Psychophysiology, 37 (2000), 92–101. Cambridge University Press. Printed in the USA.

Temporal stability of the emotion-modulated startle response

CHRISTINE L. LARSON, DANTE RUFFALO, JENNIFER Y. NIETERT, AND RICHARD J. DAVIDSON
Department of Psychology, University of Wisconsin—Madison, USA

Abstract

In the present study, we examined the stability of one measure of emotion, the emotion-modulated acoustic startle response, in an undergraduate sample. Using the acoustic startle paradigm on two different occasions, we measured stability of affective modulation of the startle response during and following the presentation of pictures selected to be of positive, negative, or neutral emotional valence. The two assessments were separated by 4 weeks. Two groups of subjects were compared: one group that viewed the same pictures at each assessment and a second group that viewed different pictures at the second assessment. We found that viewing different pictures at two assessments separated by 4 weeks yielded moderate stability of the emotion modulation of startle magnitude, whereas subjects who viewed the same pictures at both assessments showed poor stability. Furthermore, this difference was due to the stability of responses to high versus low arousal pictures, not to differences in valence.

Descriptors: Emotion, Startle response, Affective modulation, Psychometrics

Research on trait aspects of emotion has long been plagued by a lack of objective indices. In recent years, researchers have begun to use the emotion-modulated startle paradigm as an objective measure of state affect. Experiments with the startle reflex have shown that the magnitude of the eyeblink component of the response is modulated by the affective valence of a foreground stimulus (Bradley, Cuthbert, & Lang, 1993; Vrana, Spence, & Lang, 1988). Lang, Bradley, and Cuthbert (1990) suggested that this modulation is based on the match or mismatch of the motivational systems that subserve both the reflex and higher-level emotional processes. Lang and colleagues have outlined a model based on two major motivational systems, appetitive and aversive. In the framework of these two dimensional motivation systems, the startle reflex is viewed as an aversive or defensive response. Thus, they postulate that the magnitude of the startle eyeblink reflex would be enhanced during presentation of a negatively valenced foreground stimulus because the defensive reflex matches the ongoing affective state. Conversely, the eyeblink reflex is attenuated in the presence of a positively valenced foreground stimulus, as there is a mismatch between the ongoing affective state and the aversive nature of the startle reflex. This effect has been replicated reliably and robustly using various stimulus modalities, such as pictures (Bradley, Cuthbert, & Lang, 1993), odors (Miltner, Matjak, Braun, Diekmann, & Brody, 1994), and mental imagery (Cook, Hawk, Davis, & Stevenson, 1991), and has also been demonstrated in infants as young as 5 months of age (Balaban, 1995).

Although the emotion-modulated startle paradigm has been most frequently used as an index of state affect, several researchers have recently extended the use of this index to the measurement of trait patterns of affective reactivity. For example, Cook and colleagues (Cook, Davis, Hawk, Spence, & Gautier, 1992; Cook et al., 1991) have found enhanced startle magnitude to negative stimuli among individuals who scored high on trait measures of fearfulness relative to subjects who scored low on fearfulness. Other investigators have begun to examine differences in the emotion modulation of startle among individuals with personality disorders, such as psychopaths (Patrick, Bradley, & Lang, 1993). However, before widely accepting emotion-modulated startle as an objective measure of trait-like patterns of emotional responding it is important to establish the stability of the emotion modulation of startle over time, in other words, the test–retest reliability. Whereas the test–retest reliability of the startle reflex (e.g., Jennings, Dawson, Schell,

A portion of these data was presented at the 38th Annual Meeting of the Society for Psychophysiological Research, Denver, Colorado, September 23–26, 1998.

Support for this research was provided by NIMH grants MH40747, MH43454, NIMH Center grant P50-MH52354 to the Wisconsin Center for Affective Science, a grant from the John D. and Catherine T. MacArthur Foundation, an NIMH Research Scientist Award to R. J. Davidson (K05-MH00875), and an NRSA Predoctoral Fellowship Award to C. L. Larson (F31-MH12085).

We thank Steven K. Sutton, Adrian Pederson, and Darren Dottl for their assistance in implementing the study.

Address reprint requests to: Richard J. Davidson, University of Wisconsin, Department of Psychology, 1202 West Johnson Street, Madison, WI 53706, USA. E-mail: rjdavids@facstaff.wisc.edu.

Earleywine, & Runyan, 1994) and prepulse inhibition of the startle reflex (Schwarzkopf, McCoy, Smith, & Boutros, 1993) has been measured previously, there have been no published reports of the stability over time of the modulation of startle magnitude by a foreground emotional stimulus.

Thus, the major goal of this study was to examine the stability of the emotion-modulated startle response over time. Subjects participated in an emotion-modulated acoustic startle paradigm on two occasions separated by 4 weeks. Affectively laden pictures were used as emotion-eliciting stimuli. Whereas other studies have repeated still picture stimuli within the same experimental session and have demonstrated emotion modulation of the eyeblink reflex to the repeated stimuli (Bradley, Lang, & Cuthbert, 1993; Sutton, Davidson, Donzella, Irwin, & Dottl, 1997), reliability of emotion modulation in response to repeated stimuli was not examined. Furthermore, there have been no reports of the effects of presenting the same pictures on two different occasions. In a previous study conducted in our laboratory, subjects completed two emotion-modulated startle sessions separated by 4 weeks in which the identical pictures were shown at each assessment (Larson, Davidson, Sutton, Ruffalo, & Nietert, 1999). The data from this study revealed very low test–retest reliability of the emotion-modulated startle response. In light of these data, we hypothesized that one cause for the lack of stability may have been the use of the same pictures at the two different assessments. This hypothesis was based on the idea that the subjects' familiarity with the pictures may lead to a decreased overall emotional impact for both negative and positive pictures at the second assessment, and therefore to diminished test–retest reliability. An alternative hypothesis for the negative pictures is that whereas the initial emotional reaction to the same picture presented on two different occasions may be similar, subsequent processing may differ when subjects see the same picture the second time because they may be more readily able to invoke emotion regulation strategies to suppress negative affect. In light of these two hypotheses and based on our prior data, we predicted that viewing the same pictures at the two test assessments would lead to low test–retest reliability. Thus two groups of subjects were tested: one group that viewed the same pictures at the two assessments and a second that viewed a different set of pictures matched on valence and arousal to the first set. We predicted that subjects viewing different pictures would show greater test–retest reliability than subjects viewing the same pictures. In addition to measuring startle magnitude, we also measured startle latency in light of the attention that latency measures have received in the clinical neurophysiological literature on the eyeblink reflex (Kimura, 1992).

