Psychophysiology, 42 (2005), 604–610. Blackwell Publishing Inc. Printed in the USA.

DOI: 10.1111/j.1469-8986.2005.00345.x

Stability of emotion-modulated startle during short and long picture presentation

CHRISTINE L. LARSON,[a] DANTE RUFFALO,[b] JENNIFER Y. NIETERT,[b] AND RICHARD J. DAVIDSON[b,c]

[a]Department of Psychology, Michigan State University, East Lansing, Michigan, USA
[b]Department of Psychology, University of Wisconsin–Madison, Madison, Wisconsin, USA
[c]Department of Psychiatry, University of Wisconsin–Madison, Madison, Wisconsin, USA

Abstract

Following reports on improved test–retest reliability of emotion-modulated startle during a 6-s picture presentation when different pictures are presented at each assessment (Larson et al., 2000) and data suggesting that brief picture presentations also elicit affective blink modulation (Codispoti, Bradley, & Lang, 2001), we assessed test–retest reliability of blink modulation during brief picture presentations. At two acoustic startle sessions (4 weeks apart) subjects viewed different IAPS pictures for either 6 s (long group) or 300 ms (short), with emotion modulation assessed at three different points in time during and following picture viewing. Group ANOVAs revealed emotion modulation for both short and long groups. In addition, comparable and, in some cases, greater stability of emotion modulation was found for short compared to long picture presentations. Stability was generally low for individual probe times for both groups.

Descriptors: Startle reflex, Emotion-modulated startle paradigm, Emotion, Test–retest stability

The initial finding by Vrana, Spence, and Lang (1988) that the magnitude of the eyeblink startle response is potentiated in the presence of aversive foreground stimuli and attenuated in response to pleasant stimuli has resulted in a large number of studies using this technique to assess state and trait patterns of emotional responding. With the proliferation of emotion research using the emotion-modulated startle paradigm some standard practices have emerged in terms of study design parameters. One of the most notable is the use of the International Affective Picture System (Center for the Study of Emotion and Attention, 1999; Lang, Bradley, & Cuthbert, 1999) as a foreground affective stimulus. In addition, a substantial portion of studies using these pictures in conjunction with affect-modulated startle paradigms have adopted a 6-s picture presentation time (e.g., Bradley, Cuthbert, & Lang, 1993; Bradley, Lang, & Cuthbert, 1993; Cook, Davis, Hawk, Spence, & Gautier, 1992; Lang, Greenwald, Bradley, & Hamm, 1993). Furthermore, researchers have begun to use this paradigm to identify links between affective modulation of startle and traitlike individual differences in mood and affective reactivity (Cook et al., 1992; Cook, Hawk, Davis, & Stevenson, 1991; Hawk & Kowmas, 2003; Jackson et al., 2003; Patrick, Bradley, & Lang, 1993; Vanman, Mejia, Dawson, Schell, & Raine, 2003). Such data suggest that affect modulation of startle blink magnitude may be a useful index of trait affective style. To merit consideration as a trait measure of affective responding, stability of emotion modulation of the blink reflex over time is desirable.

However, the extant literature on stability of affect modulation using the 6-s presentation indicates substantial inconsistency in stability estimates. Previous work by our laboratory has found that this paradigm elicits emotion modulation effects that are only moderately stable over time (Larson, Ruffalo, Neitert, & Davidson, 2000). In that study, presentation of novel pictures at the second assessment dramatically improved test–retest reliability (*r*s between .44 and .62) compared with a group for which the same pictures were presented at the two assessments (*r*s: .04–.25). Manber, Allen, Burton, and Kaszniak (2000) found low stability of affect modulation of startle across a 2-week interval (intraclass correlations range from –.30 to .25), even when different picture stimuli were presented at the two assessments. In contrast to the low to moderate stability coefficients in these two studies, other work has reported strong stability of affect modulation of startle across a 1-week interval (Bradley, Gianaros, & Lang, 1995). Another report examining responses to affective film clips found no mean differences in magnitude of blink modulation across a 1-month interval; however, test–retest coefficients were not computed (Kaviani, Gray, Checkley, Kumari, &

Support for this research was provided by NIMH grants MH40747 and 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). The authors wish to thank Adrian Pederson and Darren Dottl for their assistance in implementing the study.

Address reprint requests to: Christine L. Larson, Michigan State University, Psychology Department, 105B Psychology Building, East Lansing, MI, 48824–1116, USA. E-mail: larson37@msu.edu.

Wilson, 1999). Finally, although not assessing stability over time, Hawk and Cook (2000) found poor internal consistency for raw affect modulation scores using the picture paradigm. Thus, the status of stability over time of affect modulation of startle remains somewhat unclear at this time, with the few studies that presented test–retest correlations indicating low to moderate stability estimates.

