Quantitative Analyzes

Subjects performed the task on three consecutive days (see Section II). Training effects occurred during those days. In Table III the effect of training on the score (SCO), the ISI and the PCD is shown. These effects were more obvious between the first and the second day than between the second and the third day. For example for the score (SCO) the difference between the first and second day was 1.1 and between the
TABLE III
THE SCO (SCORE), THE ISI AND THE PCD AS A FUNCTION OF DAY

<table>
<thead>
<tr>
<th>Variable</th>
<th>Day 1</th>
<th>Day 2</th>
<th>Day 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCO 4 signals together</td>
<td>5.8</td>
<td>6.9</td>
<td>7.1</td>
</tr>
<tr>
<td>PCD 4 signals together</td>
<td>76.5%</td>
<td>79.2%</td>
<td>80.1%</td>
</tr>
<tr>
<td>PCD S12 signal</td>
<td>93.4%</td>
<td>92.3%</td>
<td>93.1%</td>
</tr>
<tr>
<td>PCD S36 signal</td>
<td>96.3%</td>
<td>97.3%</td>
<td>97.5%</td>
</tr>
<tr>
<td>PCD F12 signal</td>
<td>51.4%</td>
<td>57.5%</td>
<td>58.1%</td>
</tr>
<tr>
<td>PCD F36 signal</td>
<td>64.4%</td>
<td>69.6%</td>
<td>73.9%</td>
</tr>
<tr>
<td>ISI 4 signals together</td>
<td>8.7 s</td>
<td>9.1 s</td>
<td>9.2 s</td>
</tr>
<tr>
<td>ISI S12 signal</td>
<td>9.8 s</td>
<td>10.6 s</td>
<td>11.2 s</td>
</tr>
<tr>
<td>ISI S36 signal</td>
<td>8.3 s</td>
<td>9.2 s</td>
<td>9.3 s</td>
</tr>
<tr>
<td>ISI F12 signal</td>
<td>9.0 s</td>
<td>9.2 s</td>
<td>9.1 s</td>
</tr>
<tr>
<td>ISI F36 signal</td>
<td>7.8 s</td>
<td>7.5 s</td>
<td>7.3 s</td>
</tr>
</tbody>
</table>

Fig. 3. The ISI as a function of bandwidth, separately for the two event conditions (Slow = 0.03 Hz, Fast = 0.09 Hz).

second and third day 0.2. For the PCD (four signals together) this was 2.7% and 0.9% and for the ISI (four signals together) 0.4 s and 0.1 s. Looking at the PCD and the ISI for the individual signals the S36 and F12 signals seem to be more stable than the S12 and F36 signals. The behavior of the subjects was rather stable on the third day. Training effects were not the primary issue of interest and the discrimination made by the subjects among signal types was at maximum on the third day, so further data description is restricted to the third day.

For two subjects, data were deleted because their average ISI for the 0.09 Hz signal with 12 events was more than 2.5 standard deviations (Sds) above the average ISI of the other subjects and because the SD itself was about four times higher than that of the other subjects. For these subjects also there was a relatively low percentage correct detections (17.1% for the deleted, compared to 55.7% for the other subjects). The means for all dependent variables are therefore based on 10 out of 12 subjects.

On the third day in total 26881 samples were recorded. This suggests that the relevance of the significance values is debatable, because even small differences in ISI between factor levels will cause high significance values. However, if an ANOVA does not reveal significant values it is plausible to presume the absence of effects. For the ANOVA concerning ISI’s therefore a significance level of $P < 0.005$ was chosen. For ANOVA’s concerning other dependent variables the commonly used significance level of $P < 0.05$ was chosen.

**B. Sampling Selectivity**

For the ISI an ANOVA [ANOVA A] was conducted with session(3 levels), bandwidth(2), events(2), signal region(6), trend(2), and trend magnitude(3). First-order interactions were derived without session, second-order interactions were calculated only for bandwidth and events with signal region, trend, and trend magnitude.

