

Running Head: EARLY ATTENTION TO EXPECTANCY VIOLATION

Event-related Potentials Reveal Early Attention Bias for Negative, Unexpected Behavior

Rita Jerónimo^{1*}, Hannah I. Volpert² and Bruce D. Bartholow²

1. Instituto Universitário de Lisboa (ISCTE-IUL), CIS-IUL, Lisboa, Portugal, Avenida das Forças Armadas, 1649-026 Lisboa – Portugal

2. Department of Psychological Sciences, 210 McAlester Hall, University of Missouri, Columbia, MO 65211. Email: bartholowb@missouri.edu, or hannah.volpert@gmail.com

To appear in *Social Neuroscience* (Note that this manuscript is not the version of record for this publication, which will be made available from the publisher's website.)

Author Note:

*Address correspondence to: Rita Jerónimo, ISCTE-IUL, Departamento de Psicologia Social e das Organizações, Avenida das Forças Armadas, 1649-026 Lisboa – Portugal. E-mail: rita.jeronimo@iscte.pt. Telephone (+351)217903000, or (+351)914169198.

Abstract

Numerous studies have documented that expectancy-violating (EV) behavior (i.e., behavior that violates existing person impressions) elicits more effortful cognitive processing compared to expectancy-consistent (EC) behavior. Some studies also have shown that this effect is modulated by the valence of behavior, though this finding is inconsistent with some extant models of expectancy processes. The current research investigated whether the valence of EV information affects very rapid attentional processes thought to tag goal-relevant information for more elaborative processing at later stages. Event-related brain potentials (ERPs) were recorded as participants read depictions of behavior that either were consistent with or violated established impressions about fictitious characters. Consistent with predictions, a very early attention-related ERP component, the frontal P2, differentiated negative from positive EV behavior but was unaffected by the valence of EC behavior. This effect occurred much earlier in processing than has been demonstrated in prior reports of EV effects on neural response, suggesting that impression-formation goals tune attention to information that might signal the need to modify existing impressions.

Keywords: Expectancy violation; Positive-negative asymmetry; Event-related potential; Attention; Valence

Most of human life is spent interacting with, thinking about, and trying to understand other people. A crucial byproduct of this ongoing social cognition is the formation of expectancies, derived from acquired knowledge of what other people are like, which we use to interpret their ongoing behaviors (see Olson, Roese, & Zanna, 1996). An interesting consequence of this practice is that expectancy-violating (EV) information about people elicits more extensive cognitive processing than does expectancy-consistent (EC) information (see Bargh & Thein, 1985), as perceivers engage effortful processes aimed at reconciling the discrepancy between existing templates and new information (see Macrae, Bodenhausen, Schloerscheidt, & Milne, 1999). Consistent with this view, psychophysiological studies have shown that, relative to EC behaviors, EV behaviors elicit enhanced neural activation in relatively long-latency event-related potential (ERP) components associated with elaboration and updating of information in working memory (e.g., Bartholow, Fabiani, Gratton, & Bettencourt, 2001; Van Duynslaeger, Van Overwalle, & Verstraeten, 2007).

In addition to whether behaviors are consistent with expectancies, behavior valence also strongly affects its processing and its influence on evaluations, with negative behaviors more strongly influencing perceivers' judgments than positive behaviors (e.g., Peeters & Czapinski, 1990). This *positive-negative asymmetry* often is explained in terms of the diagnosticity of negative versus positive moral behaviors for trait categorizations (the *cue diagnosticity model*; see Peeters & Czapinski, 1990).

Whereas studies in these domains generally have focused on processing operations that take place later in the processing sequence, this theorized difference in the certainty with which initial impressions are held should have implications for earlier operations that influence rapid engagement of attention to EV information. This was the focus of the current study. If an initial

impression of a target implies a negative trait, then the perceiver should be less motivated to monitor that target's subsequent EV (i.e., positive) actions because the initial impression will be held with relative certainty. In contrast, when an initial impression is positive (and, thus, held with less certainty), the perceiver should be motivated to attend to the target's subsequent actions. If the target subsequently behaves negatively, that new information should be particularly salient (and goal-relevant) to the perceiver, and given the tight coupling of motivation and attention (see Lang, 1995) should capture more attention. This prediction stands in contrast to some previous proposals, which held that positive and negative violations should equally influence processing (e.g., Wigboldus, Dijksterhuis, & van Knippenberg, 2003).

