Psychophysiology
Copyright © 1986 by The Society for Psychophysiological Research, Inc.

Vol. 23, No. 6
Printed in U.S.A.

Multiple P3s to Emotional Stimuli and Their Theoretical Significance

Victor S. Johnston,
Department of Psychology, New Mexico State University, Las Cruces, New Mexico

David R. Miller,
Human Factors Department, IBM, Boulder, Colorado

and Mary H. Burleson
Department of Psychology, New Mexico State University, Las Cruces, New Mexico

ABSTRACT

Event-related potentials (ERPs) were recorded to both predictive (consonant-vowel-consonant) and feedback (picture) stimuli as subjects learned associations between these stimuli. The consonant-vowel-consonants (CVCs) were selected for lack of emotional content while the pictures (PICs) varied in emotional value (Learning Group; N=20). A second group of subjects was exposed to the same CVC-PIC stimuli but was required only to count the number of different CVCs and PICs (Counting Group; N=20). A principal components analysis with varimax rotation was performed on ERPs to PICs and revealed multiple late positive components (P3 and P4) and a slow positive wave (SPW). In both groups, the P3 and P4 factors varied with the emotional value of the stimuli. The learning group had a larger SPW than the counting group, and disconfirmed predictions elicited larger P4s than confirmed predictions. For CVC stimuli, only P4 increased as subjects learned CVC-PIC relationships. From the similar scalp and temporal distributions of P3 and P4, as well as their functional similarity, it was concluded that these two factors reflect the same neural process which is activated by the emotional value of stimuli. One possible function of this process is suggested.

DESCRIPTORS: Event-related potentials, P3, P4, Slow wave, Emotion, Utility, Learning.

The concept of utility, the personal subjective value of a stimulus, was proposed as a major factor determining P3 amplitude in order to reconcile the effects of "task relevance" with the effects of manipulating the "monetary value" of a stimulus (Johnston, 1979). It is clear that both monetary value and task relevance are learned stimulus attributes. In a subsequent study, Johnston and Holcomb (1980) examined how the P3 elicited by a neutral stimulus (S1) could come to reflect the value of a subsequent high utility event (S2) as subjects learned the relationship between S1 and S2 in a paired associate learning task. The question remains, however, as to the nature of the original value system from which learned values are derived.

Address requests for reprints to: Dr. Victor S. Johnston, Department of Psychology, Box 3452, New Mexico State University, Las Cruces, New Mexico 88003.

It appears that all value must originate from an inherent emotional structure which: 1) defines certain stimuli as desirable or undesirable for biological survival, and 2) provides a cognitive mechanism for elaborating this basic emotional structure. From this perspective, all value is emotive in nature. Begleiter, Porjesz, Chou, and Aunon (1983) have discussed this issue, and concluded that "the cognitive processes for determining the 'meaning' or 'significance' of stimuli must of necessity possess a fundamental and critical emotive aspect" (p. 100). The idea that emotional value is a major determinant of P3 amplitude is supported by Simons' (1982) study which showed reduced P3 amplitude in anhedonic subjects, compared with normals, to stimuli which were predictive of high interest slides. The present experiment is an attempt to systematically manipulate the emotional value of stimuli and examine how this is reflected in P3 amplitude. Furthermore, the study is designed to determine if ERPs to neutral stimuli change as subjects learn the relationship between these stimuli and ones with high emotional value.

Method

Subjects

The subjects were 20 male and 20 female students (18–35 yrs old) who volunteered to participate in an experiment in which they would be exposed to visual stimuli which varied in emotional content. All subjects were aware that the stimulus slides included male and female models "such as might be seen in Playboy or Playgirl magazine."

Stimuli

The stimulus material consisted of 20 consonant-vowel-consonant trigrams (CVCs) and 20 picture slides (PICs). The CVCs were selected to be low in meaningfulness and association value, based upon Noble's (1961) study which investigated these attributes in 2100 CVC combinations. The first position consonants and vowels were both equally represented across the 20 stimulus slides.

The picture slides were selected on the basis of a previous study by Miller (1985) in which both male (N=24) and female (N=24) subjects were required to rate 124 different pictures on a five-point emotional scale ranging from very pleasant to very unpleasant. Based on these ratings, Miller (1985) identified five stimulus categories (four slides/category). These included three categories (pleasant, neutral, and unpleasant) whose ratings did not overlap. Their content was babies (B), ordinary people (P), and dermatological slides (D), respectively. The remaining two categories showed significant sex differences; male models (M) were rated high by females and low by males, and female models (F) received the opposite ratings. These five groups of slides were used to manipulate emotional value in the current study.

A Kodak Carousel projector, fitted with an electronic shutter, back-projected the stimulus slides onto a translucent plexiglas screen. The images subtended approximately 10 degrees of visual angle. The projector advance mechanism and shutter were under the control of a IBM PC/XT computer. A 500 Hz tone was used as the auditory signal.

Procedure

Each subject was fitted with three active nonpolarizing (Ag/AgCl) electrodes (F_Z, C_Z, P_Z), two linked reference electrodes (A_1, A_2), and both supra- and suborbital electrodes to monitor vertical EOG. All electrode resistances were adjusted to less than 10 Kohms, and signals were amplified using Grass Model 7P122 amplifiers with upper and lower half amplitude frequency responses at 60 Hz and 0.04 Hz respectively. All EEG and EOG recordings were digitized at 100 Hz, and together with trial information, were stored on an IBM PC/XT hard disk during the intertrial interval. ERP and EOG waveforms were collected during the 1000-ms interval following all slide presentations. For the purpose of data analysis, average waveforms were computed only from trials which met specified criteria in the absence of EOG artifacts. An online measure of EOG was employed, and subjects were eliminated from the study if eye movements or blinks occurred on more than 20% of the trials. Three learning and two counting subjects were rejected on this basis and subsequently replaced.

Subjects were seated in an electrically shielded, semi-dark room, with the projection screen directly in front of them and a button panel adjacent to their right hand. The panel had five buttons in one row and a single button mounted below and centered. This single button was used to initiate trials, while the top row was used to register ratings on the five-point emotional scale (very unpleasant, unpleasant, neutral, pleasant, and very pleasant).