Fig. 3 gives the effects on sampling behavior due to the simultaneous variation of bandwidth and number of events per signal. As is clear from the figure, the ISI was longer for the 0.03 Hz than for the 0.09 Hz signals $F(1,26817) = 825.4, p < 0.0001$ [ANOVA A] and longer for the 12 than for the 36 events signals $F(1,26817) = 887.2, p < 0.0001$ [A]. These data suggest that bandwidth and number of events equally affected the distribution of samples over the signal types. The interaction between bandwidth and number of events was significant $F(1,26817) = 9.1, p < 0.005$ [A], but very small, as can be seen in the figure.

Another ANOVA [ANOVA B] was conducted on all independent variables, including all first-order interactions, including “session”. The interactions between session and bandwidth $F(2,26805) = 3.0, p = NS$ [ANOVA B] and session and events $F(2,26805) = 4.8, p = NS$ [B] were not significant. This indicates that on the third day signal discrimination was not significantly different for all three sessions.

For the average total number of samples a one factor mixed model ANOVA was conducted with “session” as factor. Due to practical considerations it was conducted with the help of the BMDP program 8 V [1]. The average total number of samples for the first session was 864.8, for the second 890.3 and for the third 945.0: subjects tended to sample more frequently as a function of session $F(2,18) = 6.9, p < 0.01$.

**C. Local Sampling Behavior**

As is clear from Fig. 4, for all signal types ISI’s became smaller if the signal value approached the event region $F(5,26817) = 415.8, p < 0.0001$ [A]. Conspicuous is the difference in slopes between the 12 and the 36 events signals. The ISI differences among the four signal types became smaller near the event region. The ANOVA revealed significant interactions between bandwidth and signal region $F(5,26817) = 32.6, p < 0.0001$ [A], events and signal region $F(5,26817) = 30.7, p < 0.0001$ [A] and among bandwidth, events and signal region $F(5,26817) = 5.4, p < 0.0001$ [A].

Fig. 5 reveals that for the 4 trend, indicating a change away from the nearest event region limit, the ISI was larger for all four signal types than for the 7 trend, indicating a change in...
The IS1 as a function of signal region, separately for the four signal types.

The IS1 as a function of trend, separately for the four signal types.

The IS1 as a function of trend magnitude, separately for the four signal types.

The trend magnitude revealed for all signal types a more or less identical effect for the course of the IS1. The decrease was larger for the 12 than for the 36 events signals $F(1, 26817) = 1923.8$, $p < 0.0001$ [A]. The decrease was equal for both bandwidths $F(1, 26817) = 1.1$, $p = \text{NS}$ [A]. The decrease was the same for signal types with an equal number of events, suggesting that no interaction occurred due to the simultaneous use of bandwidth and number of events $F(1, 26817) = 5.8$, $p = \text{NS}$ [A].

The trend magnitude revealed for all signal types a more or less identical effect for the course of the IS1 (see Fig. 6). For both trend magnitudes this table reveals decreasing values as a function of the degree with which the signal value falls short of the event region. The size of the IS1 difference between the 12 events signals decreased as a function of trend magnitude, indicating that the overall influence of the global signal characteristic events was less prominent at larger trend magnitudes.

A trade-off between the importance of global signal characteristics and local features was expected. This suggests that there could be signal values where sampling behavior was dominated by signal bandwidth and number of events per signal and that there could be values where local signal features were more important. Figs. 4, 5, and 6 gave an indication of global and local effects. However, in these figures effects of signal region, trend and trend magnitude were interwoven, which makes it difficult to clearly understand the trade-off for each of the local features separately. Therefore A comparison between the IS1 for the $\delta$ (see Fig. 7(a)) and $\delta \delta \delta$ (see Fig. 7(b)) magnitude clearly demonstrates a difference in the general pattern of both trend magnitudes. For the $\delta$ magnitude, the distinction among the different signal types was larger far away from the event region than near that region. For the $\delta \delta \delta$ magnitude a similar pattern is found. But there is a clear difference in the degree of discrimination between both trend magnitudes; for the $\delta$ magnitude, offering information about relatively small changes, the distinction among IS1's for the different signal types was larger, probably indicating that in this situation global signal characteristics were more important for subjects.