Although stability of affect modulation of startle may be low to moderate, the stability over time of raw blink magnitudes is excellent (Larson et al., 2000), suggesting that any deficit in reliability is not inherent to the startle reflex itself, but is a function of the affect modulation of the reflex. This is consistent with the suggestion of Bradley, Cuthbert, and Lang (1999) that altering the properties of the foreground stimulus may impact affective modulation of startle. In particular, the duration of picture presentation may have important implications for the extent to which affective processing of the stimulus is completed prior to picture offset. In recent years a number of studies have examined affective modulation of startle blink in response to very brief picture presentations; however, the stability over time of emotion-modulated startle blink using a brief picture presentation has not yet been established. Thus, the present study set out to establish the test–retest stability across repeated assessments of emotion-modulated startle under brief picture presentation, and to compare this stability with that of 6-s picture presentation.

Although the standard 6-s picture presentation has proved useful as an affect induction in many circumstances, there are some questions for which brief presentations may be more advantageous. For example, as a number of researchers have suggested (LeDoux, 1996; Öhman & Mineka, 2001) rapid discrimination of salient environmental cues is advantageous in an evolutionary context. Indeed, affective discrimination occurs early in picture processing (Cuthbert, Schupp, Bradley, Birbaumer, & Lang, 2000; Junghöfer, Bradley, Elbert, & Lang, 2001). Brief picture presentation may offer a more appropriate test of this rapid detection as well as for assessing the immediate affective impact of a stimulus. Longer presentations allow more opportunity for elaborative processing and online emotion regulation and the large individual differences inherent in such processes may obscure the initial affective response.

Recently, a number of studies have found that brief picture presentation is effective in eliciting emotion modulation of startle. Using 150-ms presentations, Globisch, Hamm, Esteves, and Öhman (1999) found potentiation of startle for fear-relevant stimuli among subjects fearful of snakes and spiders. Codispoti, Bradley, and Lang (2001) found that in general emotion modulation of the startle blink was similar for a 500-ms picture presentation as compared to previous findings using 6-s presentations. Based on their observation of potentiation as early as 300 ms poststimulus onset during 500-ms presentation, they suggested that affective modulation of startle may be accelerated during brief relative to longer picture presentations. As an alternative hypothesis, they suggested that longer picture presentations could lead to extended attentional engagement and therefore postpone the onset of affect modulation. One possibility suggested by these data is that the rapid onset of emotion modulation of the blink response associated with brief presentations reflects a more immediate affective response that has been less influenced by the further stimulus processing allowed for during a longer presentation. As such, brief picture presentations associated with a more immediate, pure affective response may exhibit greater reliability in the affective modulation of the startle blink response.

In addition to an exploration of the relative stability of short compared to long picture presentations over time, we also sought to assess the stability of emotion-modulated startle at different time points during and after picture presentation. A number of investigators have begun to examine the time course of affective modulation across picture viewing (Dichter, Tomarken, & Baucom, 2002; Robinson & Vrana, 2000; Vanman, Boehmelt, Dawson, & Schell, 1996) and the extent to which individual differences in time course are related to trait affect (Dichter, Tomarken, Shelton, & Sutton, 2004; Jackson et al., 2003; Nitschke et al., 2002). As is the case for brief picture processing, the relative stability of affect modulation of startle at different points in time during picture viewing has not yet been assessed.

The present study offers an extension and replication of the findings on stability over time of emotion-modulated startle and the utility of brief picture presentations. Two groups of participants participated in two sessions, separated by 4 weeks. Different pictures were presented at each session. The short group viewed pictures presented for 300 ms and the long group for 6 s. Based on prior work we predicted that both short and long picture presentations would result in affective modulation of startle blinks. In addition, we predicted that short picture presentations would yield higher test–retest stability of affective modulation. As a secondary aim we also sought to assess differential stability of affect modulation of startle across different points in time during and just following picture presentation.

Method

Participants

Participants were 78 (58 female, 20 male) undergraduates (all between the ages of 18 and 20) recruited from Introductory Psychology classes at the University of Wisconsin–Madison. Participants were randomly assigned to one of two groups, a group that viewed brief picture presentations (short group, $n = 38$) and a group that viewed 6-s presentations (long group, $n = 40$). Due to participant dropout at the second session and an insufficient number of startle responses at one or both assessments (see below) 25 participants were dropped, yielding a final sample of 53 participants (27 short, 21 female, 6 male; 26 long, 20 female, 6 male).

