Emotion
2015, Vol. 15, No. 1, 35–44

1528-3542/15/$12.00 http://dx.doi.org/10.1037/a0038474

Are Valence and Arousal Separable in Emotional Experience?

Assaf Kron
University of Haifa and University of Toronto

Maryna Pilkiw and Jasmin Banaei
University of Toronto

Ariel Goldstein
Hebrew University

Adam Keith Anderson
University of Toronto and Cornell University

The bipolar valence-arousal model of conscious experience of emotions is prominent in emotion research. In this work, we examine the validity of this model in the context of feelings elicited by visual stimuli. In particular, we examine whether arousal has a unique contribution over bivariate valence (separate measures for pleasure and displeasure) in explaining physiological arousal (electrodermal activity, EDA) and self-reported feelings at the level of item-specific responses across and within individuals. Our results suggest that self-reports of arousal have neither an advantage in predicting EDA nor make a unique contribution when valence is present in the model. Acceptance of the null hypothesis was confirmed with the use of the Bayesian information criterion. Arousal also showed no advantage over valence in predicting global feelings, but demonstrated a small unique component (1.5% to 4% of variance explained). These results have practical implications for both experimental design in the study of emotions and the underlying bases of their conscious experience.

Keywords: valence, arousal, EDA, models of emotions, self reports

The bipolar valence-arousal model is a prevalent model of conscious experience of emotions in the scientific literature. Feelings in this model are estimated by using two questions. The first question asks about the degree of feelings of arousal (e.g., feelings of stimulation, excitation, arousal). The second question asks about the degree of pleasure versus displeasure. The bipolar valence-arousal model posits that when participants introspect and report on different aspects of their emotional experience, arousal, and valence are two distinct types of feeling qualia. Much as how one can separately attend to color and motion as distinct dimensions of perceptual experience that rely on distinct appraisals of external events, when attention is turned inward, emotional experience can be decomposed into qualitatively distinct perceptions of valence and arousal. In the present work, we examined the validity of the dissociation of arousal and valence in the context of internal feelings elicited by external visual stimuli.

In much of emotion research, the rule or a criterion that is used to assign stimuli to the experimental conditions is based on emotional experience. To support this claim, we note that the three largest emotion stimuli pools of pictures (IAPS: the International Affective Picture System; Lang, Bradley, & Cuthbert, 1997), audio recording (IADS: The International Affective Digitized Sounds; Bradley & Lang, 2007), and English words (ANEW: Affective Norms for English Words; Bradley & Lang, 1999) are mapped on self-reports of feelings. Each stimulus in the IAPS, IADS, and ANEW (henceforth referred as IIA) has a value that is based on self-reports of emotional experience averaged across many participants. Therefore, each time researchers use stimuli from the IIA and assign them to the experimental conditions according to their associated norms, they essentially use self-reports of feelings as criteria to assign stimuli to the experimental conditions. When using stimuli and norms from the IIA (or any other stimuli categorized by self-reports of feelings), the implicit model of emotions is a model of conscious experience of emotions. Therefore, at least when the norms are in use, models of conscious experience of emotions influence the experimental design, data, results, conclusions, and terms by which knowledge about emotions is accumulated, regardless of whether one examines brain activity, peripheral physiological activity, or any other measure that may not incorporate self-reports of feelings. As long as one uses self-reports of feelings to assign stimuli to the experimental conditions, parsing the emotional space is done in terms of self-reports of conscious experience of emotions. Whether intended or not, this is why models of conscious experience of emotions are fundamental to all emotion research.

Variants of the Bipolar Valence Arousal Model

The bipolar valence-arousal model is the prevalent model of conscious experience of emotions in the scientific literature. The popularity of the bipolar valence-arousal model is reflected in how the three largest emotion stimulus pools of pictures, audio recordings, and English words (e.g., the IIA) are mapped on self-reports

Assaf Kron, Department of Psychology, University of Haifa and University of Toronto; Maryna Pilkiw and Jasmin Banaei, Department of Psychology, University of Toronto; Ariel Goldstein, Department of Psychology, Hebrew University; Adam Keith Anderson, Department of Psychology, University of Toronto and Department of Human Development, Cornell University.

Correspondence concerning this article should be addressed to Assaf A. Kron, Department of Psychology, University of Haifa, 199 Aba Khoushy Avenue, Mount Carmel 31905, Israel. E-mail: akron@psy.haifa.ac.il

of bipolar-valence and arousal. Each stimulus in the IIA has a value that is based on self-reports of emotional experience averaged across many participants. The IIA are not only technical devices, but rather implicit distributors of the bipolar valence-arousal model. Researchers that use stimuli from the pools frequently use their adjacent norms (values of feelings of valence and arousal) as criteria to assign stimuli to the experimental conditions. In other words, using the norms of the IAPS implies parsing emotion space in terms of bipolar valence and arousal and implicitly or explicitly assuming this model. Due to this, data, conclusions, and aggregation of knowledge are made, to a large extent, in terms of bipolar valence and arousal. In this work we examined the validity of the bipolar valence-arousal model in the context of the IIA. In particular, we examined if arousal and valence were distinct components of self-reports of emotional feelings.