Some previous research supports the basic idea of an asymmetry in the influence of EV information as a function of its valence. For example, Reeder and Covert (1986) found that positive impressions underwent greater change following EV behaviors than did negative impressions (also see Ybarra, 2002). Although such findings indicate that (especially negative) EV behavior elicits enhanced working memory and impression updating, they do not address whether the valence of EV information influences early attention processes that might trigger such elaboration. The current research aimed to address this issue by measuring the amplitude of the P2 component of the ERP elicited by positive and negative, EV and EC behaviors. Initially linked to greater allocation of early attention and sensory/perceptual resources (e.g., Luck & Hillyard, 1994), enhanced P2 amplitude also has been associated with attention to unexpected or improbable stimuli (Peters, Suchan, Zhang, & Daum, 2005) and stimuli relevant to a perceiver's current goals (Amodio, 2010).

Given the hypothesis that negative EV behaviors are especially relevant to perceivers' goal of forming accurate impressions (Reeder & Covert, 1986), it should be the case that

negative EV behavior more strongly engages early attention processes, as measured by P2 amplitude, compared to positive EV behavior and to EC behavior. The current experiment was designed to test this hypothesis.

Method

Participants

Sixteen right-handed, native English-speaking, healthy university students (7 men; ages 18-32), all with normal or corrected-to-normal vision, participated in exchange for credit toward a course requirement or \$18.

Materials and Procedure

The methods used in this study were described in detail elsewhere (see Bartholow et al., 2001) and so are only briefly reviewed here.¹ Participants were asked to read 20 brief paragraphs, each describing a different target person in terms that conveyed a strong trait inference (10 positive, 10 negative), and were told to form impressions of these individuals. Following each target description, individual target behaviors were described via sentences, all six words in length, presented one word at a time in the center of a computer monitor. Words were presented at a rate of 1 every 350 ms and were displayed for 300 ms. Twelve sentences (trials) were presented for each of the 20 targets, of which two described EC behavior and two described EV behavior; the remainder depicted expectancy-irrelevant behaviors. ERPs were recorded to the final, critical word of each sentence, which conveyed the behavior's valence and congruency with the initial impression. With this design, negative behaviors were EV for positive targets and positive behaviors were EV for negative targets, permitting examination of the implications of a behavior's valence according to whether it was EC or EV.

¹ Data focusing on a different aspect of the ERP from these participants were previously reported in Bartholow et al. (2001). However, findings reported here were not included in that previous report.

Electrophysiological Recording and Analysis

The electroencephalogram (EEG) was recorded from 19 scalp locations (10-20 system) using tin electrodes fixed in an electrode cap (Electro-Cap International, Eaton, OH). All scalp electrodes were referenced online to the right mastoid; an average mastoid reference was derived offline. Ocular artifacts (blinks) were corrected offline using a standard procedure. The EEG was sampled at 100 Hz and filtered online at 0.01 to 30 Hz; a pre-stimulus baseline period was defined as the 100 ms prior to the presentation of the final word in each sentence. ERP waveforms were averaged separately as a function of the congruency and valence of sentence-ending words.

Data from one male participant were unusable due to a high proportion of EEG artifacts. Thus, the final sample used for analyses included 15 individuals. Visual inspection of the waveforms (see Figure 1) indicated that the P2 component occurred approximately 200-330 ms following the onset of sentence-ending words. Thus, the P2 was quantified as the mean amplitude from 200-330 ms following final word onset at each electrode. Data were analyzed using mixed hierarchical linear models (HLM), which have several advantages over univariate repeated-measures ANOVA for analyzing psychophysiological data (see Gratton, 2007), particularly when sample size is modest. Here, the data were modeled as 60 observations (every trial type at 15 electrodes) within 15 individuals, including random intercepts of subject and of electrodes within subjects.

Results

Mean amplitudes recorded at each of the frontal (F7, F3, Fz, F4, and F8), central (T3, C3, Cz, C4, and T4), and parietal (T5, P3, Pz, P4, and T6) electrodes were submitted to an initial HLM that included factors for Coronal location (frontal, central, parietal), Expectancy (EC, EV)

and Behavior valence (positive, negative). This analysis produced a main effect for Coronal location, $F(2, 666) = 50.40, p < .001, R^2 = .07$, indicating that the P2 was larger over frontal locations ($M = 2.67 \mu V$) than over central ($M = 2.09 \mu V$) or parietal locations ($M = 0.40 \mu V$), as well as a Coronal x Expectancy x Valence interaction, $F(2, 666) = 15.45, p < .001, R^2 = .02$. A follow-up contrast showed that the magnitude of the predicted Expectancy x Valence interaction was larger at frontal electrode locations than at either central or parietal locations, $t(666) = 4.92, p < .001, R^2 = .035$. An additional 2 (Expectancy) x 2 (Valence) HLM restricted to frontal locations produced a significant interaction, $F(1, 278) = 17.25, p < .001, R^2 = .058$ (means given Table 1). Follow-up comparisons showed that whereas positive EC and EV behaviors elicited similar P2 amplitudes, $t(222) = -0.19, p = .85, R^2 = .00016$, negative EV behaviors (i.e., negative acts committed by positive targets) elicited much larger P2 amplitudes than did negative EC behaviors, $t(222) = 5.68, p < .0001, R^2 = .127$.