The first phase of the experiment was concerned with determining the slide ratings for each subject. Initially, to acquaint the subjects with the full range of stimulus material that they would be required to rate, all CVCs and PICs were presented for 100 ms, with 1200 ms between slide presentations. Following this preview, all subjects were informed of the rating system and were then required to view and rate each slide as they were presented in a random order. To rate a slide, subjects first pressed the initiation button and 500 ms later the slide was presented for 100 ms. A 100-ms tone, one second later, informed the subjects when to register their rating on the clearly marked five-point scale. The procedure continued until all 40 CVC and PIC slides had been rated.

During the learning phase of the experiment, the 20 learning subjects (10 males, 10 females) were required to learn the relationship between the 20 CVCs and the 20 paired pictures, over the course of four sessions. Each session consisted of 100 CVC-PIC combinations presented in a random order, with the actual CVC paired with each picture being different for each of the 20 subjects. The learning subjects, therefore, saw each CVC-PIC pair a total of 5 times within each session, or 20 times altogether. A 100-ms tone informed subjects that they could initiate a trial with a button press. This response was followed 500 ms later by a CVC slide presentation of 100 ms duration. After one second, the tone again informed subjects when they could predict their rating, in other words the category, of the next slide. The associated PIC slide was then presented, after a 500-ms delay following their predictive rating response. A feedback signal, one second later, indicated a correct (double tone) or an incorrect (single tone) prediction. As an example, a prediction of neutral was considered correct if any one of the slides which were rated neutral by that subject in the first phase of the experiment appeared after the prediction. In the event of an incorrect prediction, subjects were required to identify the correct response button and receive a double tone before the next trial could be initiated.

The 20 counting subjects (10 males, 10 females) saw each slide the same number of times as the learning subjects, but there was no consistency as to which CVC preceded any PIC. Following the first tone and initiation response, they were presented with the CVC-tone PIC-tone sequence, similar to the learning subjects. However, their response to the tone following the CVC was not a predictive response, but served only to initiate the presentation of the next picture. Their task was to count the number of different CVCs and the number of different pictures which were presented during each session. This procedure required subjects to attend to and correctly identify each stimulus, but did not require any learned association between CVCs and PICs.

Results

Behavioral Analysis

To validate the use of the slide categories as a means for manipulating emotional value, it was first necessary to demonstrate that the original categories were rated differently from each other by the current group of experimental subjects. Categorical differences were tested using a loglinear analysis for categorical data. This analysis revealed that the categories were indeed rated differently, chi-square (16, N=20) = 139.76, $p<.0001$. A series of pairwise chi-squared comparisons of every category with every other category, for male and female subjects separately, demonstrated that the categories were rated differently in all but one case; although female subjects rated the baby category slightly higher than the male model category, the difference was not significant. A comparison between the model categories for each sex showed that male subjects rated the female models higher than female subjects did, chi-square (3, N=20) = 45.6, $p<.0001$, and female subjects rated the male models higher than did male subjects, chi-square (3, N=20) = 47.3, $p<.0001$. Based on these observations, we can consider the different slide categories as providing material of different emotional content.

Two measures of learning were examined. The first was the number of trials to a criterion of three consecutive correct predictions (start); the second was the total number of correct predictions (score). An ANOVA of score revealed no significant category effects, but did show a sex difference, $F(1/18)=12.97$, $p<.002$, with male subjects performing better than female subjects. Start showed a category difference, $F(4/72)=2.97$, $p<.05$, as well as the sex difference, $F(1/18)=14.48$, $p<.0013$. The category means were ordered F, B, M, P, D, from earliest to latest start.

ERPs to PICs

Average ERPs to the PIC slides as a function of emotional category are shown in Figure 1. The ERPs are divided into positive and negative affect according to the mean rating found by Miller (1985).

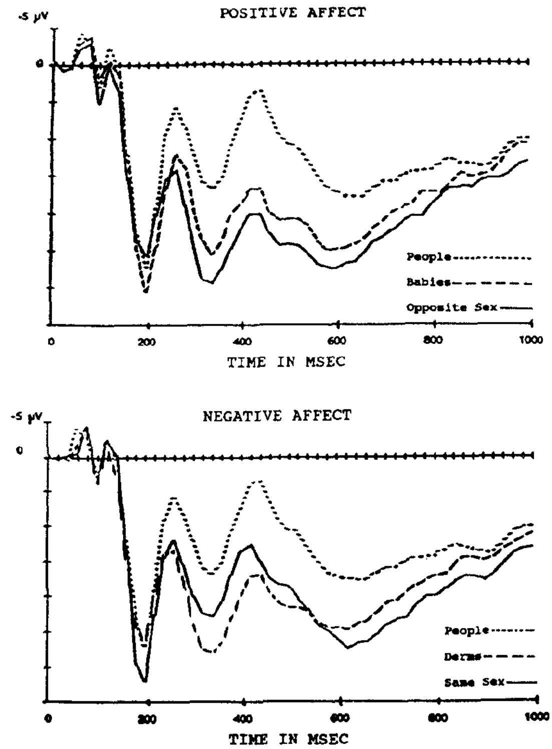

Figure 1. Average ERPs to PICs as a function of emotional category. Positive categories include people, babies, and opposite-sex models; the negative categories are people, dermatological slides, and same-sex models. The people category is shown in both graphs for comparison purposes. (P_Z electrode site; averages over all learning subjects.)

Visual inspection suggests that the emotional category of a stimulus has a major effect on late components of the ERPs.

To further examine this relationship, a Principal Components Analysis and Varimax Rotation (PCVA) was performed on the average waveforms, each being represented by 49 time points. Forty subjects, four sessions, five emotional categories, and three electrode sites yielded a total of 2400 average waveforms (mean of 18.6 waves per average), which served as the data base for the PCVA. The covariance matrix was employed to preserve the amplitude variance around the grand mean. Five factors were rotated. The grand mean waveform, together with the factor loadings for the first three factors, accounting for 78% of the variance, are shown in Figure 2.