There are two versions of the bipolar valence-arousal model: the "two-intensities" account, which holds that bipolar valence and arousal, each has its own intensity (e.g., Russell, 1980; Russell & Feldman Barrett, 1999), and the "arousal as intensity" version on which arousal is conceived as intensity of bipolar valence (e.g., Bradley, Codispoti, Cuthbert, & Lang, 2001). The theoretical model that underlies the IIA is the former. Feelings in the IIA are estimated by two questions. The first question asks about the degree of feelings of arousal (e.g., feeling of stimulation, excitation, arousal). The second question asks about the degree of pleasure versus displeasure. Hence, the bipolar valence-arousal model of the IIA has three assumptions. The first is that pleasure and displeasure are mutually exclusive. One cannot co-occur with the other. The second is that arousal and valence are two different types of feelings; the instructions for the ratings scales use different wording to make participants introspect and report on different aspects of their emotional experience. The instructions use the terms *arousal, excitation* to refer to a type of feeling associated with arousal and it uses the words *pleasure*/*displeasure* to refer to a different type of feeling associated with valence. The third assumption is that arousal and pleasure/displeasure have dissociable intensities. This assumption is evidenced by the application of two separate continuous scales, which index the degree or intensity of a feeling, one for arousal and one for valence. These assumptions are not trivial; they presuppose that while experiencing emotional feelings in the presence of stimuli, participants can distinguish and report about two qualitatively distinct types of feelings and their intensities, one for valence and the other for arousal.

However, contrary to the theoretical model that underlies the IIA, the actual distribution of pictures across the arousal and bipolar valence axes do not appear to obey the two intensities principle. Stimuli in the IIA show a consistent *V* shape, relationship between bipolar valence and arousal (see Figure 1a). It is difficult to reconcile the *V* shape relationship with the two intensities account. A *V* shape function is expected when absolute values are plotted against their original values (Figure 1b). It may be then that arousal is not separable from valence in the IIA, but it is its absolute value—its intensity. This would be more consistent with variation of the bipolar-valence model where arousal is conceived as intensity of bipolar valence (Bradley et al., 2001).

Despite the *V* shape function implying full dependence, two types of evidence do support the distinction between arousal and valence in the context of the IIA. The first comes from the limited association between self-reports of arousal and valence; modeled by absolute values (*V* shape) or a quadratic function, bipolar valence explains only around 20% of arousal variance (computed from Ito, Cacioppo, & Lang, 1998; Lang, Bradley, & Cuthbert, 1999). The second piece of evidence against arousal-as-intensity-of-valence comes from the distinct associations of arousal and valence with physiological activity. For example, self-reports of arousal predict electrodermal activity (EDA) better than bipolar valence. However, bipolar valence is negatively associated with corrugator EMG activity, whereas self-reports of arousal are not (e.g., Lang, Greenwald, Bradley, & Hamm, 1993). Indeed, to the extent that self-reports of arousal and valence reflect different qualities of emotional experience, we would expect them to show divergent correlations with physiological measures.

One conclusion from the earlier discussion is that neither of the two versions of the bipolar valence-arousal model can be fully justified in the context of the IIA. The two intensities version provides no explanation to the *V* shape function that is consistently present in the IIA. The arousal-as-intensity-of-valence version cannot explain cases in which arousal predicts variables that the *V* shape function of bipolar valence cannot. In recent work, Kron, Goldstein, Lee, Gardhouse, and Anderson, 2013 attempted to reconcile the discrepancy between the theoretical model underlying the IIA and the actual distribution of pictures along the bipolar valence and arousal axes. They suggested that the feeling of

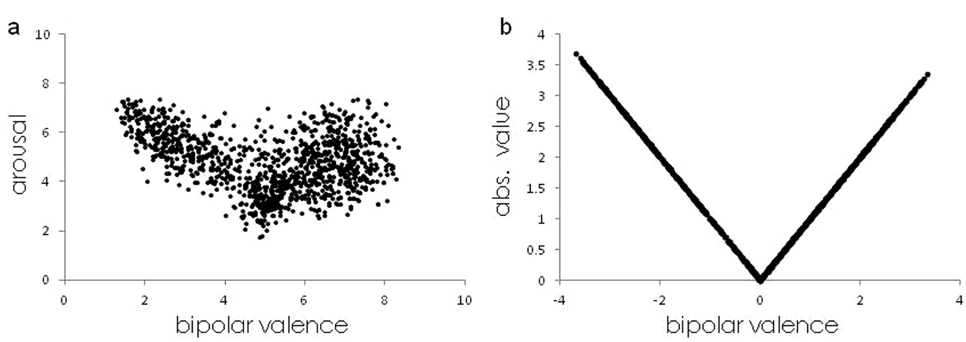

Figure 1. The graph in (a) depicts the distributions of pictures from the International Affective Picture System (Lang, Bradley, & Cuthbert, 1997) over the arousal and bipolar valence axes. The graph in (b) shows an absolute scores function.

arousal was (a) not the simple intensity of bipolar valence (b) but also not a distinct quale with an intensity distinct from the intensity of the feelings of pleasure/displeasure. Rather, both arousal and bipolar valence can be decomposed into unipolar dimensions of pleasure and displeasure and that when using two independent scales for pleasure and displeasure (unipolar valence), a strong association between valence and arousal, whether self-report or physiological, was evident.

Unipolar Valence Models

The unipolar valence perspective (also known as the bivariate valence model) maintains that valence is activated by two independent systems of pleasure and displeasure (Cacioppo & Bernston, 1994; Cacioppo, Gardner, & Berntson, 1997; Larsen, McGraw & Cacioppo, 2001) and, accordingly, that emotional experience should be estimated by separate scales for pleasure and displeasure (Larsen, Norris, McGraw, Hawkley, & Cacioppo, 2009). The unipolar valence model can be distinguished from the bipolar model in two ways: (1) it permits mixed emotions—the simultaneous activation of pleasure and displeasure, and (2) it does not necessarily assume a separate system for arousal (Kron et al., 2013). Initial evidence for the existence of mixed emotions (e.g., Larsen & McGraw, 2011; Larsen et al., 2001) and separate brain systems for pleasure and displeasure (Anderson et al., 2003; Viinikainen et al., 2010) show indirect support for the unipolar valence account of conscious experience of emotion.