Methods

Participants

Participants were 71 undergraduates (55 women, 16 men). The subjects were recruited from Introductory Psychology classes at the University of Wisconsin—Madison. All participants were between 17 and 20 years of age. Subjects were assigned randomly to one of two groups, a group that viewed the same pictures at assessments 1 and 2 (Same group, $n = 34$), and a group that viewed different pictures at the two assessments (Different group, $n = 37$). Due to an insufficient number of startle responses at one or both assessments (see below), 22 subjects were dropped, yielding a final sample of 49 subjects (26 Different: 23 women, 3 men; 23 Same:, 14 women, 9 men).

Materials

Picture stimuli designed to elicit positive, negative, or neutral emotions were chosen from Shows 1 through 10 of the International Affective Picture System (Center for the Study of Emotion and Attention, 1994). Picture selection was based on published self-report ratings of valence and arousal (Lang, Greenwald, & Bradley, 1993). Using a method used previously by our laboratory (Sutton et al., 1997), pictures were selected to ensure that the negative and positive pictures were both high on arousal, but opposite in valence, and that the neutral pictures were low on arousal and average on valence. The 42 pictures with the most similar average ratings for men and women were selected for each of three valence categories: positive, negative, and neutral. To best match the two genders on average valence and arousal ratings, slightly different picture sets were chosen for men and women. For the neutral and negative valences, the pictures were the same for each gender. For the positive pictures, 36 of the 42 pictures were used for both men and women. The 6 pictures that did not overlap between the genders depicted opposite sex nudes or heterosexual couples (see Appendixes A and B for complete list of pictures used).

Half of the subjects saw the same pictures at both assessments and half saw different pictures at the second assessment. Two different sets of 63 pictures (Set A and Set B), including 21 pictures of each valence were matched on arousal ratings. Table 1 lists the mean male and female valence and arousal ratings separately for Sets A and B.

Procedure

Subjects voluntarily participated in the study in exchange for extra credit points for their Introduction to Psychology class. Before participation, subjects were contacted by phone and informed of the procedures. Two sessions separated by 4 weeks were conducted, each at the same time of day. Each of the sessions followed the same procedure.

Upon arrival at the laboratory, the subject was seated in a comfortable chair approximately 1 m from a 43.52-cm NEC-6FG multi-sync monitor upon which pictures were displayed. Informed consent was obtained, including a reminder that unpleasant pictures were presented during the experiment. Following a brief overview of the procedure, subjects completed a set of questionnaires to measure state and trait affect. Electrodes for recording startle responses were placed and impedances checked. Before the

Table 1. *Mean (SD) Male and Female Valence and Arousal Ratings for Picture Sets A and B*

	Picture set A		Picture set B	
	Male	Female	Male	Female
Valence				
Negative	2.40 (.50)	1.70 (.33)	2.43 (.49)	1.75 (.41)
Neutral	4.98 (.20)	5.15 (.22)	4.96 (.28)	5.08 (.26)
Positive	7.45 (.42)	7.36 (.46)	7.35 (.37)	7.41 (.54)
Arousal				
Negative	6.44 (.44)	6.99 (.41)	6.42 (.35)	7.04 (.37)
Neutral	2.72 (.41)	2.90 (.50)	2.62 (.56)	2.84 (.52)
Positive	6.45 (.63)	6.12 (.48)	6.19 (.47)	5.96 (.45)

Note: These ratings are from Lang, Greenwald, & Bradley (1993).

picture presentation, the subject viewed an introductory set of 11 neutral pictures and received 9 startle probes to orient them to the procedure and habituate the subjects to the startle probe.

The subject then viewed the pictures for 21 min. Presentation of pictures and acoustic startle probes were controlled by in-house software on a 100-MHz Pentium PC. Pictures were presented in three blocks of 21 pictures, with 7 pictures of each valence included in each block. Pictures were presented in a quasi-random order, with the constraint that not more than two stimuli of a given valence were presented consecutively. Each picture was presented for 6 s with a randomized 10–18-s intertrial interval (ITI) (mean ITI was 14 s). The acoustic startle probes were a white noise burst 50 ms in duration, 95 dB, and with a nearly instantaneous rise time. Startle probes were generated with a Coulbourn S81-02 noise generator and a Coulbourn S82-24 audio-mixer power amplifier (Leheigh Valley, PA, USA), and were delivered binaurally through Radio Shack Optimus LV-20 headphones. Three probe times were used: 1.5, 4.5, and 7.5 s following picture onset.[1] For each of the three probe times, probes occurred during six trials of each valence. Three trials per valence did not contain any startle probes. Probe times were quasi-randomly assigned for each trial with the constraint that no more than two of each probe time occurred consecutively.

Before the first picture in each of the three blocks a startle probe was presented. Previous studies in our laboratory have found that the first response in a given block is often of greater magnitude than the remaining responses and may need to be dropped (Sutton et al., 1997). Therefore, these extra probes were added to aid in habituating the subject at the beginning of each block.

For the subjects in the Same group, the pictures were presented in a different order at Assessment 2 than at Assessment 1. The counterbalanced order of picture contents and probe times for the Different group was matched to that of the group that saw the same pictures.

Startle Recording and Quantification

Raw and integrated electromyograms (EMG) from the orbicularis oculi were collected using two Rochester Electro-Medical (Tampa, FL, USA) mini-electrodes placed directly below the left eye (Vrana et al., 1988). The impedance for the electrode pair was less than 20,000 Ω. Using SAI Bioelectric amplifiers (SA Instrumentation Co., Caroga Lake, NY), EMG signals were passed through bandpass filters set at 1 and 800 Hz and then amplified 10,000 times. After passing through a Rockland high pass filter set at 30 Hz, raw EMG signals were integrated and rectified using a Coulbourn S76-01 contour following integrator with the time constant set at 20 ms. All signals were digitized and stored at 250 Hz on a 100-MHz Pentium PC throughout the picture presentation using SnapStream software (HEM Data Corporation, Springfield, MI) and a 12-bit analog-to-digital board (Analogic Corporation, Wakefield, MA). Recording equipment was calibrated before and after each session. The units for raw and integrated EMG were microvolts (μV).

Orbicularis oculi EMG in response to acoustic startle probes was reduced to eyeblink reflex magnitudes using the following procedure. First, automatic peak and blink onset detection was performed on the integrated EMG response to each probe using an in-house software package. Each response was then reviewed by laboratory personnel. Approximately 14.7% of eyeblink reflexes were excluded from further analyses due to excessive noise during a 50-ms, prestartle baseline period (e.g., spontaneous blinks, unusually high amounts of integrated EMG during baseline) or because the onset of the integrated EMG eyeblink reflex began less than 20 ms following the startle probe. Eyeblink reflex magnitudes (in μV) were calculated by subtracting the amount of integrated EMG at reflex onset from the peak amplitude (maximum amount of integrated EMG between 20 and 120 ms following probe onset). Trials with no perceptible eyeblink reflex were assigned a magnitude of zero and included in analysis. Finally, eyeblink reflex magnitudes were *z*-transformed within subjects and within assessment due to large individual differences in the distribution of this measure. Blinks that were more than 3 standard deviations (*SD*) above the mean for a given subject were excluded.