Discussion

Previous studies investigating neural responses to EC and EV behaviors (e.g., Bartholow et al., 2001; Van Duynslaeger et al., 2007) have focused on differences in effortful processes often associated with updating memory representations (e.g., P300 amplitude), but to date have not examined more rapidly deployed attention processes. The current research was grounded in the premise that expectancy and valence should interact in determining the engagement of early attention to goal-relevant information, an idea informed by the cue diagnosticity model (Peeters & Czapinski, 1990). This model predicts that perceivers motivated to form accurate impressions should be especially attuned to negative, EV behaviors because of their potential to alter an uncertain positive impression.

The current findings are the first to demonstrate that neurocognitive responses associated with goal-directed attention distinguish target behavior as a function of its valence and trait consistency at such an early processing stage. These findings also generally align with previous research indicating that evaluations of valence occur almost instantaneously and require little (if any) cognitive elaboration (see Zajonc, 1980). Moreover, the pattern observed here in the P2 is generally in-line with previous work showing that written depictions of negative (but not positive) EV behavior elicit enhanced activation of the corrugator supercilii muscle (Bartholow et al., 2001), associated with spontaneous expression of negative affect (see Heller, Lapate, Mayer, & Davidson, 2014), and with research indicating that the emotional quality of words affects P2 amplitude when participants attend to their meaning rather than orthographic features (Begleiter, Projesz, & 1979). Considered together with these previous findings, the current data suggest the interaction of valence and expectancy can influence person perception much more rapidly than previously assumed.

A recent model of impression formation posits that because expectancies make inconsistent traits less available in memory, trait encoding of EV behaviors will be obstructed, making their implications more difficult to understand (Jerónimo, Garcia-Marques, Ferreira, & Macrae, 2015). However, this model establishes no clear role for the valence of behaviors in affecting trait encoding difficulty. The current results could suggest a combination of the trait inhibition and cue-diagnostics explanations in accounting for the interaction of expectancy and valence in capturing early attention to EV behaviors. According to this explanation, in the context of an impression-formation goal, very rapid valence evaluations of a behavior tune early attention to “tag” goal-relevant information (see Amodio, 2010) that might be difficult to

interpret (see Jerónimo et al., 2015) and therefore will require further, more elaborated processing downstream.

Considerably more research will be required to systematically investigate this idea and its implications for other aspects of person perception, particularly given the small sample size used here, which represents an important limitation of this work (though concerns over statistical power are mitigated, to some degree, by the use of HLM for data analysis).² It also will be important for future research to establish whether negative EV information selectively engages early attention when participants do not have an explicit impression-formation goal. Some previous research focusing on later processing stages (Van Duynslaeger et al., 2007) suggests similar neural responses to EV information regardless of whether participants are instructed to form an impression. In theory, however, having this explicit goal enhances the motivational relevance of trait-diagnostic (i.e., negative) EV information, which should bias early attention toward such information. It remains to be determined whether a similar early-stage processing bias will emerge in the absence of such a goal.

² The size of the current sample was informed by and is consistent with samples used in similar studies published prior to the design of this study (e.g., Cacioppo, Crites, Berntson, & Coles, 1993; Crites, Cacioppo, Gardner, & Berntson, 1995).

References

1. Amodio, D. M. (2010). Coordinated roles of motivation and perception in the regulation of intergroup responses: Frontal cortical asymmetry effects on the P2 event-related potential and behaviour. *Journal of Cognitive Neuroscience*, 22, 2609–2617.
2. Bargh, J. A., & Thein, R. D. (1985). Individual construct accessibility, person memory, and the recall-judgment link: The case of information overload. *Journal of Personality and Social Psychology*, 49, 1129-1146.
3. Bartholow, B. D., Fabiani, M., Gratton, G., & Bettencourt, B. A. (2001). A psychophysiological examination of cognitive processing of and affective responses to social expectancy violations. *Psychological Science*, 12, 197-204.
4. Begleiter, H., Projesz, B., & Garazzo, R. (1979). Visual evoked potentials and affective ratings of semantic stimuli. In H. Begleiter (Ed.), *Evoked brain potentials and behavior* (pp. 127–143). New York: Plenum Press.
5. Cacioppo, J. T., Crites, S. L. Jr., Berntson, G. G., & Coles, M. G. H. (1993). If attitudes affect how stimuli are processed, should they not affect the event-related brain potential? *Psychological Science*, 4, 108-112.
6. Crites, S. L. Jr., Cacioppo, J. T., Gardner, W. L., & Berntson, G. G. (1995). Bioelectrical echoes from evaluative categorization: II. A late positive brain potential that varies as a function of attitude registration rather than attitude report. *Journal of Personality and Social Psychology*, 68, 997-1013.
7. Gratton, G. (2007). Biosignal processing. In J. Cacioppo, L. Tassinary, & G. Berntson (Eds.), *Handbook of Psychophysiology* (pp. 900-923). New York, NY: Cambridge University Press.