Kron et al. (2013) examined the relationship between self-reports of arousal and EDA using unipolar scales of pleasant (PL) and unpleasant (UN) as well as bipolar valence and arousal scales. Participants viewed IAPS pictures and after each picture rated their feelings either according to the bipolar valence-arousal model or unipolar PL and UN. If the dissociation between valence and arousal reflects distinct aspects of emotional experience (dissociation is due to the instructions of the arousal and PL/UN scales referring participants to report different qualia), then reporting one's level of arousal should not be equivalent to reporting on the intensity of pleasant and unpleasant valence—arousal should be more than simply the sum of pleasant and unpleasant feelings (i.e., *pleasant* scores plus *unpleasant* scores). By contrast, if the sum of reported pleasant and unpleasant valence is equivalent to self-reported arousal and predicts its physiological correlates, then this result would question the distinction between arousal and valence as different emotional qualia. They found that arousal was a function of PL and UN, close to their addition (Arousal = PL + UN). At the same time they showed that bipolar valence was also a function of PL and UN, and was very close to their subtraction (Valence = PL − UN).

In particular, the PL + UN sum is highly correlated with the self-reports of arousal and that it can explain physiological arousal (EDA) as well as arousal scores. These results suggest that arousal and valence may not be separate qualitative dimensions of emotional experience. The traditional evidence for dissociation between valence and arousal was found because valence was measured using a bipolar valence scale.

The Present Study

A critical piece of evidence for the distinction between valence and arousal comes from the self-reports of arousal, which provide a better estimate of physiological arousal than valence (e.g., Lang et al., 1993). As discussed earlier, Kron et al. (2013) demonstrated that arousal and valence estimated physiological arousal to the same degree. Yet this test can provide only a partial test of dependence. They showed that the correlation between arousal and EDA was very similar to that of PL and UN. However, demonstrating that correlations are similar does not imply that PL and UN explained the same component of variance in EDA as arousal scores. One way to test whether arousal and PL/UN predict the same variance component of EDA is by estimating the unique variance of EDA explained by arousal when both PL and UN are present in the same model. This critical test was not performed on previous data because participants did not rate PL/UN and arousal for the same pictures. The aim of the present work was to estimate arousal and PL/UN on the same trial, which allowed us to perform this critical test of whether there is a unique variance explained by self-reported arousal.

In the present work we examined the unique contribution of self-reports of arousal over pleasure and displeasure from two perspectives: when items (pictures) are units of analysis and when participants are units of analysis. Analyzing the unique contribution of arousal at the level of items is important given that item norms are frequently used to allocate stimuli to the different conditions. This is the level that previously has been most influential in the accumulation of scientific knowledge in terms of arousal and valence. If the distinction between arousal and PL/UN is manifested at the individual level, but not evident at the item level, it would raise questions about the research that uses norms of arousal and valence as indicators of arousal and valence in planning design and data analysis—as it is frequently done. On the other hand, testing the unique contribution of arousal and valence within participants is important at the theoretical level because it directly answers the question of whether humans have distinct qualia of arousal and pleasure/displeasure in response to an external stimulus. Based on Kron et al. (2013) we predicted that, at both picture and subject levels of analysis, there would be a strong association between PL + UN and self-reports of arousal ($\sim r = .85$). In addition, given the association of PL, UN with arousal, we then also expected no substantial distinct contribution of self-reports of arousal above and beyond the effect of PL and UN in predicting electrodermal response to emotional stimuli.

Method

There were no data exclusions in this study. All manipulations that were conducted and all measures that were collected are reported in the Method section.

Participants

There were 30 University of Toronto students (19 women) who completed the experiment for monetary compensation or for course credit. Sample size was determined based on previous experience with similar design (Kron et al., 2013).

Stimuli

Images were selected from the IAPS (Lang et al., 1997), in such a way that all possible combinations of arousal and valence of the IAPS were represented. To ensure that the stimuli were randomly chosen and were equally distributed across the arousal-valence IAPS space, we used an in-house algorithm. The algorithm randomly selected a sample of 72 images[1] such that the resulting two-dimensional shape of the selected sample was the same as the original shape of the IAPS set, and all of the images were spread across this shape in a uniform manner.

Self-Report Scales of Emotional Experience

Standard self-reports of feelings elicited by pictorial stimuli reflected not only one's feelings, but also nonexperiential knowledge, such as beliefs about expected emotions and semantic knowledge (Levenson, 2003; Robinson & Clore, 2002). The tendency to rely on semantics might be stronger when participants were asked to report their feelings, but they experienced very little or no feelings (Robinson & Clore, 2002). This may result in biased data, especially when estimating valence. Participants may understand valence as both semantic knowledge about stimuli (stimuli are labeled as having negative, positive content) and feelings (feelings are unpleasant or pleasant). Based on these observations, we carefully structured the instructions to refer participants to their feelings. We emphasized two aspects: that (a) self-reports should reflect the participant's internal feelings and not the value of a picture, belief, or semantics; and (b) the rating of *no feelings* should be selected if feelings are not detected (see the Appendix). A Power Point presentation, with all slides and an audio recording of the instructions, is available on request from A. Kron.