Although Fig. 7(a) and (b) speak for themselves, sums of squares of the differences among the IS1 for the four separate signal types and the mean IS1 of those signal types that is, $\Sigma (\text{IS1}_i - \bar{M})^2$, were separately calculated for the $\delta$ and the $\delta \delta \delta$ magnitude, as a function of signal region. For both trend magnitudes this table reveals decreasing values as a function of the degree with which the signal value falls short of the event
D. Detections and Score

For the percentage correct detections and the number of incorrect detections, separate three factor mixed models ANOVA’s were conducted with session, bandwidth and number of events as factors. As is clear from Fig. 8, the percentage correct detections (PCD) was higher for the 0.03 Hz signals, the signal types with the relatively high predictability, than for the 0.09 Hz signals $F(1,9) = 30.9, p < 0.01$. The PCD was smaller for the 12 than for the 36 events signals $F(1,9) = 13.3, p < 0.01$. For the 0.03 Hz signals, the difference in PCD between 12 and 36 events signals was smaller than for the 0.09 Hz signals. But the interaction between bandwidth and number of events was not significant. Session effects were absent.

The average number of incorrect detections per session (6.0) was low in comparison with the average number of correct detections. No effect of session, bandwidth or events was found. For the score an one factor mixed model ANOVA was conducted with session as factor. The average score per experimental trial varied from 7.4 in the first to 7.6 in the second and to 6.5 in the third session $F(2;18) = 4.0, p < 0.05$. The low score for the third session was caused by the high number of samples.

IV. DISCUSSION

A. Experimental Method

The purpose of the present study is to investigate the effect of actually observed signal characteristics on human sampling or observing behavior. The observing response method is used to gather data, while for example Senders [18], [19] used eye movement recordings. Data gathered with the observing response method are by no means visual sampling data such as those gathered with experiments using eye movement recordings. But although data collected with these different...
techniques cannot be compared directly [8], this does not mean that the data of the present experiment cannot be discussed in the framework of sampling models in which a subject’s knowledge about actual signal values is incorporated. Senders was the first one who developed such models, the Conditional Sampling Models (CSM models). In 1983 he wrote [18]:

"...Clearly, if we are to understand WHY people do what they do, we must in fact learn WHAT they in fact do. It is necessary to record not only the positions of the eyes (that is the monitoring behavior), but also the values of the signals which are observed. It is only in the relation between these two sets of data that will tell us whether there is anything at all in the idea that observers make use of information that they see, in deciding when to look again \ldots\".

In fact that was one of the main reasons to use the observing response method, and the results indeed give strong evidence in favor of the conditional sampling models (see further on).

Another difference between the present and other studies on monitoring behavior is the presentation of the signals. This study uses a digital presentation with an additional rate information indication, while most other studies used analog signals, in which rate information can be deduced directly from the changing signal value. Again this is an important choice. We believe that, as Senders [18] described it "...that it is the nature of the signals which drives the behavior, rather than the method of display itself \ldots\". So, it is our opinion that, although the presentation mode is important and will have an effect on monitoring behavior, this effect is smaller than the effect of for example bandwidth, signal value or actual signal state. So the results of this study are at least qualitatively valid.

B. Sampling Selectivity

Both global signal characteristics, bandwidth and number of events, equally affected the way in which subjects distributed samples over signals (see Fig. 3). The ISI decreased with an increase in bandwidth and number of events. In an earlier experiment of Bohnen and Leermakers [3] only bandwidth differences could be used to discriminate among different signals. The observed slope had a coefficient of 1.17 while in this study it was 0.37. Although this comparison concerns two different task situations, this can possibly be seen as support for results of Van Delft [6], [7], concerning binary signals, which showed that no fixed relation exists between signal bandwidths and the frequency with which signals are sampled. This means that the normative indications derived from the theorem of Shannon [20] are only of limited use. The theorem gives a rather inaccurate image of the relation between signal bandwidth and sampling frequency in this kind of experiments.