At the first assessment, 14 subjects did not display an eyeblink response to acoustic startle probes or more than half of the trials were bad (see above), yielding a sample size of 57 for the first assessment. Sixty subjects returned for the second assessment. Of these 60, 5 subjects did not respond to the startle probes or were dropped because more than half of the trials were bad, leaving a sample size of 55 for the second assessment. In addition, of the 6 trials per cell (Probe Time × Picture content × Assessment) subjects were required to have at least 3 responses that were not excluded from the analyses for reasons listed above ("good" responses). To ensure that data from all three probe times contributed to the average for all three picture categories and both assessments, analyses were restricted to subjects with at least three good responses in every cell. This yielded a final sample of 49 subjects (26 Different: 23 women, 3 men; 23 Same: 14 women, 9 men)[2]. Finally, mean blink magnitude and latency for each picture valence category was calculated by averaging across all three probe times.

Self-Report Measures

At the beginning of each experimental session, a set of questionnaires was administered to assess affect and mood and anxiety symptoms. The Behavioral Activation/Inhibition Scales (BAS/BIS; Carver & White, 1994; coefficient alpha = .72 and .73 for present sample) were administered to examine activation of these two biobehavioral systems described by Gray (1994). The BAS scale is comprised of three subscales: Drive, Reward Responsivity, and Fun Seeking. The Positive and Negative Affect Schedule was administered in its trait version (PANAS-Trait; Watson, Clark, & Tellegen, 1988; coefficient alpha = .80 for trait positive affect, and .81 for trait negative affect among the present sample).

Data Analysis

Replication of emotion-modulated startle effect. To determine whether or not the valence-dependent modulation of startle occurred, a three-way repeated measures analysis of variance (ANOVA) on blink magnitude was performed, with Picture Content

[1] These three probe times were chosen based on a previous study conducted in our laboratory that focused on measuring the time course of affective responses to the pictures. A major goal of the prior study was to examine affective reactions following the offset of the stimulus compared with reactions during the stimulus presentation. Thus, the 7.5-s probe time was chosen to assess affective reactivity following picture offset. The 1.5- and 4.5-s probes were intended to allow for comparison of the degree of emotion modulation of startle responding following picture offset with these two mid-picture probe times. In the current study we aimed to examine the stability of different probe times. However, the stability for each probe time was low, likely due to a low number of trials per cell, thus the average of the three probe times are the only data presented here.

[2] The power to detect a .5 correlation (moderate stability) at a *p*-value of .05 two-tailed is .76 for the Different group (n = 26) and .71 for the Same group (n = 23).

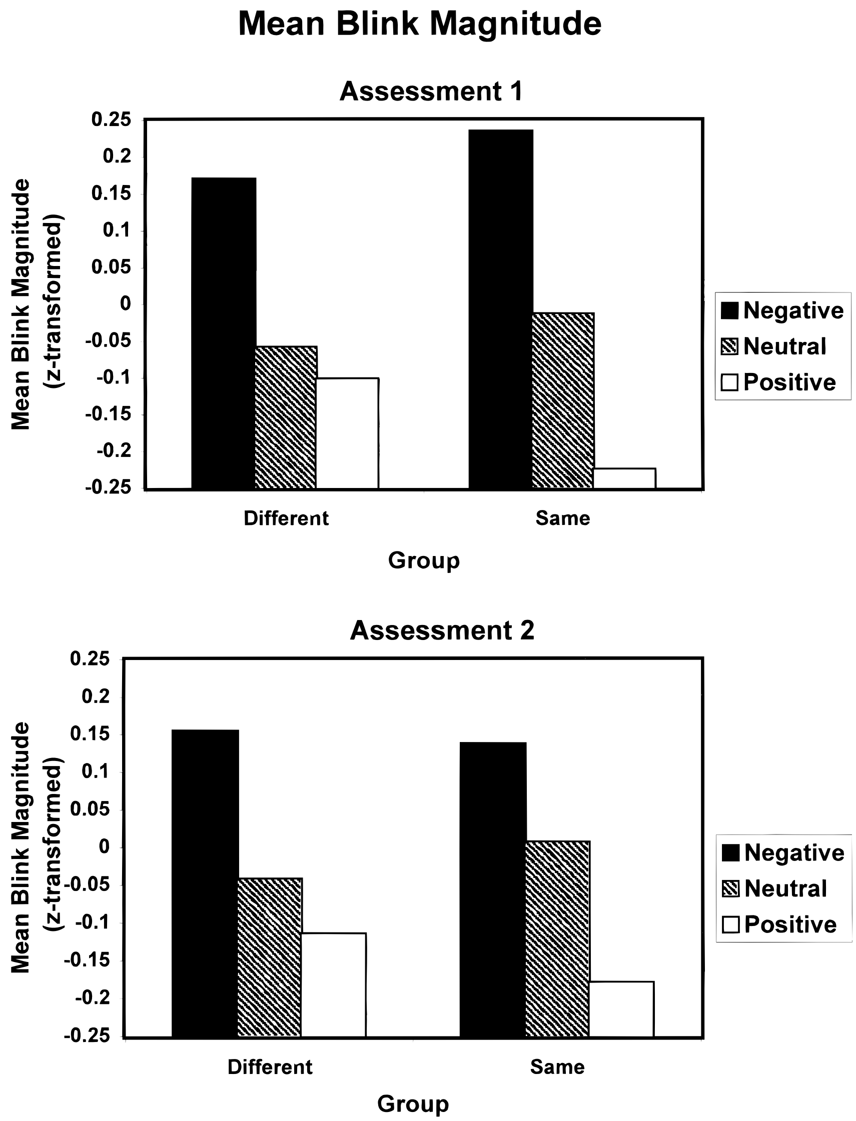

Figure 1. Standardized (*z*-score) blink magnitude averaged across all three probe times for the Same and Different subject groups at each assessment.

(negative, neutral, and positive) and Assessment as the repeated measures factors, and Group (Same, Different) as the between-groups factor. A second analogous ANOVA was calculated for startle blink latencies[3]. All ANOVAs using a within-subjects factor used a Huynh–Feldt correction (Huynh & Feldt, 1970).