The instructions for self-reports had the following four components:

1. Explaining to the participants the distinction between semantics and their own experience of feelings. In particular we emphasized two cases (a) confusing evaluation of feelings with evaluation of the content of the picture (e.g., you feel pleasure vs. the content of the picture is pleasant), and (b) confusing feelings with belief or expectation about what "one should feel" while looking at a picture.

2. Framing the task as a "feelings detection task," for example, "You can think of this scale as a volume knob that indicates the intensity of your feelings. The question is whether you detected any feelings, and if yes, what was their intensity?"

3. Reducing the accessibility bias (e.g., reports about semantics in the absence of strong feelings) by legitimizing cases of no feelings: for example, "If you did not detect any feelings inside you, press [0] and press [1] only if you are sure that you detect some feelings."

4. Specific instructions for the ratings scales are as follows: All scales were composed as volume graphs (see the Appendix) ranging from 0 (*low*) to 8 (*high*).

- Feelings scale: Instructed to rate the maximum value of any type of feelings (arousal, pleasure, displeasure, or any other feeling).
- Arousal scale: Instructed to rate feelings of being excited, frenzied, jittery, and/or wide awake.
- Pleasure scale: Instructed to rate feelings of pleasure, happiness, and/or any other pleasant feelings.
- Displeasure scale: Instructed to rate feelings of displeasure, sadness, and/or any other unpleasant feelings.

Design

Pictures were presented randomly in two blocks of 36 pictures, with 1-min breaks in between. During each trial, participants rated their feelings according to four scales: They first rated their overall feelings making sure they focus on feelings and not semantic knowledge (see details in the Appendix) and then rated pleasure, displeasure, and feeling of arousal. Due to this, the scale for overall feelings always appeared first whereas the order of pleasure, displeasure, and arousal scales was counterbalanced across participants. Trials began with a blank screen presented randomly for 10 to 21 s. Then, a picture was presented for 6 s followed by the rating scales. Each picture was presented full screen onto a 19“ monitor situated 0.5 m away from the participant.

Procedure

Participants provided consent and measuring devices were connected. Then instructions for the rating scales were delivered followed by a short practice (six trails). Participants then performed the experiment alone in a dimly lit room.

EDA Measurement

EDA data was acquired using (GSR100C Biopac Systems, Goleta, CA) amplifier (Gain = 20μ Ω, Low Pass = 1.0 Hz, HP:DC). Two 6-mm Ag-AgCl, nonpolarizable electrodes (TP–TSD203 Biopac Systems, Goleta, CA) were placed on the hypothenar eminence of the left index and middle fingers with 0.5% saline electrode paste (GEL101, Biopac Systems, Goleta, CA).

EDA Preprocessing

The 6 s of each picture presentation were divided into 12 half seconds. Mean EDA response was computed for each half second using ANSLAB 2.4 (2006–2014). EDA activity was measured by log transformation of the maximum change occurring between 1 and 4 s after picture onset: maximum and minimum half-second EDA average scores (microsiemens, μs) were computed both from the time interval between 1 s and 4 s after picture presentation, subtracted and transformed (log [(max − min) + 1]); Bradley et al., 2001; Lang et al., 1993).

Data Reduction and Statistical Analysis

All self-report scores were standardized ($M = 0$, $SD = 1$) across participants so differences in standard deviations of the self-report scales would not affect estimators of regression coefficients. However, using raw scores instead of standardized scores did not change the pattern of results and did not change the significance of the hypothesis testing.

[1] Picture numbers in the IAPS: 1050, 1321, 1441, 1505, 1525, 1605, 1650, 1710, 1720, 2055, 2057, 2075, 2205, 2279, 2309, 2347, 2388, 2491, 2606, 2620, 2630, 2720, 2730, 3010, 3053, 3110, 3168, 3190, 3211, 3550, 4005, 4310, 4525, 4664, 4668, 4669, 4697, 5000, 5200, 5300, 5470, 5621, 5740, 5760, 6212, 6230, 6242, 6900, 6910, 7002, 7004, 7031, 7040, 7054, 7058, 7351, 7509, 7900, 8021, 8160, 8163, 8230, 8480, 8501, 9000, 9001, 9090, 9253, 9360, 9403, 9417, 9432.

Item-level analysis. All self-report and EDA scores were averaged across participants resulting in 72 data points. Each data point reflected a specific picture, mean EDA for that picture, and four means of self-report scores (arousal, PL, UN, and overall feelings) for that picture.

To evaluate EDA's variance explained exclusively by arousal we used three approaches:

1. Estimating the increment of EDA's variance explained by arousal scores over variance explained by PL and UN. We used a significance test for regression coefficient of arousal when it was entered simultaneously with PL and UN to the same model.

$$\widehat{EDA} = b_0 + b_{PL}(PL) + b_{UN}(UN) + b_{AROUSAL}(AROUSAL) .$$

Statistical significance of this test is equivalent to testing the difference between full and partial models ($R^2_{Y.AROUSAL,\,PL,UN} - R^2_{Y.PL,UN}$) (Pedhazur, 1997).

2. Because conclusions may involve accepting the null hypothesis (no difference between arousal and pleasure/displeasure), in addition to relying on hypothesis testing and the significance test, we also compared the two models by Bayesian model selection criterion, or the Bayesian information criterion (BIC; Schwarz, 1978; see also Burnham & Anderson, 2004). BIC is computed for the partial (PL, UN) and the full (PL, UN, AROUSAL) model. Application of other model fit criteria (e.g., AIC, Akaike, 1974) provides the same results and conclusions.