In order to facilitate the reporting of results, the ERP peaks have been initially named according to their polarity and order of occurrence. Also, the PCA factors have been designated according to their

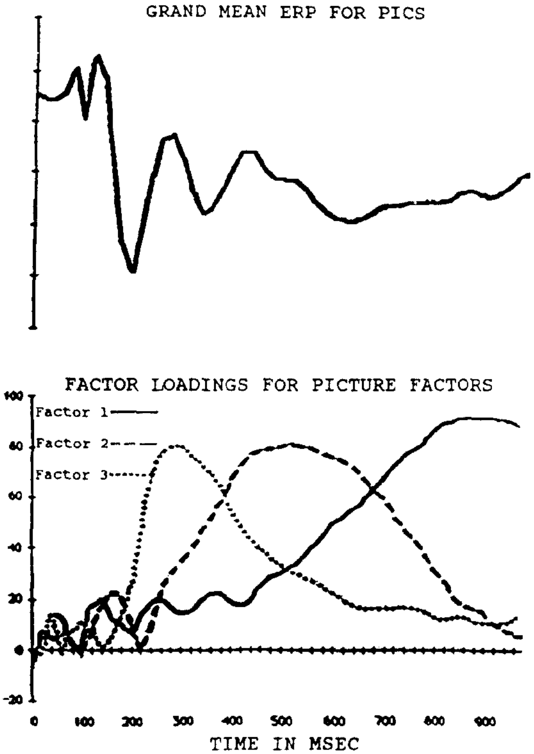

Figure 2. Grand mean ERP to PIC stimuli, over all subjects and conditions, together with the factor loadings of the first three rotated factors from the PCVA.

polarity and the ERP peak which corresponds to their maximum activity. The former was determined from a PCA using the crossproducts matrix, and the latter was determined from the temporal location of the maximum factor loadings. Using this nomenclature, Factor 1 (max at 920 ms) is a late-developing positive slow wave (SPW), Factor 2 (max at 540 ms) corresponds to P4, and Factor 3 is most active at the P3 peak of the ERP (max at 300 ms). The relationship between these components and those identified in other studies will be discussed more fully following an examination of their scalp distribution and their sensitivity to the independent variables.

An analysis of variance for mixed designs was used to determine whether the factor scores for any factor were different as a function of the experimental variables. Treatment group and sex were between-subject variables, while emotional category, session, and electrode site were within-subject variables (Myers, 1979). Because the assumption of homogeneity of variance cannot be justified for ERP data, all F values were interpreted using the Greenhouse-Geisser (1959) correction.

Factor 1 (SPW) showed a main effect of Site, $F(2/72)=8.53$, $p<.0001$, and both Group $\times$ Site, $F(2/72)=7.46$, $p<.004$, and Category $\times$ Site, $F(8/288)=3.98$, $p<.003$, interactions. Bonferroni t tests revealed the scalp distribution to be $C_Z>F_Z>P_Z$. The Group $\times$ Site interaction was due to the larger SPW at P_Z than F_Z in the learning, but not the counting group; C_Z was largest for both groups, and especially large for the learning subjects. A smaller SPW factor score for people slides at the P_Z site, compared to C_Z and F_Z, accounted for the Category $\times$ Site interaction.

The P4 factor scores revealed main effects due to Group, $F(1/36)=8.68$, $p<.006$, Category, $F(4/144)=40.71$, $p<.0001$, and Site, $F(2/72)=48.78$, $p<.0001$. Bonferroni t tests revealed P4's scalp distribution to be $P_Z>C_Z>F_Z$, and P4 to be larger in the learning, compared with the counting group. Similar tests showed P4 to be largest for the female model slide category and smallest for ordinary people (F>B=M=D>P). There were also significant Sex $\times$ Category, $F(4/144)=4.95$, $p<.002$, and Category $\times$ Site, $F(8/288)=11.69$, $p<.0001$, interactions. The Sex $\times$ Category interaction resulted from the P4 being smaller in male compared with female subjects for all categories except for the female model category; in this case the males had a much larger P4 than the female subjects. The P4 factor scores had a similar pattern across categories for all three electrode locations except for a larger P4 to the dermatology slides at the P_Z recording site; this resulted in the Category $\times$ Site interaction. A Session $\times$ Category $\times$ Site interaction was also significant.

Like P4, the P3 factor also showed main effects for Category, $F(4/144)=30.01$, $p<.0001$, Site, $F(2/72)=88.80$, $p<.0001$, and a Category $\times$ Site interaction, $F(8/288)=13.09$, $p<.0001$. In addition, there was a Group $\times$ Category interaction, $F(4/144)=5.68$, $p<.0006$. Bonferroni t tests established a parietal dominance in the scalp distribution of the P3 component ($P_Z>F_Z=C_Z$), largest P3s to the female model slides, and smallest P3s to the people category (F>M=B=D>P). All categories produced a $P_Z>F_Z>C_Z$ scalp distribution except babies ($P_Z>C_Z>F_Z$); this effect was responsible for the observed Category $\times$ Site interaction. Finally, the Group $\times$ Category interaction resulted from the P3 factor scores for all categories, except female models, being lower for the learning group; for this category, the learning group exhibited larger P3s.

Since a PCVA analysis of ERPs can be misleading when latency differences are present in the data, a baseline-to-peak analysis was performed to validate the above results. Using the mean of the first two points as a baseline, the P3 and P4 peaks were defined as the maximum values in the windows

300–380 ms and 600–680 ms respectively. These peak amplitudes were analyzed using the same ANOVA model as described above.

The P3 peak data replicated the P3 factor findings. Again, there were main effects of Category, $F(4/144)=36.04$, $p<.0001$, and Site, $F(2/72)=101.87$, $p<.0001$, as well as the Category × Site, $F(8/288)=3.16$, $p<.007$, and Group × Category, $F(4/144)=2.95$, $p<.03$, interactions noted for P3 factor scores.

The P4 peak data also confirmed the main effects of Group, $F(1/36)=10.55$, $p<.003$, Category, $F(4/144)=22.08$, $p<.0001$, and Site, $F(2/72)=59.2$, $p<.0001$, observed for the P4 factor. Unlike the P4 factor, there was an additional interaction of Group × Site, $F(2/72)=4.89$, $p<.02$, as well as the Sex × Category, $F(4/144)=2.76$, $p<.05$, and Category × Site, $F(8/288)=3.56$, $p<.004$, interactions observed in the P4 factor scores. The Group × Site interaction was due to a larger P4 peak at C_Z than P_Z ($C_Z>P_Z>F_Z$) in the learning group compared to the counting group ($P_Z>C_Z>F_Z$). It appears that the peak analysis has confounded the SPW (large at C_Z in the learning group) with the P4 factor; this limitation of peak analysis in separating these two overlapping components will be discussed later.