Materials

Based on published self-report ratings of valence and arousal (Lang et al., 1999) 42 positive, negative, and neutral pictures were selected from Shows 1–12 of the International Affective Picture System (Center for the Study of Emotion and Attention, 1999; see Larson et al., 2000, for details of picture selection).[1] All

---

[1] The IAPS pictures used were: 1650, 1710, 2030[a], 2190, 2200, 2570, 2730, 2870, 2880, 2890, 3000, 3010, 3030, 3053, 3060, 3071, 3080, 3100, 3102, 3110, 3120, 3130, 3150, 3170, 3400, 3500, 3530, 4532[b], 4533[b], 4572[b], 4599[b], 4607[a], 4608, 4609[b], 4640[b], 4652[a], 4660, 4664[a], 4680, 4690, 5450, 5460, 5470, 5510, 5520, 5530, 5531, 5534, 5621, 5623, 5626, 5629, 5700, 5731, 5740, 5910, 6150, 6230, 6250, 6260, 6312, 6313, 6350, 6360, 6370, 6510, 6540, 6550, 6560, 6570, 6821, 7000, 7002, 7004, 7006, 7009, 7010, 7020, 7025, 7034, 7035, 7050, 7080, 7090, 7100, 7160, 7175, 7185, 7205, 7207, 7217, 7230, 7233, 7235, 7270, 7490, 7491, 7502, 7710, 7950, 8030, 8034, 8080, 8090, 8170, 8180, 8190, 8200, 8210, 8260[a], 8300, 8340[a], 8370, 8380, 8400, 8470, 8490, 8500, 8501, 8502, 9050, 9070, 9250, 9252, 9410, 9570, 9600, 9700, 9800, 9810, 9910, 9921. Due to gender differences in the normed ratings for pleasantly valenced pictures, particularly erotic stimuli, the pleasant picture set used included six pictures presented just to men (a) and six presented just to women (b).

participants saw different pictures at the two assessments. Two different sets of 63 pictures (21 of each valence) were matched on arousal ratings and presented in a counterbalanced order across participants.

Procedure

Two sessions separated by 4 weeks were conducted, each at the same time of day and following the identical procedure (detailed in Larson et al., 2000). Pictures were presented in a quasirandom order for either 300 ms (short group) or 6 s (long group) with a randomized 10- to 18-s intertrial interval (mean ITI = 14 s). Acoustic startle probes (50 ms, 95 dB white noise burst) were presented at three different probe times: 1.5, 4.5, and 7.5 s following picture onset. The three probe times used here were chosen to maintain consistency with prior work; the 7.5-s probes occurred after picture offset for the long group and all probes occurred following picture offset in the short group. Participants were not instructed to maintain an image of the picture after its offset. For each of the three probe times, probes occurred during six trials of each valence.

Startle Recording and Quantification

Raw and integrated EMG from the orbicularis oculi were collected using two Sensormedics mini-electrodes placed directly below the left eye. 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. 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 using SnapStream software (HEM Data Corporation, Springfield, MI) and a 12-bit analog-to-digital board (Analogic Corporation, Wakefield, MA).

Approximately 11.7% of eyeblink reflexes were excluded from analyses due to excessive noise during a 50-ms, prestartle baseline period (e.g., blinks, unusually high amounts of integrated EMG) or because the onset of the integrated eyeblink reflex began less than 20 ms following the probe. Eyeblink reflex magnitudes (in microvolts) were calculated by subtracting the amount of integrated EMG at reflex onset from that at 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. Eyeblink reflex magnitudes were z-transformed within a given session (separately for each assessment for each participant). Data were z-transformed within, not across, assessment in order minimize the influence of mean differences in raw blink magnitude at the two sessions on affect modulation scores for the standardized data. Blinks greater than 3 standard deviations above the mean for a given participant were excluded.

At the first assessment, 13 participants did not display a probe-induced eyeblink or more than half of the trials were bad (see criteria above) yielding an n of 65 for the first assessment. Of the 53 participants that returned for the second assessment, 4 participants did not respond to the startle probes or were dropped because more than half of the trials were bad, leaving a maximum n of 49 for the second assessment. 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 participants with at least three good responses in every cell (Probe Time × Picture Content × Assessment). This yielded a final sample of 43 participants (21 short, 15 female, 6 male; 22 long, 16 female, 6 male).[2] Mean blink magnitude for each cell was computed for each participant. ANOVA and test–retest correlation analyses were conducted on these means for individual probe time data as well as for aggregated data (mean across all three probe times).

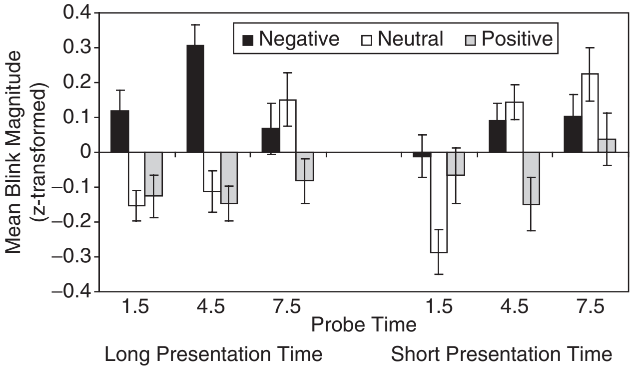

Figure 1. Standardized (z-score) blink magnitude averaged across assessment for each probe time, picture content, and group.