3. Partial regression coefficients (i.e., 1) and Bayesian criterion (i.e., 2) are informative about whether the model including arousal and PL, UN is different from the model including only PL and UN, without arousal. However, these measures are less informative about the degree to which arousal contributes to PL and UN. To estimate the effect sizes and contribution of arousal over PL and UN we computed the squared semipartial correlation (this is the squared correlation between arousal and EDA when variance related to PL and UN is partialed out from arousal scores) and compared it to the square correlation of PL and UN. Squared semipartial measures were computed using Type II error to estimate the unique component of arousal. It is reported in results as spr^2_{II}.

Within individual level analysis. To evaluate EDA's variance explained exclusively by self-reported arousal at the individual level we take the same approach as in item analysis using the following measures: (a) Estimating an increment of EDA's variance explained by arousal scores over variance explained by PL and UN was done by a simultaneous multilevel regression with the participant's intercept as a between-second level random variable and PL, UN and arousal as within-second level fixed variables (PROC MIXED, SAS). Such analysis provides estimation for the average association between the predictor (e.g., arousal) and the criterion (e.g., EDA) within participant (Van den Noortgate & Onghena, 2006). (b) The Bayesian factor was computed based on the multilevel regression model. (c) Effect sizes and partial effect sizes were computed according to Cohen's f^2 for linear mixed models with random effect (Selya, Rose, Dierker, Hedeker, & Mermelstein, 2012).

Results

Norm Validity

Norm validity for arousal was computed by a correlation of arousal ratings (72 pictures) with IAPS norms and showed high association, $r = .80$, similar to previous research (Ito et al., 1998).

Item Level Analysis

Association of PL, UN with self-reported arousal. PL + UN scores were associated highly with arousal scores, $r = .88$, $p < .001$ (see Figure 2a). Given that the reliability of arousal scores (a correlation of arousal scores with themselves) can reach a maximum of $r = .86$ (Ito et al., 1998), the correlation here between PL + UN and arousal explained all nonerror variance in arousal scores, suggesting a unity between PL + UN and arousal in this analysis. Together in the same regression model, both PL and UN scores are positively associated with arousal scores, explaining 77% of its variance, $b_{PL} = .55$, CI [.45, .66], $t(69) = 10.6$, $p < .001$, $spr^2_{II} = .37$; $b_{UN} = .60$, CI [.52, .68], $t(69) = 15.2$, $p < .001$, $spr^2_{II} = .75$.

Association of PL, UN, and arousal with EDA. PL + UN and arousal show similar significant association with EDA, $r = .52$, $p < .001$; $r = .52$, $p < .001$, respectively (see Figure 2b). Together in the same regression model, both PL and UN scores are positively associated with EDA, explaining 28% of its variance, $b_{PL} = .001$, CI [.0003, .002], $t(69) = 2.53$, $p < .01$, $spr^2_{II} = .06$, $b_{UN} = .002$, CI [.001, .003], $t(69) = 5.22$, $p < .001$, $spr^2_{II} = .21$.

Unique EDA's variance explained by arousal. No significant evidence was found for distinct variance explained by arousal beyond PL and UN, $b_{arousal} = .002$, CI [−.0005, .004], $t(68) = 1.59$, $spr^2_{II} = .025$, $BIC_{EDA=PL,UN} = -633.3$, $BIC_{EDA=PL,UN,AROUSAL} = -624.4$ (smaller BIC implies better fit).

Within Individual Level Analysis

Association of PL, UN, and arousal with EDA. Both PL + UN and arousal ratings show significant association with EDA, $b_{PL+UN} = .001$, CI [.0008, .0018], $F(1, 2{,}129)^2 = 24.82$, $p < .001$, $f^2 = .009$; $b_{arousal} = .001$, CI [.0007, .0017], $F(1, 2{,}129) = 18.86$, $p < .001$, $f^2 = .009$, respectively. Together in the same model, PL and UN are positively associated with EDA, $b_{PL} = .0006$, $F(1, 2{,}128) = 5.74$, $p < .01$; $b_{UN} = .001$, $F(1, 2{,}128) = 29.06$, $p < .003$.

Unique EDA's variance explained by arousal. No significant evidence was found for unique variance explained by arousal, $b_{arousal} = .0005$, $t(2{,}227) = 1.87$, $BIC_{EDA=PL,UN} = -13319$, $BIC_{EDA=PL,UN,AROUSAL} = -13307$, $spr^2_{II} = .0008$, $f^2 = .00000$.

[2] Default denominators' degrees of freedom for fixed effect in PROC MIXED with random effect (subject intercept) are calculated here as (*N*-*s*-*p*) when *N*-number of observations (30 subjects × 72 pictures), where *s* equals the number of subjects and *p* equals the number of predictors.