To help clarify the effects of learning, new average ERPs were computed for the learning subjects according to whether the PIC stimulus confirmed, or disconfirmed, the subject's prediction (outcome). Since there were no session effects, these averages were computed over sessions, and the factor scores were derived using the factor score coefficients from the previous analysis. An ANOVA revealed that P4 was the only factor sensitive to outcome, with disconfirmation resulting in larger P4s than confirmation, $F(1/36)=11.9$, $p<.002$; this was also true of the P4 peak, $F(1/36)=17.54$, $p<.0002$. A separate analysis of each outcome revealed that the category effects on P3 and P4 noted above were also significant within each outcome condition.

ERPs to CVCs

All 2400 CVC average waveforms were analyzed using the same procedure as for the PIC data. The grand mean waveform for CVCs and factor loadings for the first three factors, accounting for 76% of the variance, are shown in Figure 3. Average CVC waveforms for the learning subjects, for each session, are shown in Figure 4. Despite large differences in the waveforms to CVCs and PICs, the sources of variance in these waveforms, as reflected in the factor loadings, are very similar in both temporal distribution and polarity. With the exception of Factor 1, where scalp distribution was not significant in the CVC waveforms, the other two factors had the same scalp distributions as the corresponding picture factors; that is, a main effect of Site for Factor 2, $F(2/72)=24.5$, $p<.0001$ ($P_Z>C_Z>F_Z$), and for Factor 3, $F(2/72)=47.2$, $p<.0001$ ($P_Z>F_Z>C_Z$). Because of these similarities in spatial and temporal distribution, the CVC factors have been named SPW (Factor 1), P4 (Factor 2), and P3 (Factor 3) to reflect their relationship to the picture factors.

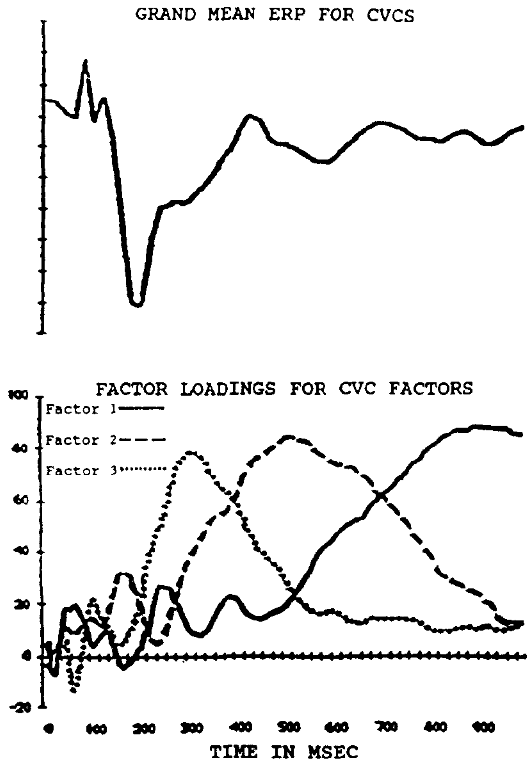

Figure 3. Grand mean ERP to CVC stimuli, over all subjects and conditions, together with the factor loadings of the first three rotated factors from the PCVA.

The only significant effect on the CVC SPW factor was a Sex × Session interaction, $F(3/108)=4.42$, $p<.01$. Over the first three sessions, this factor increased for female subjects but decreased for male subjects.

As well as the main effect of site noted above, there were three significant two-way interactions on the CVC P4 factor. These were a Group × Site interaction, $F(2/72)=10.11$, $p<.001$, a Group × Session interaction, $F(3/108)=3.5$, $p<.03$; and a Sex × Category interaction, $F(4/144)=3.25$, $p<.02$. The Group × Site interaction was due mainly to a much larger P4 in the learning group at the P_Z site, but a smaller P4 at the F_Z site. The observed Group × Session interaction is shown in Figure 4;

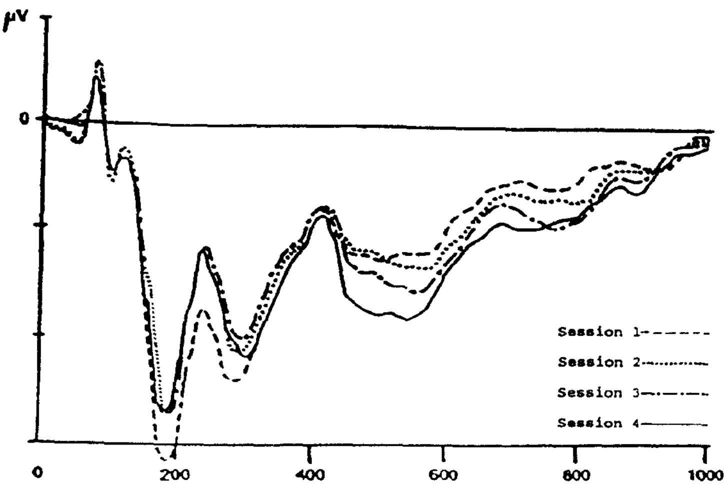

Figure 4. Mean ERPs to CVC stimuli for each session (P_Z electrode site, learning subjects only). Session averages show a systematic increase in positivity (downwards on figure) from Session 1 to Session 4, when the P4 factor is most active (550–600 ms after stimulus onset).

only the learning subjects showed an increase in this factor over sessions. Finally, the Sex × Category interaction was due to lower P4s in male subjects for all categories except female models; males had much larger P4s than female subjects for this category.

In addition to the main effect of site noted above, the only other significant effect on the CVC P3 factor was a main effect of session. P3 decreased systematically over sessions, $F(3/108)=14.7$, $p<.0001$ (S1>S2>S3=S4 from Bonferroni t tests).