Results

Emotion Modulation Effects

Standardized data. To determine whether affect modulation of the startle blink was present a four-way Huynh–Feldt (1970) corrected ANOVA with Group as the between-subjects factor (short, long), and Assessment (1, 2), Probe Time (1.5, 4.5, and 7.5 s), and Picture Content (negative, neutral, positive) as within-subjects factors was calculated. A significant main effect was found for Picture Content, $F(2,82) = 10.20$, $p = .0001$, $\varepsilon = 1.03$ (see Figure 1). A priori comparisons revealed that across Group, Assessment, and Probe Time blinks were larger for negative compared to neutral, $t(42) = 2.52$, $p < .02$, and positive pictures, $t(42) = 4.14$, $p = .0002$. Attenuation of blink magnitude to positive relative to neutral pictures was not significant, $t(42) = 1.80$, $p < .08$. A main effect was also found for Probe Time, $F(2,82) = 8.51$, $p = .0004$, $\varepsilon = 1.00$. Post hoc tests for this main effect and all interactions were corrected for multiple comparisons using a Bonferroni adjustment. Across Group, Assessment, and Picture Content, blink responses were significantly smaller at the 1.5-s probe time compared with the 7.5-s probe time, $t(42) = 3.78$, $p = .01$.

These main effects were qualified by significant Picture Content × Probe Time, $F(4,164) = 7.73$, $p = .0001$, $\varepsilon = 1.01$, and Picture Content × Probe Time × Group, $F(4,164) = 2.78$, $p < .03$, $\varepsilon = 1.04$, interactions. Across both groups, potentiation of blink magnitude for negative compared to neutral pictures was present at 1.5 s, $t(42) = 4.59$, $p < .001$. Blink magnitude in response to negative pictures was greater than that for positive pictures for both groups at the 4.5-s probe time, $t(42) = 5.45$, $p < .0005$. Among the group viewing the pictures for 6 s, blink magnitude was significantly potentiated in response to negative compared to neutral pictures at 1.5 s, $t(21) = 4.08$, $p = .005$, and 4.5 s, $t(21) = 4.32$, $p = .001$. When participants viewed pleasant

---

[2] The power to detect a .5 correlation (moderate stability) at a p value of .05 two-tailed is .63 for the long group ($N = 22$) and .62 for the short group ($N = 21$).

compared to neutral pictures, blink magnitudes for the short group were attenuated at 4.5 s, $t(20) = 3.40$, $p < .05$. The long group showed significantly greater potentiation to negative compared to neutral pictures than did the short group at the 4.5 s probe time, $t(41) = 3.85$, $p = .01$.

The main effect for Group did not reach significance, $F(1,41) = 3.10$, $p = .09$. There were also no significant effects involving the Assessment factor (ps > .15).

Raw data. An identical ANOVA computed for the raw data revealed a similar pattern of results with two exceptions. The Picture Content × Probe Time × Group interaction did not reach significance ($p = .11$). In addition, in contrast to the z-transformed data, there was a significant main effect for Assessment, $F(1,41) = 4.89$, $p < .03$. Across all probe times, picture contents and both groups, raw blink magnitude was smaller at Time 2 (mean = 84.66 μV, $SD = 63.21$) than at Time 1 (mean = 97.17 μV, $SD = 64.14$), $t(42) = 2.05$, $p < .05$ (uncorrected).

Stability of Emotion Modulation of the Startle Response

Stability of emotion modulation of startle was calculated by correlating potentiation (negative – neutral) and attenuation (positive – neutral) difference scores at the first assessment with those from the second assessment. Previous research has demonstrated that these contrasts are of use in comparing each valence with neutral, but that they confound valence and arousal (see Lang et al., 1993; Vrana et al., 1988). As a result, two subsequent contrasts were computed to compare high versus low arousal (mean blink magnitudes for negative and positive minus magnitude for neutral) and effects of valence independent of arousal (negative – positive). Test–retest correlations were then performed for all four contrasts for standardized and raw data for both groups at each probe time as well.[3] Of the 48 correlations performed for the raw and standardized data (four contrasts, three probe times, two groups) only four achieved statistical significance (uncorrected for multiple tests), and one of these correlations was negative. The mean test–retest stability correlations across probe time and contrast were as follows: short group standardized data: mean = .07, $SD = .25$, long group standardized data, mean = .07, $SD = .25$, short group raw data, mean = .08, $SD = .23$, and long group raw data, mean = .04, $SD = .32$. Thus, test–retest stability for the individual probe time data was generally quite low.