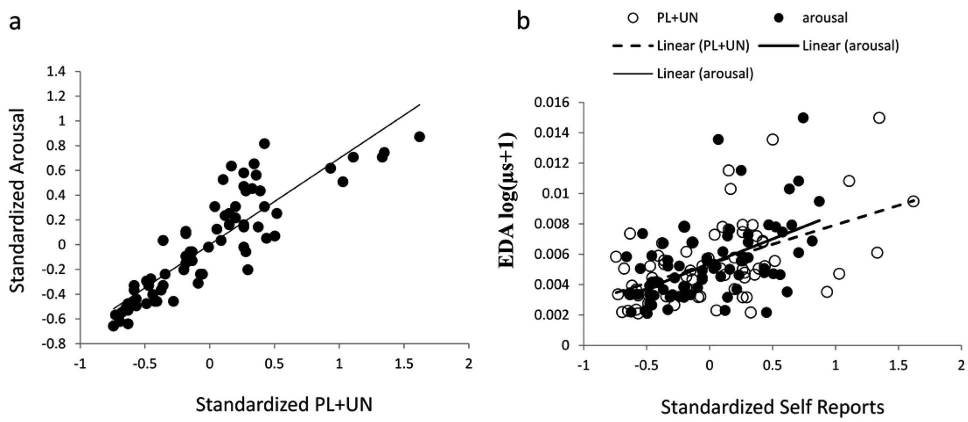

Figure 2. (a) Item analysis scatter plot and regression line for the association between PL + UN and arousal standardized self-report scores. (b) Item analysis scatter plot with regression lines for the association between standardized self-reports (PL + UN and arousal) with EDA scores. PL = pleasant; UN = unpleasant; EDA = electrodermal activity.

Global Emotional Experience

In this study we use the overall feelings scale as a way of focusing participants onto their global feeling state and to ensure that they understood that ratings of PL, UN, and arousal all reflect emotional experience and not the semantic/evaluative dimensions of the stimuli. Next we examined the relative association between arousal and PL and UN with ratings of overall feelings. Such an analysis is informative because arousal may be a distinct component of global emotional experience even though it does not provide a better estimation specifically of physiological arousal.

Item level analysis.

Association of PL, UN, and arousal with overall feelings. PL + UN and arousal showed similar high association with overall feelings, $r = .98$, $p < .001$; $r = .92$, $p < .001$, respectively. Together in the same regression model, both PL and UN scores were positively associated with overall feelings, explaining 96% of its variance, $b_{PL} = .72$ CI [.67355, .77446], $t(69) = 28$, $p < .001$, $spr^2_{II} = .38$, $b_{UN} = .87$, CI [.83303, .90953], $t(69) = 45$, $p < .001$, $spr^2_{II} = .96$. There was, however, weak evidence for distinctive variance accounted for by arousal, $b_{AROUSAL} = .33$, CI [.25465, .42051], $t(68) = 8.12$, $p < .001$, $\text{BIC}_{EDA=PL,UN} = -111$, $\text{BIC}_{EDA=PL,UN,AROUSAL} = -153.7$, explaining 1.5% of unique variance ($spr^2_{Il} = .015$) beyond 96% that was explained by PL and UN.

Within individual level analysis.

Association of PL, UN, and arousal with overall feelings. Both PL + UN and arousal ratings show significant association with overall feelings, $b_{PL+UN} = .86$, CI [.8404, .8816], $b_{PL+UN} = .86$, CI [.8404, .8816], $F(1, 2{,}129) = 6{,}734$, $p < .001$, $f^2 = .75$; $b_{AROUSAL} = .78$, CI [.7529, .8102], $F(1, 2{,}129)$, $= 2{,}862$, $p < .001$, $f^2 = .81$, respectively. Together in the same model, PL and UN were positively associated with overall feelings, $b_{PL} = .61$, CI [.5927, .6338], $F(1, 2{,}128) = 3{,}415$, $p < .001$, $spr^2_{II} = .37$, $b_{UN} = .77$, CI [.7548, .7945], $F(1, 2{,}128) = 5{,}860$, $p < .001$, $spr^2_{II} = .59$. Once again, arousal explained a small distinct portion of variance in overall feelings, $b_{AROUSAL} = .32$, $F(1, 2{,}127) = 606$, $p < .001$, $\text{BIC}_{EDA=PL,UN} = 2{,}470$, $\text{BIC}_{EDA=PL,UN,AROUSAL} = 1{,}943$, accounting for 4% of unique variance ($f^2 = .06$) over the 75% explained by PL and UN.

General Discussion

In this work we examined whether arousal and valence are distinct components of emotional experience. The hypothesis that predicts valence and arousal to be distinct feelings in the context of picture viewing relies on two types of evidence: the dissociative relationship of arousal and valence with physiological arousal (e.g., Lang et al., 1993), and the limited association between self-reports of valence and arousal (Ito et al., 1998). The current experiment addressed this question with three tests. First, we showed that self-reports of arousal did not contribute to prediction of physiological arousal over the contributions of feelings of valence. Second, we replicated previous findings (Kron et al., 2013) and showed a strong association between self-reports of PL + UN and arousal. Third, we demonstrated there was diminishingly small relative contribution of arousal over valence in predicting feelings at two levels: the items level and the within-individual level.

The results from the item-level analysis reflected the aggregation of PL and UN scores across participants and therefore are primarily informative for the experimental designs that model emotions by selecting stimuli according to the IIA's item's norms of valence and arousal. The model of emotions one uses is most accurately reflected in the rules or criteria by which stimuli are assigned to the different conditions in research design. Currently, criteria that are used to assign stimuli to experimental conditions are based on self-reports of emotional experience (e.g., the IIA norms of valence and arousal). Due to this, the role of the IIA norms in research design and the accumulation of knowledge about emotions in terms of valence and arousal should not be underestimated.