Stability of emotion-modulated startle. The second question to be examined was the test–retest stability of the emotion-modulated startle measures. First, test–retest reliability was examined for raw blink magnitude averaged across all trials to examine stability of the startle reflex itself. Second, to examine the test–retest stability of the emotion modulation of the startle response, using the *z*-transformed data correlations were performed for potentiation (negative − neutral) and attenuation (positive − neutral)[4]. Although these two contrasts are useful in comparing positive and negative pictures against neutral, they confound the effects of arousal and valence (Lang, Greenwald, Bradley, & Hamm, 1993; Vrana et al., 1988). Thus, a second set of correlations was performed to differentiate between the contributions of valence and arousal to the stability of the emotion-modulated startle response. Two different contrasts were computed. First, to compare high versus low arousal pictures, the mean of blink magnitudes for negative and

[3]Latency data have been included because others have reported on the influence of the emotional foreground on the latency of the startle response elicited by background probes (e.g., Cook et al., 1992). However, it must be noted that when designing this study our primary focus was on blink magnitude, not latency. In light of this focus, we elected to use a 250-Hz sampling rate, which substantially diminishes the resolution of the latency data. Thus, this major limitation should be kept in mind when interpreting the latency effects.

[4]Pearson correlations were computed rather than intraclass correlations due to our concerns that habituation across sessions may reduce the magnitude of the intraclass correlations. The main question was whether the relative difference between conditions (e.g., negative and neutral) was stable, regardless of whether or not the means shifted across assessments.

positive pictures minus the magnitude for neutral pictures was computed. Second, a negative − positive contrast was computed to examine the effects of valence independent of arousal. Test–retest correlations were then performed using these two contrasts. A set of identical correlations was conducted on the blink latency data.

Relations between emotion-modulated startle and self-report measures. The four blink magnitude contrasts described above were correlated with the BIS/BAS scales and PANAS-Trait positive and negative affect. To determine whether aggregating across assessment better predicted measures of affect, averages across assessment were computed for each contrast and then correlated with self-report measures separately for each group. A parallel set of correlations was performed for the blink latency data.

Results

Emotion-Modulation Effects

Blink magnitude. A three-way ANOVA with Group as the between-subjects factor (Same, Different), and Assessment (1, 2), and Picture Content (negative, neutral, positive) as within-subjects factors revealed a significant main effect for Picture Content, $F(2,94) = 19.15$, $p = .0001$. Follow-up t tests revealed that across assessment and group, blink magnitudes were greater for negative than neutral pictures, $t(1,48) = 3.52$, $p = .001$ (Figure 1). Also, blink magnitudes were significantly smaller for positive than for neutral pictures, $t(1,48) = 6.63$, $p = .0001$. There were no main effects for Group or Assessment (ps $> .10$). Also, there were no significant interactions (ps $> .25$).

Blink latency. Mean blink latency for each picture valence was calculated. A three-way Group × Assessment × Picture content ANOVA revealed no significant main effects or interactions. The main effect for Picture content approached significance, $F(2,94) = 2.67$, $p = .085$ (Table 2). Follow-up t tests revealed that blink latencies were significantly shorter during the negative compared with neutral ($t = 2.16$, $p < .05$) and positive pictures ($t = 2.70$, $p < .01$; see Table 2 for means).

Stability of Emotion Modulation of the Startle Response

Stability of raw data. Raw blink magnitudes (in μV) for each valence were highly correlated across the two assessments for both groups (rs separately for group and picture valence range from .79 to .90, ps $< .001$). There were no differences in these correlations between groups or valences. Thus, mean blink magnitude across all trials was computed for each assessment and a test–retest correlation performed for both groups combined ($r = .89$, $p = .0001$, see Figure 2).

Stability of emotion modulation. Stability of emotion modulation of startle was calculated by correlating potentiation and attenuation difference scores (using z-transformed data) at the first assessment with these same scores from the second assessment. For negative − neutral valences, subjects who saw different pictures showed moderate stability ($r = .55$, $p < .004$), compared with subjects who viewed the same pictures ($r = .21$, $p = ns$; see Figure 3). The difference between these two correlations approached significance ($z = 1.32$, $p = .10$). A similar pattern was seen for the positive − neutral contrast. Subjects who viewed different pictures showed greater test–retest reliability (for different group, $r = .44$, $p < .03$, for same group, $r = -.09$, $p = ns$; see Figure 4). The difference between these two correlations approached significance ($z = 1.57$, $p = .06$).

Whereas the comparisons of positive and negative to responses to neutral pictures have been examined often, these contrasts confound valence and arousal differences among the picture categories (Lang et al., 1993; Vrana et al., 1988). Thus, a second set of correlations was performed to disentangle the contributions of arousal (mean[negative, positive] − neutral) and valence (negative − positive). For the arousal contrast the group that viewed different pictures showed greater reliability ($r = .62$, $p < .001$) then the group that viewed the same pictures ($r = .04$, $p = ns$; see Figure 5). These two correlations were significantly different ($z = 2.24$, $p < .01$). Neither group showed reliability across the two testing sessions for the negative − positive contrast (for different group, $r = .25$, $p = ns$; for same group, $r = .22$, $p = ns$; see Figure 6).

Stability of latency data. Test–retest reliability was computed for each picture category. For both groups, all categories showed

Table 2. *Mean (SD) Startle Blink Onset Latency (in ms)*

	Same picture group		Different picture group	
	Time 1	Time 2	Time 1	Time 2
Negative	39.5 (6.6)	41.4 (6.7)	39.1 (5.1)	38.3 (5.6)
Neutral	40.5 (6.7)	42.1 (7.1)	40.5 (6.5)	39.8 (5.4)
Positive	40.3 (6.8)	40.5 (10.8)	40.2 (5.1)	39.7 (5.6)

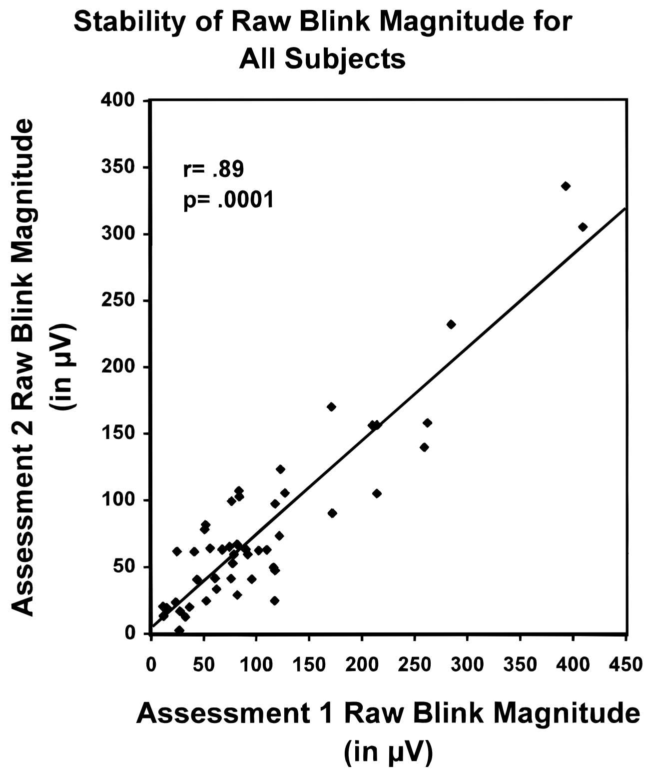

Figure 2. Scatter plot of the relation between raw blink magnitudes averaged across all probe times and picture valences for assessments 1 and 2. Subjects from both groups are included ($N = 49$).