Standardized data aggregated across probe time. To enable comparison with previous work (Larson et al., 2000) and to assess the impact of including more trials on the stability coefficients, correlations between the two assessments were also computed for the data aggregated across probe times. For negative – neutral valences, both groups showed moderate stability (long: $r = .50$, $p < .02$, 90% CI = .26–.84; short: $r = .46$, $p < .04$, 90% CI = .19–.80; see Figure 2). There were also no group differences in stability for the negative versus positive contrast stability (long: $r = .37$, $p < .09$, 90% CI = .10–.68; short: $r = .52$, $p < .02$, 90% CI = .27–.88). In contrast, for positive – neutral contrast the short group ($r = .47$, $p = .03$, 90% CI = .2–.81) showed significantly greater test–retest stability than the long group ($r = -.06$, p = n.s., 90% CI = –.35–.23; $z = 1.73$, $p < .05$). A similar but nonsignificant group difference was found for the arousal contrast (mean(negative, positive) – neutral), such that the short group ($r = .44$, $p < .05$, 90% CI = .17–.78) showed moderate stability, but the long group ($r = .22$, $p = .33$, 90% CI = –.07–.52) did not.

Raw data aggregated across probe time. Test–retest stability correlations for the aggregated raw data were much less robust than for the standardized data. Neither group showed significant test–retest stability for the negative versus neutral (long: $r = .14$, p = n.s., 90% CI = –.15–.43; short: $r = .35$, $p < .12$, 90% CI = .06–.67), positive versus neutral (long: $r = .38$, $p = .08$, 90% CI = .11–.69; short: $r = .17$, p = n.s., 90% CI = –.13–.47), or arousal contrasts (long: $r = .13$, p = n.s., 90% CI = –.16–.42; short: $r = .26$, p = n.s., 90% CI = –.04–.57). For the negative versus positive contrast moderate stability was found for the short ($r = .54$, $p < .01$, 90% CI = .30–.90), but not the long group ($r = .20$, p = n.s., 90% CI = –.09–.50).

How many assessments are needed to achieve adequate stability? The Spearman–Brown formula was applied to estimate the number of assessments, assuming the identical paradigm, that would be required to achieve adequate test–retest stability correlations, defined here as .07 (Nunnally & Bernstein, 1994). Qualitatively, using the more stable aggregated across probe time z-transformed data, the number of assessments needed to achieve this degree of stability was generally lower for the short (positive – neutral = 5.0, negative – neutral = 4.7, negative – positive = 3.7, arousal contrast = 5.1) than long group (positive – neutral = 70.7, negative – neutral = 4.0, negative – positive = 6.8, arousal contrast = 14.2).

Discussion

These data confirm reports indicating that brief pictures, in this case 300-ms presentations, do in fact elicit emotion modulation of startle and further support the claim of Bradley et al. (1999) that the presence of the picture is not necessary for the modulation of startle. As all of the probes presented during the 300-ms presentation followed stimulus offset, and the first two of these probe times elicited emotion modulation of startle, the data from the present study are congruent with this assertion. Furthermore, subjects viewing 6-s pictures exhibited attenuation in response to pleasant pictures at the 7.5-s probe time. However, no emotion-modulation effects were observed among the short group at this late probe time, suggesting that for the 300-ms picture presentation emotion modulation effects disappear at some point between 4 and 7 s following picture offset. This is consistent with Bradley and colleagues' (1999) proposal that adjusting the duration of the foreground stimulus may impact the time course of affect modulation. In this case, although brief pictures do result in affect modulation following picture offset, these effects do not persist for as long following picture onset as is the case for the 6-s pictures. In addition, as would be expected across both groups, the raw data showed a significant decrease in blink magnitude at the second assessment. These data are in keeping with a large body of work indicating that the startle blink reflex habituates over time (Abel, Waikar, Pedro, Hemsley, & Geyer, 1998; Bradley, Lang, et al., 1993). However, as has been found previously (Bradley, Lang, et al., 1993; Bradley et al., 1995), this main effect for assessment was not qualified by any interactions;

---

[3] 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.