Peak amplitude measures of P3 and P4 were also examined, using the windows 280–360 ms for P3, and 540–660 ms for P4. An ANOVA of the P4 peak confirmed the main effect of Site, $F(2/72)=30.8$, $p<.0001$, the Category × Sex interaction, $F(4/144)=3.15$, $p<.03$, and the Group × Site interaction, $F(2/72)=9.5$, $p<.002$, already observed for the P4 factor. In addition, however, the peak analysis revealed a main effect of Sex, $F(1/36)=6.14$, $p<.02$ (Females>Males), and two three-way interactions, Group × Sex × Session, $F(3/108)=3.08$, $p<.05$, and Group × Session × Site, $F(6/216)=2.79$, $p<.04$. The Group × Sex × Session interaction appears to be the result of the peak analysis not separating the Group × Session effect (observed on the P4 factor) from the Sex × Session effect (localized to the SPW factor using PCVA). This is a good example of a case where PCVA may be superior to peak analysis due to an overlap in the temporal distribution of the components. The final three-way interaction, Group × Session × Site, is a consequence of the much greater increase in P4 amplitude over sessions at the P_Z site of the learning group than the other electrode locations.

An analysis of the P3 peaks duplicated the main effects of Site, $F(2/72)=58.3$, $p<.0001$, and Session, $F(3/108)=12.18$, $p<.0001$, which were observed for the P3 factor.

Using the same strategy as for the PIC waveforms, the CVC data of the learning group were examined to determine the effects of outcome (confirmation or disconfirmation). No factors or peaks in the CVC data were found to be sensitive to this variable.

Discussion

Central to the interpretation of the current study is the issue of whether emotional value was indeed the independent variable being manipulated by the presentation of the PICs. Both the consistent ratings between groups on the emotional value scale and the sex reversal in ratings for the nude model slides suggest an emotional interpretation of the stimulus material. It is unlikely that any physical differences in the slides accounted for the results since, except for the presence of a single human figure, all other physical aspects of the stimulus slides varied widely within each stimulus category. A more plausible hypothesis is that the stimulus materials within each category differed in their subjective probability of occurrence in the real world, even though they were equally probable within the experimental situation. However, if this were so, we should expect the dermatological slides, which were extreme cases taken from textbooks, to be the most improbable slides and thus to elicit the largest late components. This was not the case. Given the consistent ratings on the emotional scale, the sex differences in those ratings, and the lack of any other non-emotional attentional attribute (size, color, probability, complexity, etc.) which varied consistently between categories, it appears that emotional value is the critical variable being manipulated by the PIC categories.

An examination of the factor loadings for PICs and CVCs reveals that despite large differences in the stimulus material and associated ERPs, most of the variance in both sets of data can be accounted for by three overlapping positive components. These factors, which were derived independently from the variance in the PIC and CVC waveforms, have the same scalp as well as temporal distribution. Since a peak analysis, using windows to reduce the effects of latency differences between subjects, replicated most of the PCVA findings, we can conclude that the PCVA components are not the result of latency differences in the ERPs. Where peak and factor analyses differed, the latter provided the most parsimonious interpretation. All the major differences in amplitude, latency, and scalp distribution between PIC and CVC waveforms appear to be accounted for by the relative contribution of these three principal components. To simplify the discussion, the PIC data and the CVC data will be discussed separately before attempting to integrate the findings.

ERPs to PIC Slides

Multiple late positive components of ERPs to feedback stimuli have been reported by several authors (Stuss & Picton, 1978; Stuss, Toga, Hutchison, & Picton, 1980) and their theoretical significance has been discussed in a thoughtful paper by Johnson and Donchin (1985). The feedback tones used by Johnson and Donchin (1985) varied in discriminability, resulting in latency differences of the positive components both within and between subjects, and an inevitable spread of the ERP peaks on the average waveforms. However, when latency adjustments were made on individual waveforms, they observed P3 and P4 factors (additional P3s in their interpretation) with the latter often superimposed on a positive slow wave, apparent even on individual trial waveforms. These authors argue that the two late positive components are indicative of serial successive decisions which are time-locked to some extent since the second decision is dependent upon completion of the first decision. The consequences of this model are: 1) factors which delay the first decision inevitably delay the second decision; 2) there is a minimum time, perhaps constant in some cases, between the first and the second decisions; and 3) the first decision involves "stimulus related" variables whereas the second or subsequent decisions are sensitive to "task related" variables.

Stuss et al. (1980), using feedback lights to indicate the correctness or incorrectness of a complex classification task, also report clearly defined P3 (354 ms) and P4 (590 ms) peaks on the average ERPs to feedback stimuli. These factors bear a close resemblance to the current study both in their temporal distribution and their sensitivity to the experimental variables. Although both factors are delayed in the current study, the time between the factor peaks is 240 ms, very similar to the 236 ms found by Stuss et al. (1980). Using the model outlined above, we should expect both factors to be delayed in the current study due to the additional processing time required for picture analysis compared with a single light flash.

The effect of emotional value on ERPs, as revealed by a significant effect of slide categories, was evident in both the P3 and P4 components of the picture ERPs recorded from both learning and counting subjects. Category differences were evident in the learning group for both confirmed and disconfirmed outcomes, but disconfirmation resulted in significantly larger P4s than confirmation. It appears therefore that two separate effects are being reflected in the P4 factor: the emotional value of the stimulus, and an evaluation of the trial prediction. In contrast, the P3 factor was sensitive only to the emotional value of the stimulus; task outcome, confirmation or disconfirmation, having no effect. These observations are consistent with the Johnson and Donchin (1985) serial stage model, where stimulus factors are registered on the first positive factor, and task variables are reflected in later positive components. In keeping with this hypothesis, counting subjects, who require no evaluation of outcome, have significantly smaller P4 components than learning subjects.

In simple prediction tasks disconfirmation normally results in a larger P3 or P300 component than confirmation. That the P4 component was sensitive to outcome in the current study supports the hypothesis by Johnson and Donchin (1985), "that these additional positive peaks represent P300 activity" (p. 182). These authors, however, noted a different scalp distribution between their P3 and P4 factors, somewhat weakening the argument for their functional equivalence. They suggest that perhaps "the scalp distribution is shifted by the superimposition of additional processes" (p. 192) or that the two components may reflect the "manifestation of similar functional activities that are undertaken by different populations of neurons" (p. 192). In the current study, however, the scalp distributions of P3 ($P_Z>C_Z=F_Z$) and P4 ($P_Z>C_Z>F_Z$) were very similar. This suggests that the same population of neurons is responsible for both factors, but it may be necessary to use a PCVA to separate them from other components; specifically, the overlapping SPW which has a $C_Z>F_Z>P_Z$ distribution.