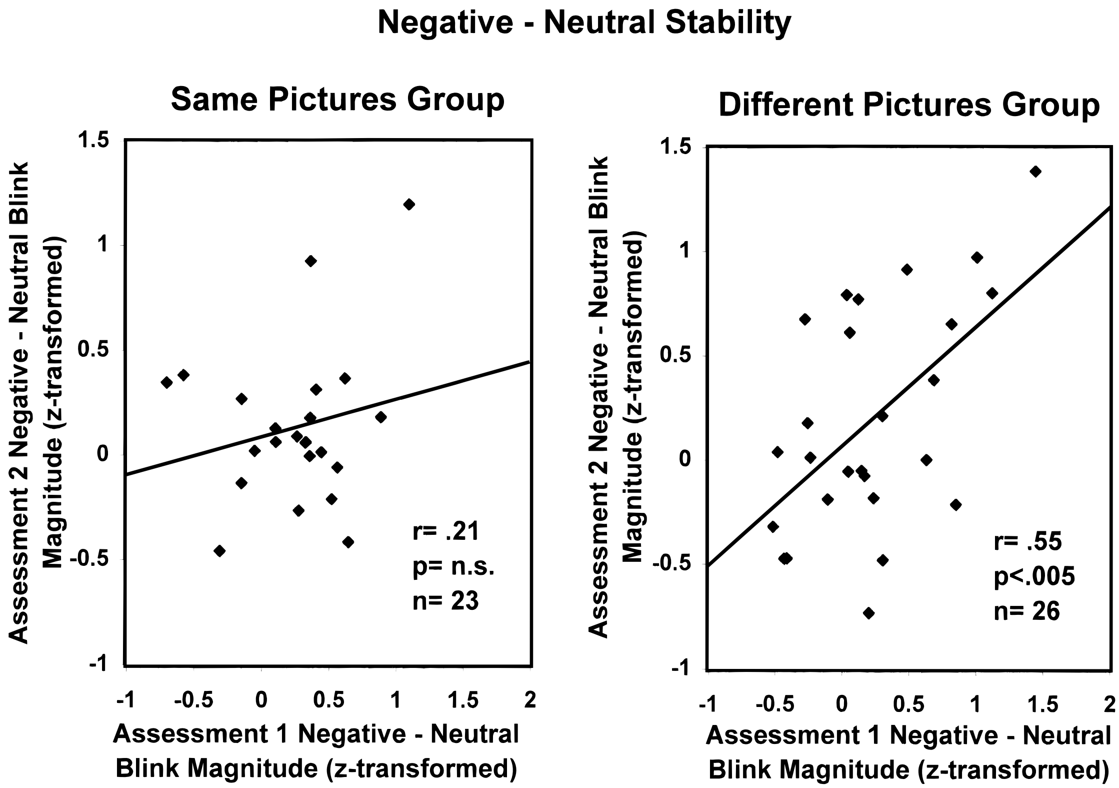

Figure 3. Scatter plot of the relation between negative-neutral (potentiation) standardized blink magnitude (averaged across probe time) for the Same and Different subject groups at each assessment.

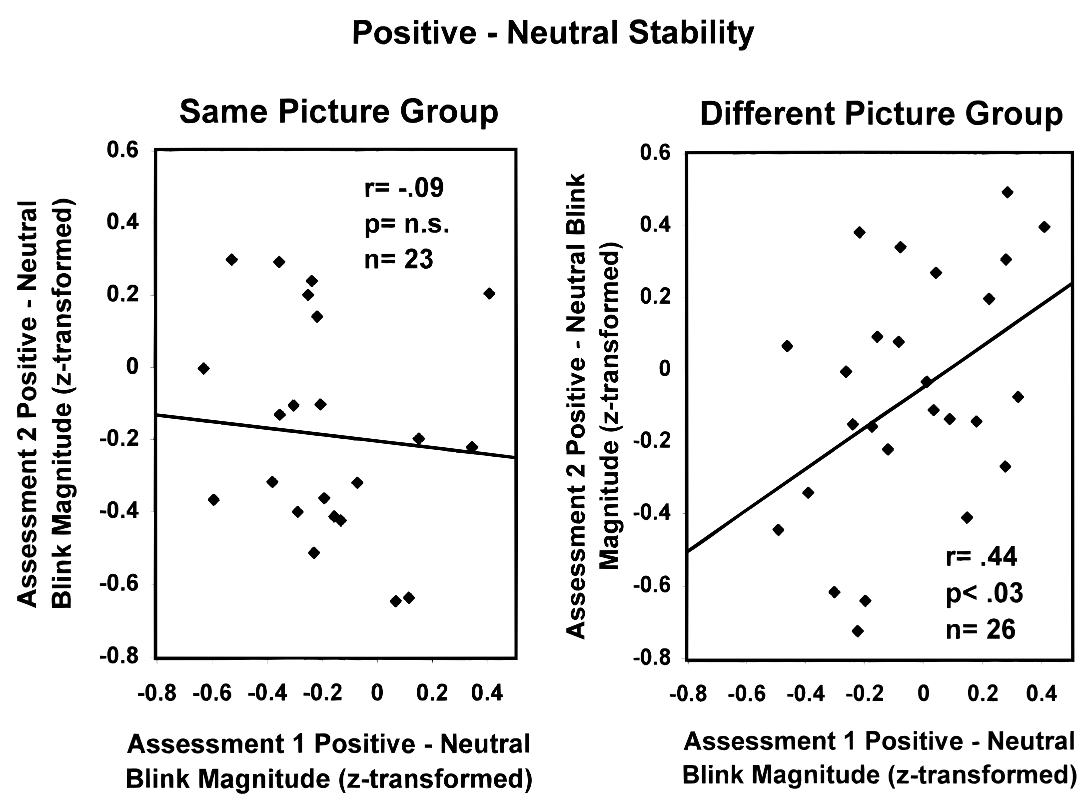

Figure 4. Scatter plot of the relation between positive-neutral (attenuation) standardized blink magnitude (averaged across probe time) for the Same and Different subject groups at each assessment.