8. Heller, A.S., Lapate, R.C., Mayer, K.E., & Davidson, R.J. (2014). The face of negative affect: trial-by-trial corrugator responses to negative pictures are positively associated with amygdala and negatively associated with ventromedial prefrontal cortex activity. *Journal of Cognitive Neuroscience*, 26, 2102-2110.
9. Jerónimo, R., Garcia-Marques, L., Ferreira, M., & Macrae, N. C. (2015). When expectancies harm comprehension: Encoding flexibility in impression formation. *Journal of Experimental Social Psychology*, 61, 110-119.
10. Lang (1995). The emotion probe: Studies of motivation and attention. *American Psychologist*, 50, 372-385.
11. Luck, S. J., & Hillyard, S. A. (1994). Electrophysiological correlates of feature analysis during visual search. *Psychophysiology*, 31, 291-308.
12. Macrae, C. N., Bodenhausen, G. V., Schloerscheidt, A. M., & Milne, A. B. (1999). Tales of the unexpected: Executive function and person perception. *Journal of Personality and Social Psychology*, 76, 200-213.
13. Olson, J. M., Roese, N. J., & Zanna, M. P. (1996). Expectancies. In E.T. Higgins & A.W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 211-38). New York: Guilford.
14. Peeters, G. & Czapinski, J. (1990). Positive-negative asymmetry in evaluations: The distinction between affective and informational negativity effects. *European Review of Social Psychology*, 1, 33-60.
15. Peters, J., Suchan, B., Zhang, Y., & Daum, I. (2005). Visuo-verbal interactions in working memory: Evidence from event-related potentials. *Cognitive Brain Research*, 25, 406 – 415.

16. Reeder, G. D., & Coover, M. D. (1986). Revising an impression of morality. *Social Cognition, 4*, 1–17.
17. Van Duynslaeger, M., Van Overwalle, F., & Verstraeten, E. (2007). Electrophysiological time course and brain areas of spontaneous and intentional trait inferences. *Social Cognitive and Affective Neuroscience, 2*, 174-188.
18. Wigboldus, D. H. J., Dijksterhuis, A., & van Knippenberg, A. (2003). When stereotypes get in the way: Stereotypes obstruct stereotype-inconsistent trait inferences. *Journal of Personality and Social Psychology, 84*, 470-484.
19. Ybarra, O. (2002). Naïve causal understanding of valenced behaviors and its implications for social information processing. *Psychological Bulletin, 128*, 421–441.
20. Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American psychologist, 35*, 151-175.

Table 1

Mean P2 Amplitude Values (and SDs) at Frontal Electrodes as a Function of Condition

Target behaviors	Electrodes					Mean
	F7	F3	Fz	F4	F8	
Positive						
EC	2.19	2.31	2.32	1.85	1.36	2.01
	(3.59)	(4.25)	(5.09)	(4.57)	(3.66)	(4.23)
EV	1.56	1.97	2.22	1.07	1.85	1.74
	(3.08)	(3.71)	(3.77)	(3.98)	(3.28)	(3.56)
Negative						
EC	1.62	2.31	2.61	2.48	1.89	2.18
	(3.25)	(3.23)	(3.74)	(3.57)	(3.01)	(3.36)
EV	3.62	3.92	3.99	3.48	3.29	3.66
	(3.89)	(4.63)	(4.49)	(4.15)	(3.70)	(4.17)

Note. EC = expectancy-consistent; EV = expectancy-violating. Values in parentheses are standard deviations. Amplitude values represent the average voltage measured 200-330 ms following onset of the final (target) word in each sentence. Electrode position in the table from left to right mirrors electrode position on the scalp.

Figure Caption

Figure 1. ERP waveforms measured at frontal electrodes as a function of target behavior valence and consistency with previous expectancies. Zero on the x-axis (in ms) indicates the onset of the final word in each sentence, which conveyed both the valence (positive or negative) and consistency (expectancy-consistent [EC] or expectancy-violating [EV]) of the behavior. The P2 is the prominent positivity in the waveform peaking around 280 ms post-stimulus.

Figure 1