A peak analysis of the current data revealed a P4 component with a C_Z dominant distribution, similar to that observed by Johnson and Donchin (1985). However, the PCVA analysis separated this into a P_Z dominant P4 factor, and a C_Z dominant SPW factor; this may explain why Johnson and Donchin (1985), who used only a peak analysis, found different scalp distributions for the two positive components. This interpretation becomes more convincing when we examine what happens when subjects are required only to count the feedback stimuli. Johnson and Donchin (1985) report a "shift from a C_Z maximum to a P_Z maximum" when their subjects' task was changed to a counting task; this is the very condition when SPW activity at C_Z is reduced and results in a P_Z dominant peak for counting subjects in the current experiment. It appears therefore, that not only are P3 and P4 functionally similar, they also have similar scalp distributions ($P_Z>C_Z>F_Z$) when separated from the confounding effects of the overlapping SW.

The effects of emotional value on the P3 and P4 factors can be seen in the categorical differences in male and female subjects of the counting group (Figure 5). For the three totally independent categories which Miller (1985) found to be rated as pleasant (B), neutral (P), and unpleasant (D), there is a U-shaped function in P3 and P4 scores for both male and female subjects. Both positive and negative affect are reflected in larger P3 and P4 factor scores. For the slide categories showing sex differences in rating, both P3 and P4 scores also show sex differences: opposite sex models resulting in larger factor scores than same sex models. Together, these data provide strong support for the emotional value of a stimulus being a major factor determining P3 and P4 amplitudes.

It is somewhat puzzling that both P3 and P4 were sensitive to the same independent variable, emotional value. Why is there an apparent re-evaluation of the stimulus at P4, similar to that already observed at P3? Some indication of a possible difference between these two evaluations can be seen in Table 1. In this table, the factor scores for the five slide categories have been expressed as a function of the neutral category (P) factor scores.

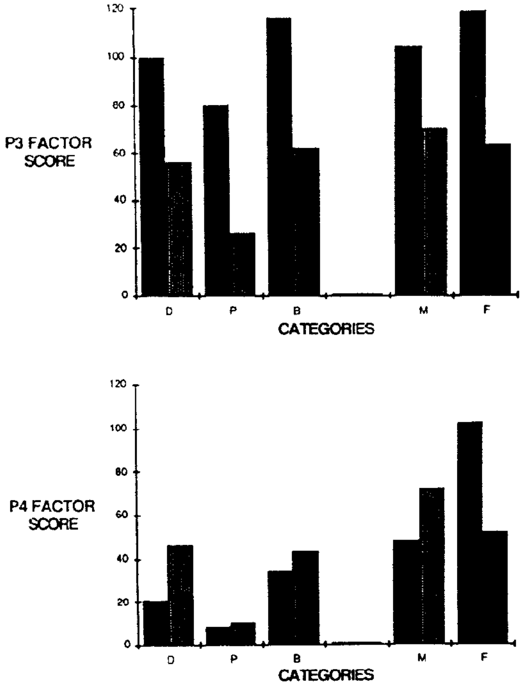

Figure 5. P3 and P4 factor scores to PICs in both male (black) and female (shaded) subjects, as a function of the five emotional slide categories: dermatological slides (D), people (P), babies (B), male models (M), and female models (F).

For all categories, both male and female subjects show relatively larger scores and a greater degree of discrimination between the categories at P4 compared with P3. This suggests that the generator for the late positive components, if they are indeed the same, is behaving more discriminantly at P4 than P3.

Unlike P3 and P4, the final SPW component was insensitive to both category and outcome variables; it also had a scalp distribution ($C_Z>F_Z>P_Z$) quite

Table 1

Factor scores for P3 and P4 factors for each emotional slide category, expressed as a function of the score to the neutral category (people)

Emotional Slide Categories	Factor Scores: Males P3	Males P4	Females P3	Females P4
Dermatological	1.2	2.6	1.1	4.6
People	1.0	1.0	1.0	1.0
Babies	1.4	4.2	2.4	4.3
Male Models	1.3	6.0	2.7	7.2
Female Models	1.5	12.7	2.4	5.2

different from the two earlier components. Several authors have noted a similar late positive slow wave (SW) having a parietally positive and frontally negative scalp distribution, and have proposed that it reflects a final evaluation process (Ruchkin, Sutton, & Stega, 1980; Ruchkin, Sutton, Kietzman, & Silver, 1980). This SW component has been shown to increase at both C_Z and P_Z electrode locations as task complexity increases. In the current study, the increase in the SPW at these locations in the learning group, compared with counting subjects, suggests a similarity between these two late positive slow waves, despite their different scalp distributions.

ERPs to CVCs

In a previous study of ERPs during a paired associate learning task, Johnston and Holcomb (1980) found that as subjects learned the relationship between the first (S1) and second stimulus (S2), then the P3 factor to S1 increased in amplitude. In the current study the P3 factor to CVCs decreased over sessions in both the learning and counting subjects; it was the P4 factor in the learning group which increased systematically over sessions (Session 1<2<3<4). This observation again supports the close functional similarity between the P3 and P4 factors; they also show similar scalp distributions in the CVC waveforms, $P_Z>C_Z=F_Z$ and $P_Z>C_Z>F_Z$ respectively. The learning group differed from the counting group, not only in P4 increasing over sessions, but also in the parietal dominance of this factor in the learning subjects.

Johnston and Holcomb (1980) also reported that when subjects learned the S1-S2 relationship, the P3 factor to S1 reflected the value of the S2 event. This relationship has been confirmed by Simons (1982) but was not observed for either P3 or P4 in the current study. The only significant category effect on P4 was due to male subjects exhibiting larger P4 factor scores than female subjects to CVCs predictive of female models; females exhibited larger P4 scores to all other CVCs in the learning subjects and to all CVCs in the counting group. In view of this trend, it seems reasonable to conclude that the failure to find a category effect on CVCs was due to the greater complexity of the current learning task compared with these previous studies. More sessions may be required to establish the relationship under these conditions.