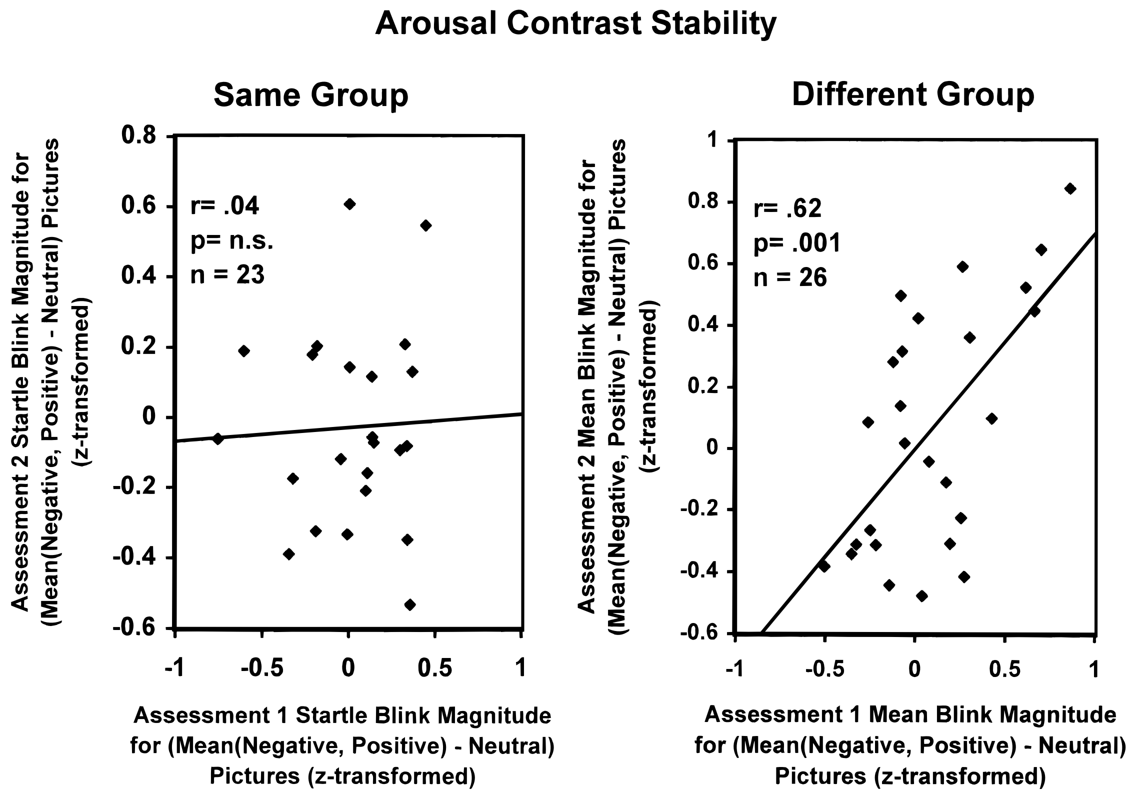

Figure 5. Scatter plot of the relation between the arousal contrast metric (mean[negative, positive] − neutral standardized blink magnitude) for the Same and Different subject groups at each assessment.

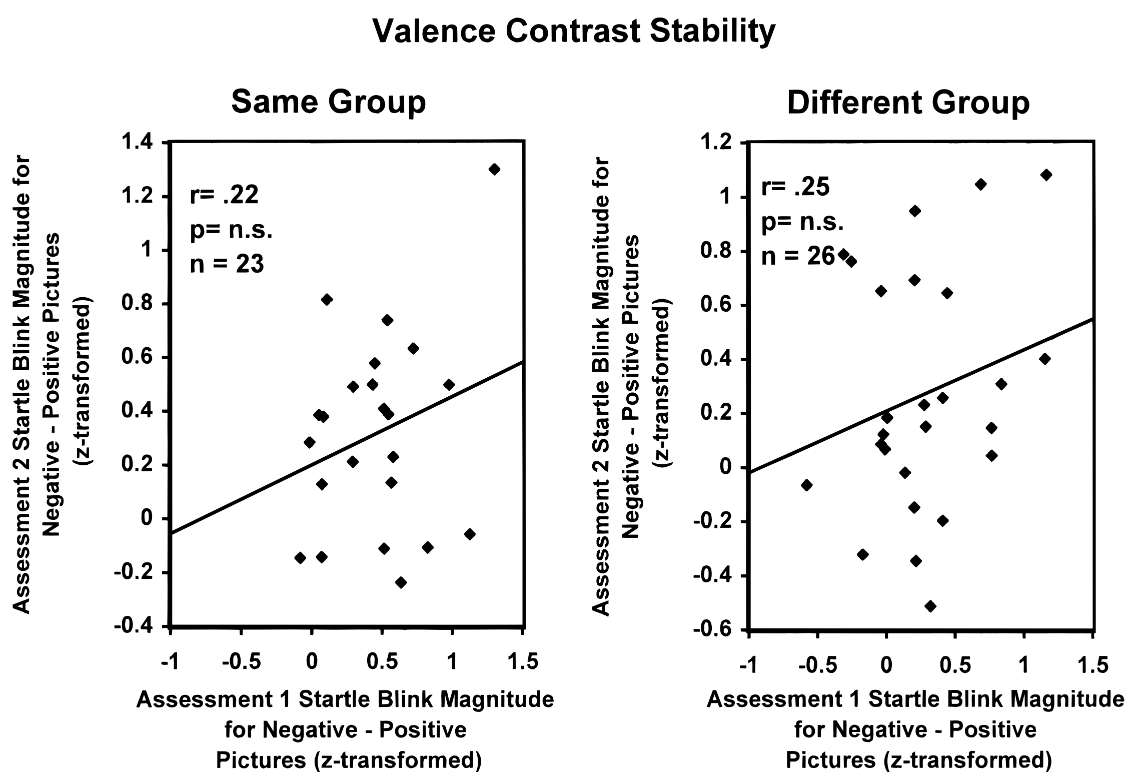

Figure 6. Scatter plot of the relation between the valence contrast metric (negative − positive standardized blink magnitude) for the Same and Different subject groups at each assessment.

Table 3. *Test–Retest Reliability for Startle Blink Latency*

	Same picture group	Different picture group
Raw latency by picture content		
Negative	.55**	.66**
Neutral	.55**	.80**
Positive	.49*	.63**
Latency difference score contrasts		
Negative − neutral	.07	.44*
Positive − neutral	−.06	.11
Mean(negative, positive) − neutral	−.09	.35
Negative − positive	.29	.15

*$p < .05$, **$p < .01$.

stability across the two assessments, with slightly, but not significantly higher test–retest correlations for the Different group (see Table 3).

Correlations across the two assessments were then performed for the four emotion-modulation contrasts. None of the four contrasts showed stability across the 4-week interval with the exception of the negative − neutral contrast for the Different group ($r = .44$, $p < .03$; see Table 3).

Relations Between Emotion-Modulated Startle and Self-Report

There were no significant correlations between the PANAS-Trait positive and negative affect scores or any of the BIS/BAS scales and any of the four emotion-modulation blink magnitude or latency contrasts, either when examined separately for each assessment or when averaged across the two assessments (ps $> .10$).