Although the results are not in accordance to what would be expected from Shannon and Weavers theorem they seem quite rational. Bandwidth is related with the signal predictability which is relevant for an efficient distribution of samples. This confirms experimental results [5], [10], showing that people are sensitive to the rate with which uncertainty is generated by monitored processes. Events can be seen as a kind of reward or as a reinforcer [9] and as an important factor which can determine sampling behavior, especially in this task situation, where subjects were rewarded for correct detected events. This is supported by the results presented in Fig. 8 showing that the percentage correct detected events was higher for the 36 than for the 12 events signals.

The average number of samples made by the subjects increased as a function of session. Admittedly, a post-hoc explanation can be fatigue as a function of session, whereby the level of activation decreased. An updating of the internal representation with the help of samples [12] may require less cognitive effort than some kind of reasoning with the help of this representation [2], which puts a heavy load on working memory. The latter is a controlled process, whereas sampling can be quite automatic and thus requires less cognitive effort.

C. Local Sampling Behavior

The present study confirms earlier results for numerically displayed signals [3], [11] that, sampling behavior is influenced by the degree with which the signal value falls short of the event region. The ISI decreased for all signals if signal values approached the event region (see Fig. 4). Near the event region (region V) local information dominated sampling behavior and the influence of both global signal characteristics (bandwidth and events) was small.

The present study also demonstrates that sampling behavior was influenced by the value of the rate of change indication, divided in a trend and a trend magnitude (see Sect. II-E), and the results are in line with earlier results [3]. For all signals the ISI was smaller for the \( \uparrow \) trend (see Fig. 5), indicating a higher probability that an event will occur in the near future. The decrease was equal for both bandwidths. This is understandable: bandwidth by itself is statistically not related to the direction of the changes. The number of events revealed an interaction with the trend (see Fig. 5). The ISI decrease was larger for 12 than for 36 events signals, indicating that the influence of local features was greater for signals with a small number of events. The same can be remarked on the value of the trend magnitude (see Fig. 6). The ISI is smaller for larger trend magnitudes, indicating a higher probability that an event will occur in the near future and the influence of trend magnitude was greater for signals with a small number of events.

The results concerning local features (see Figs. 4-6) revealed that the distinction in ISI among signals became smaller when the attentional demands brought about by the local signal features become of greater importance. Thus, the influence of global signal characteristics diminished in these situations. This is confirmed by results concerning pure local features (see Fig. 7 and Table IV). The distinction among signals is larger in region I then in region V and larger for the \( b \) than for the \( b \delta b \) magnitude. Where the \( \uparrow \) trend indicates changes in the direction of the event region, the probability that an event will occur in the near future is higher for the \( b \delta b \) magnitude than for the \( b \) magnitude. This is an important issue; subjects
likely have some kind of internal representation that indicates how to distribute samples over signals. However, they change their strategy when important local features appear, the so-called conditional sampling behavior [18], [19]. This suggests that subjects do not sample with fixed sampling sequences, but that they change strategy to make it possible to focus on more relevant information.

The importance of the local features was dependent on both global characteristics. When local information with a high attentional value is presented, bandwidth seems to dominate over the number of events. This is not the case when local information is of less importance. This means that not predictability as governed by the global signal characteristics as such, but rather predictability given certain local features is a crucial factor which can determine sampling behavior. Pure global sampling behavior indications, in which the influence of local features are omitted (see Fig. 3) give a rather limited image of the influence of global characteristics. Indications in which the influence of local features are taken into consideration (see e.g., Fig. 6) will contribute to a more accurate image of the structure and state of the internal representation. In fact, it is necessary to be able to make a clear distinction between samples associated with actually observed local features and samples made to update the internal representation [12]. Such a distinction could reveal a more accurate image of mechanisms governing sampling behavior. However, we did not have a method to make definitely a distinction between these two kind of samples.