The behavioral analysis revealed that male subjects learned the CVC-PIC combinations more rapidly than female subjects. An examination of the SPW in males revealed a steady decrease in this factor over the learning sessions. In contrast, the SPW in female subjects increased over the first three sessions and declined only in the final session. These observations are again consistent with the task complexity hypothesis (Ruchkin, Sutton, Kietzman, & Silver, 1980b) which predicts a large late positive component while stimuli require additional processing and a decrease in this component as the task demands are met.

Implications for a Model of Cognitive Processing

The amount of information received and the utility of a stimulus are generally accepted as the major factors determining P3 amplitude (Ruchkin & Sutton, 1978; Johnston, 1979; Johnson & Donchin, 1985). Sutton, Tueting, Hammer, and Hakerem (1978), and more recently Begleiter et al. (1983), have suggested that there may be a common factor underlying these two variables. If utility is redefined as the emotional value of a stimulus, then changes in the uncertainty of such events may also be considered to elicit emotional reactions, whose value is related to the degree to which the uncertainty is changed. Emotional value, then, may serve as the common currency determining the amplitude of late positive components of ERPs.

This hypothesis is consistent with an observation which is difficult to explain within any of the existing theoretical formulations. In the Johnston and Holcomb (1980) study cited above, one variable involved was the degree to which S1 was predictive of S2. All S1 events were equally probable, but the conditional probability of the high utility event (S2) following S1 varied between 0.5, 0.75, and 1.0. When subjects, who were initially uninformed of these relationships, had experience with the contingencies, the P3 to S1 reflected the conditional probability of S2 following S1, and not the probability of occurrence of S1, as might be expected. More surprising, however, was the observation that the lower the conditional probability of S2 following S1, the greater the P3 to the S1 event! If uncertainty reduction in the occurrence of S2 was the relevant variable then the reliable predictors (1.0 conditional probability), by virtue of having reduced more uncertainty, should have elicited the largest P3 amplitudes; exactly the opposite was true. The emotional value hypothesis, however, states that it is the change in uncertainty, either an increase or a decrease, which is the relevant variable governing the emotional reaction and the associated increase in P3. When subjects have learned the reliable relationships, then other S1s signal an increase in the uncertainty of S2, and P3 increases, compared with the certain condition.

Taken together, the CVC and PIC data suggest a general model of cognitive processing involving three independent overlapping serial stages. The first two stages are similar in nature and appear to involve two sequential activations of the same neural process. The primary activation appears to be governed by an emotional evaluation of the stimulus content. The second stage, which is temporally locked to the first, becomes more active when new associations are required; for example, following disconfirmed outcomes or when neutral events are becoming meaningful for the subject. The third phase of processing, which appears to be different in nature, corresponds to conditions when new hypothesis testing may be required.

The CVC-PIC learning task involves both an identification of the picture stimulus and a comparison of this stimulus to the predicted one. For learning to proceed, memory storage is required following each of these stages. The first activation of memory storage ensures that the identified stimulus is stored in the context in which it occurred (i.e., which CVC it followed); the reactivation, following disconfirmation, ensures that this new CVC-PIC association becomes more established at the expense of the incorrect association. It appears that it would be desirable to have both of these memory processes modulated by the importance (i.e., emotional value) of the events to the subject. It is therefore postulated that P3 and P4 reflect the activation of a memory storage mechanism regulated by an emotionally based value system.

Little is known concerning the neural processes underlying memory storage. One of the most promising mechanisms, however, is the observed long-term potentiation (LTP) of hippocampal neurons following bursts of stimulation (Swanson, Teyler, & Thompson, 1982). Such stimulation has been found to be most effective when the bursts are separated by about 200 ms (Larson, Wong, & Lynch, in press). More recent evidence suggests that a single burst applied to one input of hippocampal neurons will generate LTPs if the same neurons are stimulated by different input fibers at about 200 ms after the first excitation (Lynch, 1986). Since volume conduction of synchronous hippocampal activity is thought to be the origin of late positive components of ERPs (Halgren et al., 1980), the P3 and P4 factors may reflect dual activations of the hippocampus at the optimal rate for generating LTPs. From this perspective, the first activation, at P3, would facilitate the association between any ongoing activity and the current event; the second activation, at P4, would further enhance the LTP induced by the prior activation. This second activation, at the optimal rate for generating LTPs, would ensure that the new CVC-PIC association would become more established at the expense of any prior incorrect associations.

REFERENCES

Begleiter, H., Porjesz, B., Chou, C.L., & Aunon, J.I. (1983). P3 and stimulus incentive value. *Psychophysiology, 20*, 95–101.

Greenhouse, S.W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika, 24*, 95–112.

Halgren, E., Squires, N.K., Wilson, C.L., Rohrbaugh, J.W., Babb, T.L., & Crandall, P.H. (1980). Endogenous potentials generated in the human hippocampal formation and amygdala by infrequent events. *Science, 210*, 803–805.

Johnson, R., Jr., & Donchin, E. (1985). Second thoughts: Multiple P300s elicited by a single stimulus. *Psychophysiology, 22*, 182–194.

Johnston, V.S. (1979). Stimuli with biological significance. In H. Begleiter (Ed.), *Evoked brain potentials and behavior* (pp. 1–12). New York: Plenum Press.

Johnston, V.S., & Holcomb, P.J. (1980). Probability learning and the P3 component of the visual evoked potential in man. *Psychophysiology, 17*, 396–400.

Larson, J., Wong, D., & Lynch, G. (in press). Patterned stimulation at the theta frequency is optimal for the induction of hippocampal long-term potentiation. *Brain Research*.

Lynch, G. (1986, February). *Organization, activity and learning in simple cortical structures*. Paper presented at the Conference on Biological Computation, Tucson, AZ.

Miller, D. (1985). *Some electrophysiological correlates of affect and memory*. Unpublished doctoral dissertation, New Mexico State University, Las Cruces, NM.