Discussion

These data indicate that when using the same stimuli to manipulate affective state, the emotion modulation of the startle response is not reliable over time. Despite the fact that the mean group data indicate emotion-modulation effects were present for both groups at both assessments when aggregating across probe times, correlations between the two assessments indicated that subjects viewing two different sets of pictures showed greater stability of emotion modulation of blink reflex magnitude than those viewing the same pictures on both occasions. Furthermore, this greater stability for the group that saw different pictures appears to be due to higher stability for the arousal component of the pictures and not the valence of the pictures. The effect size for the stability of emotion-modulated startle when viewing different pictures is comparable to what has been reported for other biological indices related to affective style (e.g., the correlation for electroencephalographic measures of anterior activation asymmetry was found to range from .53 to .72 [Tomarken et al., 1992]).

Interestingly, the raw blink magnitudes showed very high test–retest stability. It may well be that raw magnitude is heavily influenced by nonpsychological factors such as muscle size and fatty deposition that obviously change little over a 1-month period. These findings with raw blink magnitude are congruent with previous data (Jennings et al., 1994). Similarly, raw blink latency showed good stability for each picture category separately, but generally poor test–retest reliability for the change scores.

Data from our laboratory (Sutton et al., 1997) and Bradley, Lang, and Cuthbert (1993) have shown that emotion-modulation effects are still present when picture stimuli are repeated in the same session. Similarly, these data show that there is a significant main effect for picture valence for both groups at both assessments, indicating that at the second assessment the group that saw the same pictures still exhibited emotion modulation of startle. Thus, our hypothesis that the repeated viewing of pictures on different occasions may lead to decreased emotional impact was not confirmed. The group that saw the same pictures did show emotion modulation of startle responding at the second assessment. However, the fact that this group exhibited poor test–retest reliability indicates that on the average, although subjects showed an ordering of startle magnitude going from negative to neutral to positive, the relative differences among these conditions was not preserved over time. Perhaps when viewed multiple times the familiarity of the stimulus becomes an important factor in the modulation of the blink response. Our data suggest that the group that viewed different pictures showed elevated arousal effects at the second assessment compared with the group that was exposed to the same pictures. This pattern implies that arousal may be the key component that leads to increased test–retest reliability for different pictures. In contrast, when viewed a second time the pictures may elicit lower levels of arousal, which could contribute to the decreased reliability for repeated picture exposure.

In addition to blink magnitude, startle reflex latency has also been examined as a measure of emotional responsiveness. However, the literature on emotion modulation of blink latency has yielded numerous inconsistencies. In the current study, blink onset latency was significantly shorter during negative than during neutral or positive pictures. Although some researchers using affective pictures have also found results consistent with these findings (Bradley, Cuthbert, & Lang, 1991), others have not (Vrana et al., 1988). Additionally, Cook et al. (1992) found facilitation of blink latency to negative pictures only among high fear subjects. Cook has suggested that startle latency is modulated specifically by arousal and not valence (Cook et al., 1991). In the present study, however, positive pictures did not lead to facilitation of blink latency, which is consistent with the data of Bradley et al. (1990). In addition, the data presented here do not suggest that more arousing pictures lead to greater stability over time. In considering the latency data, however, it must be noted that our resolution for measuring latency was coarse at best and thus our findings with this measure should be regarded with caution.

The data from this study indicate clearly that the stability of the emotion-modulated startle response requires further investigation. Future studies are needed to address a number of limitations of the current study. One limitation concerns the small number of male participants, and the disproportionately small number of male participants in the group that viewed different pictures. Because of this limitation, these data may not be generalizable across gender. A second limitation concerns the probe times that were used. One of the probe times occurred following stimulus offset. It is conceivable that better test–retest stability will be achieved with probe times that are closer together in time and that all occur during rather than after the picture presentation.

The findings from this study suggest that when considering the use of emotion-modulated startle for assessing trait-like characteristics, it is important to recognize the moderate degree of temporal stability, at least as the paradigm is commonly used. These data highlight the importance of using different picture stimuli across assessments. However, we must also recognize that whereas this paradigm may not show robust reliability, its validity may not be compromised. As demonstrated by Cook et al. (1991, 1992), potentiation of startle is associated with trait measures of fearfulness and lack of potentiation has been found in a population hypothesized to be low in fearfulness, psychopaths (Patrick et al., 1993). Furthermore, emotion modulation of startle responding has been found to be significantly correlated among monozygotic but not dizygotic twins, suggesting that affective modulation may be partly heritable (Carlson, Katsanis, Iacono, & McGue, 1997). Thus, the emotion-modulated startle paradigm may be a valid indicator of individual differences and trait measures of personality constructs, such as fearfulness. However, the low reliability of the paradigm as commonly implemented with affective pictures presented for 6 s (Bradley, Lang, and Cuthbert, 1993; Cook et al., 1992) may cap the predictive validity of the measure. Perhaps by modifying parameters of stimulus presentation, such as presenting pictures for a briefer period, or by utilizing more potent stimuli such as idiographically tailored pictures or imagery the reliability of emotion-modulated startle responding can be improved.

REFERENCES

Balaban, M. T. (1995). Affective influences on startle in five-month-old infants: Reactions to facial expressions of emotion. *Child Development*, *66*, 28–36.

Bradley, M. M., Cuthbert, B. N., & Lang, P. J. (1990). Startle reflex modification: Emotion or attention? *Psychophysiology*, *27*, 513–522.

Bradley, M. M., Cuthbert, B. N., & Lang, P. J. (1991). Startle and emotion: Lateral acoustic probes and the bilateral blink. *Psychophysiology*, *28*, 285–295.

Bradley, M. M., Cuthbert, B. N., & Lang, P. J. (1993). Pictures as prepulse: Attention and emotion in startle modification. *Psychophysiology*, *30*, 541–545.

Bradley, M. M., Lang, P. J., & Cuthbert, B. N. (1993). Emotion, novelty, and the startle reflex: Habituation in humans. *Behavioral Neuroscience*, *107*, 970–980.

Carlson, S. R., Katsanis, J., Iacono, W. G., & McGue, M. (1997). Emotional modulation of the startle reflex in twins: Preliminary findings. *Biological Psychology*, *46*, 235–246.

Carver, C. L., & White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS scales. *Journal of Personality and Social Psychology*, *67*, 319–333.

Center for the Study of Emotion and Attention. (1994). *The international affective picture system* [photographic slides]. Gainesville, FL: Center for Research in Psychophysiology, University of Florida.

Cook, E. W., Davis, T. L., Hawk, L. W., Spence, E. L., & Gautier, C. H. (1992). Fearfulness and startle potentiation during aversive stimuli. *Psychophysiology*, *29*, 633–645.