D. Detections and Score

The significance of bandwidth for the percentage correct detections per signal (see Fig. 8) could be explained through the lower predictability for 0.09 Hz signals. Another explanation could be the smaller duration of events for these signals (average of 3.5 s compared to 11.8 s for 0.03 Hz signals). Therefore it was probably more difficult to obtain a correct image of the number of events to detect. The difference between the percentage correct detections for 12 and 36 events signals could be partially explained with the help of this statement too; the average duration of events for 12 events signals was 6.9 s compared to 8.4 s for 36 events signals. Another explanation can be an incorrect estimation of the number of events, probably the same as incorrect estimations of proportions in static environments [16]. Possibly this was facilitated by the higher percentage of time that 36 events signals were in the region ≥25.0 (see Table I, region VI). An additional explanation can be that 36 events signals received more attention just because of the higher number of events and with that the higher possibility for the subjects of receiving rewards.

Concerning the relatively low percentage correct detections for the 0.09 Hz signals (see Fig. 8), it is conspicuous that these signals were not sampled more frequently. Yet, it is too easy to conclude that this is an indication for a nonoptimal sample distribution. Considered in the light of the average score of 7.2 it is even a reasonable effective sampling distribution.

V. Conclusion

The importance of local features on sampling behavior was dependent on both signal bandwidth and number of events per signal. It is remarkable that local features had greater effects for 12 than for 36 events signals. In monitor situations in process industries, where commonly only a few disturbances occur, local features can probably be of an even greater importance, suggesting that the way in which the information is presented can be most important.

Intuitively, it seems desirable to give subjects the chance to handle the uncertainty associated with the sampling process in the way they prefer. This can be done by offering them the opportunity to determine themselves the moment or the condition under which they like to get a sample of the signal which they want to examine. This also can give a more accurate view of the processes that govern sampling behavior and possibly puts a lower demand on working memory.

In this study both global signal characteristics equally affected the way in which subjects distribute samples over signals. In real-life situations, however, the influence of events will depend on the value of the outcome for the operator. This means that pay-off structures will be an important factor in the control of sampling behavior.

The present data form a too small basis for a thorough discussion. The same holds for applying the several quantitative models known from literature (e.g., [4], [10], [11], [21], [22]). In our opinion these models are too complex for a behavioral validation because the strategy of subjects may depend on minor changes in condition so that exact defined models cannot be applied on the data of this study.

The bandwidths used in this study are high compared to those in the slow responding monitor situations in the process industries. It is assumed that bandwidth will be a less important determinant of sampling behavior in that kind of situations. This can be explained by the fact that endogenous sources of uncertainty [14], such as forgetting, will play a more important role in the control of sampling in that kind of situation instead of the uncertainty arising from the dynamics of the system (exogenous sources). This suggests that psychological instead of quantitative models need to be further explored, since the consequences of events outside the tolerance region in process monitoring situations can be even larger than in this study, or can indeed lead to serious industrial accidents.

Acknowledgment

The authors would like to thank J. Moraal for comments on earlier drafts of the manuscript.

References

Harrie G. M. Bohnen received the M.Sc. degrees in both experimental psychology and ergonomics from Tilburg University in 1987 and 1989, respectively. From June 1989 to April 1991, he was employed at the Eindhoven University of Technology, Eindhoven, The Netherlands. He is now employed at the National Aerospace Laboratory NLR. His research interests focus on pilot performance, stress, mental workload assessment and prediction, human computer interaction and pilot-vehicle interface design.

Martinus A. M. Leermakers received the M.E. degree in physics engineering from Eindhoven University of Technology, Eindhoven, The Netherlands, in 1979. Since 1986, he has been with the faculty of Industrial Engineering, Eindhoven University. His current research interests are human performance and human computer interaction.

Piet J. Venemans received the M.E. in physics engineering/measurement and control in 1987 from Delft University of Technology, Delft, The Netherlands. From 1972 to 1991, he was with the Department of Psychology, Tilburg University, for the Human Factors programme. In 1991, he joined the Faculty of Building, Eindhoven University of Technology, Eindhoven, The Netherlands, for research on way finding and orientation in public buildings. His research interests include the user aspects of control systems, the usability aspects of buildings and the translation of user research findings to design support tools.