Myers, J.L. (1979). *Fundamentals of experimental design* (3rd ed.). Boston: Allyn and Bacon.

Noble, C.E. (1961). Measurements of association value, rated associations and scaled meaningfulness for the 2100 CVC combinations of the English alphabet. *Psychological Reports, 8*, 487–521.

Ruchkin, D.S., & Sutton, S. (1978). Equivocation and P300 amplitude. In D. Otto (Ed.), *Multidisciplinary perspectives in event related brain potential research* (pp. 175–177) (EPA-60019-77-043). Washington, DC: U.S. Government Printing Office.

Ruchkin, D.S., Sutton, S., Kietzman, M.L., & Silver, K. (1980). Slow wave and P300 in signal detection. *Electroencephalography & Clinical Neurophysiology, 50*, 35–47.

Ruchkin, D.S., Sutton,S., & Stega, M. (1980). Emitted P300 and slow wave event-related potentials in guessing and detection tasks. *Electroencephalography & Clinical Neurophysiology, 49*, 1–14.

Simons, R.F. (1982). Physical anhedonia and future psychopathology: An electrocortical continuity? *Psychophysiology, 19*, 433–441.

Stuss, D.T., & Picton, T.W. (1978). Neurophysiological correlates of human concept formation. *Behavioral Biology, 23*, 135–162.

Stuss, D.T., Toga, A., Hutchison, J., & Picton, T.W. (1980).

Feedback evoked potentials during an auditory concept formation task. In H.H. Kornhuber & L. Deecke (Eds.), *Motivation, motor and sensory processes of the brain. Progress in brain research* (Vol. 54, pp. 403–409). Amsterdam: Elsevier.

Sutton, S., Tueting, P., Hammer, M., & Hakerem, G. (1978). Evoked potentials and feedback. In D. Otto (Ed.), *Multidisciplinary perspectives in event related brain potential research* (pp. 184–188) (EPA-60019-77-043). Washington, DC: U.S. Government Printing Office.

Swanson, L.W., Teyler, T.J., & Thompson, R.F. (Eds.). (1982). *Hippocampal long-term potentiation: Mechanisms and implications for memory. Neurosciences Research program bulletin* (Vol. 20). Cambridge, MA: MIT Press.

(Manuscript received March 11, 1986; accepted for publication July 19, 1986)

Coming Soon:

Manuscripts in Press, *Psychophysiology* (in alphabetical order by author, accepted at press time)

"Cooperative dyadic interactions, perceived control, and task difficulty in Type A and Type B individuals: A cardiovascular study," by Janice Abbott, Colin Sutherland, and Douglas Watt.

"Shock threat and sexual arousal: The role of selective attention, thought content, and affective states," by J. Gayle Beck, David H. Barlow, David K. Sakheim, and Daniel J. Abrahamson.

"Correcting psychophysiological measures for individual differences in responsivity should be based on typical response parameters: A reply to Stemmler," by Gershon Ben-Shakhar.

"An event-related potential analysis of incongruity in music and other non-linguistic contexts," by M. Besson and F. Macar.

"Individual differences in stages 3 and 4 sleep," by Donald L. Bliwise and Bernard M. Bergmann.

"Efficiency of coping with a real-life stressor: A multimodal comparison of aerobic fitness," by Sam T. Brooke and Bonita C. Long.

"Heart rate and oxygen consumption during mental arithmetic, a video game, and graded static exercise," by Douglas Carroll, J. Rick Turner, and Stephen Rogers.

"A data based method for bivariate outlier detection: Application to automatic blood pressure recording devices," by L.A. Clark, L. Denby, D. Pregibon, G.A. Harshfield, T.G. Pickering, S. Blank, and J.H. Laragh.

"The effect of piracetam upon visual event-related potentials in dyslexic children," by C.K. Conners, M. Reader, J. Caldwell, R. Clymer, A Reiss, L. Caldwell, A. Adesman, L. Mayer, and M. Berg.

"Effects of arousal on habituation of electrodermal versus vasomotor responses," by Bram C. Goldwater.

"The effects of caffeine and stress on blood pressure in individuals with and without a family history of hypertension," by Walter Greenberg & David Shapiro.

"Effects of auditory stimulation on respiration," by Andrew Harver and Harry Kotses.

"The sleeping rat shows signs of orienting response to neutral stimuli," by Marga Johnen.

"Cardiovascular responses to stress: II. Relationships to aerobic exercise patterns," by Kathleen C. Light, Paul A. Obrist, Sherman A. James, and David S. Strogatz.

"Cardiovascular responses to stress: I. Measures of myocardial response and relationships to high resting systolic pressure and parental hypertension," by Paul A. Obrist, Kathleen C. Light, Sherman A. James, and David S. Strogatz.

"Comparison of P300 from a passive tone sequence paradigm and an active discrimination task," by John Polich.

"Contingent negative variation, P3 evoked potentials, and antisocial behavior," by Adrian Raine and Peter H. Venables.

"Functional differences between members of the P300 complex: P3e and P3b," by D.S. Ruchkin, S. Sutton, and D. Mahaffey.

"Event related desynchronization and P300," by Joseph Sergeant, Reint Geuze, and Wim van Winsum.

"Is mediation of sweating cholinergic, adrenergic, or both?" by Stephanie A. Shields, Kathleen A. MacDowell, Sarah B. Fairchild, and Mary L. Campbell.

"Comparison of discrete cardiovascular fitness groups on plasma catecholamine and selected behavioral responses to psychological stress," by Mark S. Sothmann, Thelma S. Horn, Barbara A. Hart, and Anthony B. Gustafson.

"Standardization within subjects: A critique of Ben-Shakhar's conclusions," by Gerhard Stemmler.

"Electrogastrography: Current issues in validation and methodology," by Robert M. Stern, Kenneth L. Koch, William R. Stewart, and Michael W. Vasey.

"The proximal effects of two and four hour naps within extended performance without sleep," by Wilse B. Webb.

"The effects of psychological and environmental stressors on peristaltic esophageal contractions in healthy volunteers," by Larry D. Young, Joel E. Richter, Karen O. Anderson, Laurence A. Bradley, Philip O. Katz, Lisa McElveen, W.F. Obrecht, Chris Dalton, and R. Miller Snyder.