Cook, E. W., III, Hawk, L. W., Davis, T. L., & Stevenson, V. E. (1991). Affective individual differences and startle reflex modulation. *Journal of Abnormal Psychology*, *100*, 3–13.

Gray, J. A. (1994). Three fundamental emotion systems. In P. Ekman & R. J. Davidson (Eds.), *The nature of emotion: Fundamental questions* (pp. 243–247). New York: Oxford University Press.

Huynh, H., & Feldt, L. S. (1970). Conditions under which mean square ratios in repeated measurements designs have exact *F*-distributions. *Journal of the American Statistical Association*, *65*, 1582–1589.

Jennings, P. D., Dawson, M. E., Schell, A. M., Earleywine, M., & Runyan, M. D. (1994). Test–retest reliability of startle blink modification [Abstract]. *Psychophysiology*, *S31*, 1994.

Kimura, J. (1992). The blink reflex as a clinical test. In M. J. Aminoff (Ed.), *Electrodiagnosis in clinical neurology* (3rd ed., pp. 369–402). London: Churchill-Livingstone.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1990). Emotion, attention, and the startle reflex. *Psychological Review*, *97*, 377–395.

Lang, P. J., Greenwald, M. K., & Bradley, M. (1993). *The international affective picture system: Standardization procedure and results for affective judgments—Technical Report 1D.* Gainesville, FL: University of Florida.

Lang, P. J., Greenwald, M. K., Bradley, M. M., & Hamm, A. O. (1993). Looking at pictures: Affective, facial, visceral, and behavioral reactions. *Psychophysiology*, *30*, 261–273.

Larson, C. L., Davidson, R. J., Sutton, S. K., Ruffalo, D., & Nietert, J. Y. (1999). *The time course and stability of the emotion modulated startle response: Relations with anterior brain asymmetry.* Unpublished manuscript.

Miltner, W., Matjak, M., Braun, C., Diekmann, H. D., & Brody, S. (1994). Emotional qualities of odors and their influence on the startle reflex in humans. *Psychophysiology*, *31*, 107–110.

Patrick, C. J., Bradley, M. M., & Lang, P. J. (1993). Emotion in the criminal psychopath: startle reflex modulation. *Journal of Abnormal Psychology*, *102*, 82–92.

Schwarzkopf, S. B., McCoy, L., Smith, D. A., & Boutros, N. N. (1993). Test–retest reliability of prepulse inhibition of the acoustic startle response. *Biological Psychiatry*, *34*, 896–900.

Sutton, S. K., Davidson, R. J., Donzella, B., Irwin, W., & Dottl, D. A. (1997). Manipulating affective state using extended picture presentations. *Psychophysiology*, *34*, 217–226.

Tomarken, A. J., Davidson, R. J., Wheeler, R. W., & Kinney, L. (1992). Psychometric properties of resting anterior EEG asymmetry: Temporal stability and internal consistency. *Psychophysiology*, *29*, 576–592.

Vrana, S. R., Spence, E. L., & Lang, P. J. (1988). The startle probe response: A new measure of emotion? *Journal of Abnormal Psychology*, *97*, 487–491.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, *54*, 1063–1070.

(Received August 4, 1998; Accepted April 17, 1999)

APPENDICES

Appendix A: Picture Set A

Pleasant Pictures Description	Number	Gender	Unpleasant Pictures Description	Number	Neutral Pictures Description	Number
leopard by water	1650		boy licking cow	2730	male, capped	2190
couple	4608		mafia hit	3010	man with hat	2570
couple	4680		burn victim	3053	person's shadow	2880
astronaut	5460		throat slash	3071	layered mushrooms	5510
space shuttle	5450		burn victim	3100	mushrooms in clover	5530
windsurfers	5623		burn victim	3110	three dry mushrooms	5534
hiker along ridge	5629		mutilated body	3130	green plant and soil	5740
fireworks	5910		baby w/ eye tumor	3170	rolling pin	7000

continued

Appendix A: Picture Set A (*Continued*)

Pleasant Pictures Description	Number	Gender	Unpleasant Pictures Description	Number	Neutral Pictures Description	Number
chocolate drink	7270		subway robbery	3500	spoon	7004
ski jump (90 M)	8030		aimed gun	6230	blue mug	7009
extended sailor	8080		aimed pistol	6260	fan	7020
sailing	8170		knife assault at door	6313	claw hammer	7034
powder skiing	8190		man slapping wife	6360	hair dryer	7050
catamaran	8210		man w/ ski mask	6510	old book	7090
shooting rapids	8370		knife assault	6550	painted fabric	7160
rafting duo	8400		suicide	6570	square block	7185
roller coaster	8490		plane evacuation	9050	fabric and beads	7207
US currency	8501		torture & beating	9252	decorated plate	7233
male w/ groceries	4532	female	decaying animal	9570	window & blue wall	7490
couple on street	4599	female	male neo-Nazi	9800	unmade bed w/ white	7710
couple, gazing	4609	female	auto wreck	9910	block boy pondering	9070
couple, sex	4607	male				
couple, nude	4664	male				
motorcycle racing	8260	male				

Appendix B: Picture Set B

Pleasant Pictures Description	Number	Gender	Unpleasant Pictures Description	Number	Neutral Pictures Description	Number
three puppies	1710		mutilated face	3000	neutral male, young	2200
couple, kissing	4660		face laceration	3030	itnerant Latino boy	2870
couple, nude	4690		mangled face	3060	two people	2890
astronaut & earth	5470		missing face	3080	three mushrooms	5520
skydivers in a circle	5621		burn victim, left chest	3102	mushroom	5531
hang-glider	5626		mutilated body	3120	blue door	5731
rocky mountain peak	5700		mutilated hand	3150	electric outlet	6150
turkey dinner	7230		severed hand	3400	blue towel	7002
Disney castle	7502		man w/ gun in mouth	3530	white bowl	7006
slalom ski racer	8034		aimed pistol	6250	wicker basket	7010
gymnast	8090		man grabbing woman	6312	wooden stool	7025
cliff divers	8180		knife assault	6350	glass mug	7035
water skier	8200		man w/ ski mask	6370	fork, silverware	7080
pilot & prop plane	8300		knife assault	6540	yellow fire hydrant	7100
victorious relay team	8380		man holding gun	6560	brass and wood lamp	7175
victorious gymnast	8470		gang attacking car	6821	scarves, etc., pastel	7205
gold bars	8500		medics carrying victim	9250	clothes rack	7217
stack of money	8502		man w/ dead child	9410	wooden chair	7235
male, volleyball	4533	female	sinking ship (nosedive)	9600	lab building	7491
fireman/nice chest	4572	female	Klansmen & cross	9810	box of Kleenex	7950
couple, kissing	4640	female	firefighters	9921	environmental workers	9700
couple	4652	male				
smiling female	2030	male				
pilot towing hang-glider	8340	male				