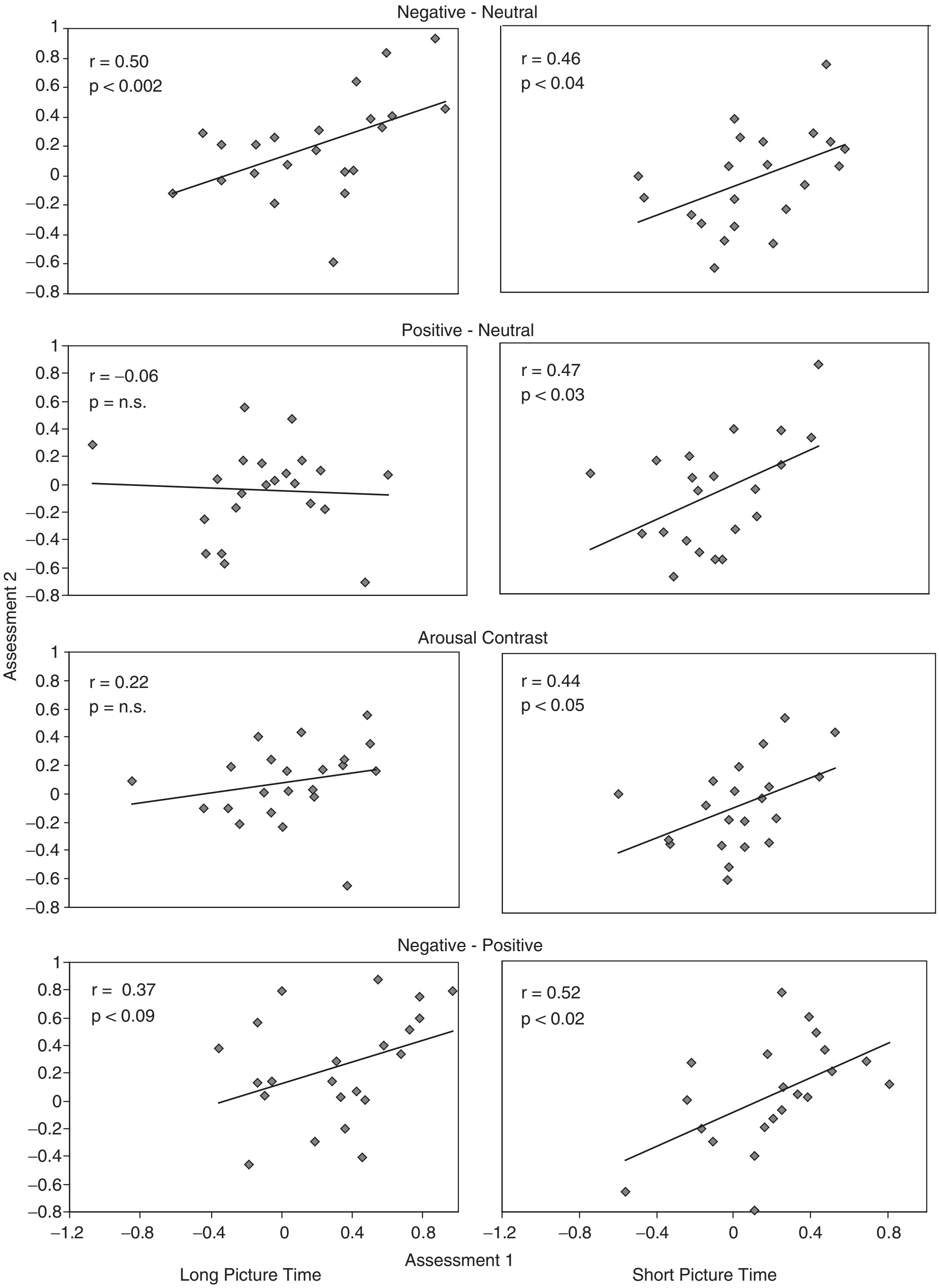

Figure 2. Scatter plots demonstrating the test–retest correlation for the standardized data aggregated across all three probe times for each of the four contrasts (negative – neutral, positive – neutral, arousal: mean (negative, positive) – neutral, and valence: negative – positive) separately for the long ($N = 22$) and short ($N = 21$) groups.

thus the modulation patterns present at the first session were maintained 4 weeks later for both short and long picture presentations.

With respect to stability over time, consistent with the Larson et al. (2000) data, across probe latencies we found moderate test–retest stability of emotion-modulated startle. As predicted, stability of emotion-modulated blink responses during brief presentations was at least equivalent and in one case superior to longer presentations. This slight advantage in stability for brief compared to long pictures may reflect a minimization of more complex picture processing tasks, such as scanning and more extended encoding, that occurs during longer picture presentations. Along with the now well-replicated finding that brief pictures effectively modulate startle blink, these data on the stability of emotion modulation provide additional support for the use of brief picture presentations. Notably, however, the stability estimates from the current study also indicate differences between the z-transformed and raw data. In virtually all cases, stability estimates were higher for z-transformed as compared to raw data.