At the level of items, PL and UN feelings predict both physiological activity (EDA) and self-reports of arousal. Self-reported arousal showed a small unique variance in predicting overall feelings (1.5% over 96% explained by valence). These results suggest that, arousal does not appear to be substantially distinct from dual unipolar valence. Although previously used to establish the distinction between valence and arousal (e.g., Lang et al., 1993), this item based is based on aggregated scores across individuals. Therefore they may not answer the question of whether humans are able to distinguish between arousal and valence as separate components of feelings during a specific emotional experience. When focused at the resolution of individual (trial-based) responses, we found only weak evidence for such a distinction even. Feelings of arousal and valence showed no difference in predicting physiological arousal, with no evidence of self-reported arousal accounting for unique variance over valence. In predicting self-reports of overall feelings, we found a 4% unique variance contributed by arousal over the 76% explained by bivariate valence. This small portion of unique variance may represent a distinct feeling of arousal, or some other aspect of experience that is captured in one's assessment of overall feelings, such as uncertainty or ambivalence. Thus, if arousal is dissociated from bipolar valence, then it is not because participants can introspect on highly distinct types of feelings. Given such a small contribution of self-reported arousal, the results of the individual-level analysis question the importance of the unique component of emotional experience of arousal in human emotional experience, at least those responses that are evoked by an external stimulus.

By claiming that self-reported arousal does not show a substantial contribution over valence (pleasure and displeasure) we do not suggest that physiological arousal is not a unique subsystem of emotional response. In the emotion literature, the term *arousal* is used in at least three different contexts. First, arousal is used to describe systems that relate to physiological responses (e.g., reticular activation system, autonomic nervous system; Jones, 2003). The second meaning of the term refers to the name that is given to a factor in a latent variable models (e.g., Yik, Russell, & Barrett, 1999). In the latent variable, participants are not asked to report about the level of arousal but, arousal is a title given by the researcher to a group of items that are believed to reflect commonality of arousal. Finally, the term arousal refers to the observed (not latent) self-report scale (e.g., Lang et al., 1999) on which participants are asked to report about feelings named arousal. In this work we do not argue against the specificity of physiological arousal and we do not claim that there is no such experience as arousal that is independent of valence (such as a feeling of arousal when a person walks up the stairs). Our claim is concerned with the meaning of arousal in the context of observed self-report scales, such as those implemented in the IIA stimuli pools. Within the context of emotional response, arousal and pleasure/displeasure are not separable in emotional experience, and, consequently, using self-reports of arousal in addition to bivariate valence (unipolar positive and negative affect) valence to model emotions, as is done in numerous experiments is empirically redundant. When employing bipolar valence (pleasant vs. unpleasant), rather than capturing a unique aspect of emotional experience, the additional measurement of self reported arousal makes up for valence information loss inherent to bipolar scales, such as mixed valence experiences (Kron et al., 2013).

In this study we used the IAPS pictures. There are several limitations that come with using only emotional pictures. First, the IAPS pool includes only a limited number of content categories. Second, there is a limit to the degree that pictures can elicit a strong emotional response (e.g., attending a funeral vs. viewing a picture of it). Although the other two emotion stimuli pools (IADS—for audio recording; ANEW—for English words) show the same relationship between bipolar valence and arousal, it is important to support these findings in other content categories and modalities. In addition, in the current study we examined only immediate response to pictorial stimuli. Thus, it is not clear whether the strong link between self-reported arousal and valence also will be manifested in prolonged affective states that are not necessarily coupled with external stimuli (e.g., moods, current affect; Yik et al., 1999).

Another limitation of our study is that the models of conscious experience of emotions are estimated by self-reports, which themselves are only proxies to their underling qualia and might involve metacognitive description and interpretation of the experience, biases and thus reflect not only emotions but also attitudes toward events (Robinson & Clore, 2002). In this work we tried to eliminate such alternative explanations by using a physiological criterion (EDA) and self-report scales that focus on reporting actual feelings. However, the gap between self-reports and experience itself is valid and challenging and should be kept in mind when applying models based on self-reports of experience of emotions.

Finally, our results suggest a strong analytical link between the bipolar valence-arousal model and the two-unipolar valence model. Although strongly related (bipolar valence $\sim$ PL $-$ UN; arousal $\sim$ PL $+$ UN), the two models result in a different parsing of the emotional space and influence how physiological and neural activation associated with the emotional response is mapped and interpreted. Although the results earlier suggest that the assumptions underlying the bipolar valence-arousal model are not justified, at the same time, they emphasize the challenge in selecting between the two models, given the strong analytical relationship between them. One factor in deciding between the two models is the research on mixed emotions (simultaneous activation of pleasure and displeasure). According to the bipolar valence-arousal model, pleasure and displeasure are mutually exclusive, while the two-unipolar model assumes separate representations of pleasure and displeasure (and consequently permits its parallel activation). Thus, if mixed emotions exist, they strongly support the two-unipolar valence model over the bipolar valence arousal model. Currently, accumulating evidence suggests that participants can report about experience of mixed emotions (e.g., Ersner-Hershfield, Mikels, Sullivan, & Carstensen, 2008; Hunter, Schellenberg & Schimmack, 2008; Kron et al., 2013; Larsen & McGraw, 2011; Larsen, McGraw & Cacioppo, 2001; Stanley & Meyer, 2009; Williams & Aaker, 2002). However, strong evidence for parallel activation of pleasure and displeasure is still missing, given that all the previous evidence for mixed emotions is primarily based on self-reports and it is not clear if pleasure and displeasure are activated simultaneously or serially (Larsen & McGraw, 2011). Future research that would provide support for parallel or serial activation of pleasure and displeasure in mixed emotions has the potential to contribute substantially to the decision between the models.

References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*, 716–723.

Anderson, A. K., Christoff, K., Stappen, I., Panitz, D., Ghahremani, D. G., Glover, G., . . . Sobel, N. (2003). Dissociated neural representations of intensity and valence in human olfaction. *Nature Neuroscience*, *6*, 196–202.