The stability estimates from the aggregated standardized data presented here are largely but not completely consistent with our earlier work (Larson et al., 2000). Although all subjects were shown different pictures at the two assessments, within the long group the negative versus neutral but not positive versus neutral contrast showed significant stability across assessments. Although the cause of this partial failure to replicate is unclear, work by others has indicated that significant attenuation of blink responses occurs in response to pictures of erotica but not necessarily for other pictures rated as being pleasant, including adventure, sports, food, families, and nature (Bradley, Codispoti, Cuthbert, & Lang, 2001). The inconsistency in test–retest stability of responses to positive pictures may be a reflection of the fact that pleasant pictures, with the exception of erotica, are less robust and consistent modulators of the startle blink response than are negative pictures.

In addition, stability of emotion modulation at individual probe times was, as might be expected given the smaller number of trials represented, more variable and generally lower than that of the data aggregated across probe time. No one particular time point seemed to yield better stability over time for both pleasant and unpleasant picture stimuli, regardless of picture duration. Previous work with the 6-s picture paradigm has suggested that the mid-picture epoch (Bradley, Cuthbert, et al., 1993; Vanman et al., 1996) tends to result in maximum emotion modulation of the blink reflex. However, in the current study, among the group viewing the 6-s pictures there was no apparent advantage in terms of stability for this probe time. In fact, despite the fact that aggregating across the probe times used in the present study results in the inclusion of very different phases of picture processing, this procedure still results in more robust stability estimates.

Two psychometric issues are relevant to the present findings. First, the use of standardized scores in calculating test–retest stability has been questioned on the grounds that it artificially equates the variation in scores from different assessments and thus can be misleading (Willett, Singer, & Martin, 1998). Thus, both standardized and raw scores have been reported here. As these data show, stability over time is low to moderate for both measures, but is typically somewhat lower for raw compared to standardized scores. In addition, difference scores, such as those used here, tend to be less reliable, particularly when the two original scores are strongly correlated (Nunnally & Bernstein, 1994). Although the correlations between each condition were not presented here, the correlations between conditions for the standardized scores were typically low ($r = .3$ or lower), whereas correlations among positive, negative, and neutral pictures for the raw scores were quite high ($r = .6$ or greater). However, as our data indicate, both measures led to generally low to moderate stability over time, with only very slight improvement in stability estimates for the standardized scores, despite the substantially lower magnitude of correlations among conditions. Furthermore, low internal consistency of affect modulation has also been found with nondifference score metrics, such as percent modulation (Hawk & Cook, 2000). Although the use of difference scores here may have compromised test–retest stability coefficients, the fact that in the present study stability was not high even for the standardized scores, taken together with the growing corpus of data finding low reliability for emotion-modulation scores (Hawk & Cook, 2000; Larson et al., 2000; Manber et al., 2000), suggests that the reliability of affect modulation of startle may be inherently limited.

In the current data, although the aggregated standardized data indicate equivalent and, in some cases, enhanced stability of emotion-modulated startle for brief presentations, it should be noted that the test–retest stability coefficients for this condition were still only in the moderate range. Thus, although stability over time of affective blink modulation may in some cases be improved through the use of brief picture presentations, clearly this does not yield an affect measure with high reliability across assessments. The extent to which the ceiling of reliability for affect modulation of the startle blink has been reached for this basic paradigm is unclear at this time. Taken together these data along with the Larson et al. (2000) data indicate that adjustments to the foreground stimulus, including the use of novel stimuli and brief presentations, can improve stability over time of emotion-modulated startle to some degree. Additional research manipulating other aspects of the experimental design in search of improved reliability is clearly required to further address this possibility. However, it may be the case that affective modulation of startle simply is not a robust indicator of trait affective reactivity. Although high test–retest reliability is desirable for stable traits, it is less so for transient states. The apparent limits on the stability of emotion modulation of startle may indeed reflect the degree to which transient state affect, rather than trait patterns of affective style, govern emotion modulation of startle blink. Given that there is some degree of stability over time, a likely possibility is that emotion modulation of startle reflects a mixture of both state and trait components of affective reactivity.

REFERENCES

Abel, K., Waikar, M., Pedro, B., Hemsley, D., & Geyer, M. (1998). Repeated testing of prepulse inhibition and habituation of the startle reflex: A study in healthy human controls. *Journal of Psychopharmacology*, *12*, 330–337.

Bradley, M. M., Codispoti, M., Cuthbert, B. N., & Lang, P. J. (2001). Emotion and motivation I: Defensive and appetitive reactions in picture processing. *Emotion*, *1*, 276–298.

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., Cuthbert, B. N., & Lang, P. J. (1999). Affect and the startle reflex. In M. E. Dawson, A. M. Schell, & A. H. Böhmelt (Eds.) *Startle modification: Implications for neuroscience, cognitive science, and clinical science* (pp. 157–183). New York: Cambridge University Press.

Bradley, M. M., Gianaros, P., & Lang, P. J. (1995). As time goes by: Stability of affective startle modulation. *Psychophysiology*, *32*, S21.

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

Center for the Study of Emotion and Attention (CSEA-NIMH) (1999). *The international affective picture system: Digitized photographs*. Gainesville, FL: The Center for Research in Psychophysiology, University of Florida.

Codispoti, M., Bradley, M. M., & Lang, P. J. (2001). Affective reactions to briefly presented pictures. *Psychophysiology*, *38*, 474–478.

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.

Cuthbert, B. N., Schupp, H. T., Bradley, M. M., Birbaumer, N., & Lang, P. J. (2000). Brain potentials in affective picture processing: Covariation with autonomic arousal and affective report. *Biological Psychology*, *52*, 95–111.

Dichter, G. S., Tomarken, A. J., & Baucom, B. R. (2002). Startle modulation before, during, and after exposure to emotional stimuli. *International Journal of Psychophysiology*, *43*, 191–196.

Dichter, G. S., Tomarken, A. J., Shelton, R. C., & Sutton, S. K. (2004). Early- and late-onset startle modulation in unipolar depression. *Psychophysiology*, *41*, 433–440.

Globisch, J., Hamm, A. O., Esteves, F., & Öhman, A. (1999). Fear appears fast: Temporal course of startle reflex potentiation in animal fearful subjects. *Psychophysiology*, *36*, 66–75.

Hawk, L. W., & Cook, E. W., III. (2000). Independence of valence modulation and prepulse inhibition of startle. *Psychophysiology*, *37*, 5–12.

Hawk, L. W., & Kowmas, A. D. (2003). Affective modulation and prepulse inhibition of startle among undergraduates high and low in behavioral inhibition and approach. *Psychophysiology*, *40*, 131–138.

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.

Jackson, D. C., Mueller, C. J., Dolski, I., Dalton, K. M., Nitschke, J. B., & Urry, H. L., et al. (2003). Now you feel it, now you don't: Frontal brain electrical asymmetry and individual differences in emotion regulation. *Psychological Science*, *14*, 612–617.

Junghöfer, M., Bradley, M. M., Elbert, T. R., & Lang, P. J. (2001). Fleeting images: A new look at early emotion discrimination. *Psychophysiology*, *38*, 175–178.

Kaviani, H., Gray, J. A., Checkley, S. A., Kumari, V., & Wilson, G. D. (1999). Modulation of the acoustic startle reflex by emotionally-toned film clips. *International Journal of Psychophysiology*, *32*, 47–54.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1999). *International affective picture system: Technical manual and affective ratings*. Gainesville, FL: The Center for Research in Psychophysiology, 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., Ruffalo, D., Nietert, J., & Davidson, R. J. (2000). Temporal stability of the emotion-modulated startle response. *Psychophysiology*, *37*, 92–101.

LeDoux, J. E. (1996). *The emotional brain*. New York: Simon & Schuster.

Manber, R., Allen, J. J. B., Burton, K., & Kaszniak, A. W. (2000). Valence-dependent modulation of psychophysiological measures: Is there consistency across repeated testing? *Psychophysiology*, *37*, 683–692.

Nitschke, J. B., Larson, C. L., Smoller, M. J., Navin, S. D., Pederson, A. J. C., & Ruffalo, D., et al. (2002). Startle potentiation in aversive anticipation: Evidence for state but not trait effects. *Psychophysiology*, *39*, 254–258.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw Hill.

Öhman, A. E., & Mineka, S. (2001). Fears, phobias, and preparedness: Toward an evolved module of fear and fear learning. *Psychological Review*, *108*, 483–522.

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.

Robinson, J. D., & Vrana, S. R. (2000). The time course of emotional and attentional modulation of the startle eyeblink reflex during imagery. *International Journal of Psychophysiology*, *37*, 275–289.

Vanman, E. J., Boehmelt, A. H., Dawson, M. E., & Schell, A. M. (1996). The varying time courses of the attentional and affective modulation of the startle eyeblink reflex. *Psychophysiology*, *33*, 691–697.

Vanman, E. J., Mejia, V. Y., Dawson, M. E., Schell, A. M., & Raine, A. (2003). Modification of the startle reflex in a community sample: Do one or two dimensions of psychopathy underlie emotional processing? *Personality and Individual Differences*, *35*, 2007–2021.

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.

Willett, J. B., Singer, J. D., & Martin, N. C. (1998). The design and analysis of longitudinal studies of development and psychopathology in context: Statistical models and methodological recommendations. *Development and Psychopathology*, *10*, 395–426.

(Received November 3, 2003; Accepted May 26, 2005)