ANSLAB Version 2.4. (2006–2014) [Computer software]. Frank H. Wilhelm, Jens Blechert, Vitaliy Kolodyazny, Michael Liedlgruber & Peter Peyk. http://www.anslab.net

Averill, J. R. (1975). A semantic atlas of emotional concepts (Manuscript No. 421). *JSAS Catalog of Selected Documents in Psychology, 5,* 330.

Borgatta, E. F. (1961). Mood, personality, and interaction. *The Journal of General Psychology, 64,* 105–137. http://dx.doi.org/10.1080/00221309.1961.9920431

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. http://dx.doi.org/10.1037/1528-3542.1.3.276

Bradley, M. M., & Lang, P. J. (1999). *Affective norms for English words (ANEW): Instruction manual and affective ratings* (Tech. Rep. No. C–1). Gainesville: Center for Research in Psychophysiology, University of Florida.

Bradley, M. M., & Lang, P. J. (2007). *The international affective digitized sounds (2nd edition, IADS–2): Affective ratings of sounds and instruction manual* (Tech. Rep. No. B–3). Gainesville: Center for Research in Psychophysiology, University of Florida.

Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference understanding AIC and BIC in model selection. *Sociological Methods & Research, 33,* 261–304.

Cacioppo, J. T., & Berntson, G. G. (1994). Relationship between attitudes and evaluative space: A critical review, with emphasis on the separability of positive and negative substrates. *Psychological Bulletin, 115,* 401–423. http://dx.doi.org/10.1037/0033-2909.115.3.401

Cacioppo, J. T., Gardner, W. L., & Berntson, G. G. (1997). Beyond bipolar conceptualizations and measures: The case of attitudes and evaluative space. *Personality and Social Psychology Review, 1,* 3–25.

Ersner-Hershfield, H., Mikels, J. A., Sullivan, S. J., & Carstensen, L. L. (2008). Poignancy: Mixed emotional experience in the face of meaningful endings. *Journal of Personality and Social Psychology, 94,* 158–167. http://dx.doi.org/10.1037/0022-3514.94.1.158

Hunter, P. G., Schellenberg, E. G., & Schimmack, U. (2008). Mixed affective responses to music with conflicting cues. *Cognition & Emotion, 22,* 327–352.

Ito, T. A., Cacioppo, J. T., & Lang, P. J. (1998). Eliciting affect using the international affective picture system: Trajectories through evaluative space. *Personality and Social Psychology Bulletin, 24,* 855–879. http://dx.doi.org/10.1177/0146167298248006

Jones, B. E. (2003). Arousal systems. *Frontiers in Bioscience, 8,* s438–s451. http://dx.doi.org/10.2741/1074

Kron, A., Goldstein, A., Lee, D. H., Gardhouse, K., & Anderson, A. K. (2013). How are you feeling? Revisiting the quantification of emotional qualia. *Psychological Science, 24,* 1503–1511. http://dx.doi.org/10.1177/0956797613475456

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1997). *International affective picture system*. Gainesville: Center for the Study of Emotion and Attention, University of Florida.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1999). *International affective picture system (IAPS): Technical manual and affective ratings*. Gainesville: Center for the Study of Emotion and Attention, 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. http://dx.doi.org/10.1111/j.1469-8986.1993.tb03352.x

Larsen, J. T., & McGraw, A. P. (2011). Further evidence for mixed emotions. *Journal of Personality and Social Psychology, 100,* 1095–1110. http://dx.doi.org/10.1037/a0021846

Larsen, J. T., McGraw, A. P., & Cacioppo, J. T. (2001). Can people feel happy and sad at the same time? *Journal of Personality and Social Psychology, 81,* 684–696. http://dx.doi.org/10.1037/0022-3514.81.4.684

Larsen, J. T., Norris, C. J., McGraw, A. P., Hawkley, L. C., & Cacioppo, J. T. (2009). The evaluative space grid: A single-item measure of positivity and negativity. *Cognition and Emotion, 23,* 453–480. http://dx.doi.org/10.1080/02699930801994054

Levenson, R. W. (2003). Autonomic specificity and emotion. *Handbook of Affective Sciences, 2,* 212–224.

Pedhazur, E. J. (1997). *Multiple regression in behavioral research: Explanation and prediction* (3rd ed.). Boston, MA: Wadsworth.

Robinson, M. D., & Clore, G. L. (2002). Belief and feeling: Evidence for an accessibility model of emotional self-report. *Psychological Bulletin, 128,* 934–960. http://dx.doi.org/10.1037/0033-2909.128.6.934

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39,* 1161–1178. http://dx.doi.org/10.1037/h0077714

Russell, J. A., & Feldman Barrett, L. (1999). Core affect, prototypical emotional episodes, and other things called emotion: Dissecting the elephant. *Journal of Personality and Social Psychology, 76,* 805–819. http://dx.doi.org/10.1037/0022-3514.76.5.805

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics, 6,* 461–464.

Selya, A. S., Rose, J. S., Dierker, L. C., Hedeker, D., & Mermelstein, R. J. (2012). A practical guide to calculating Cohen's f2, a measure of local effect size, from PROC MIXED. *Frontiers in Psychology, 3*(111), 2–6. http://dx.doi.org/10.3389/fpsyg.2012.00111

Stanley, D. J., & Meyer, J. P. (2009). Two-dimensional affective space: A new approach to orienting the axes. *Emotion, 9,* 214–237. http://dx.doi.org/10.1037/a0014612

Van den Noortgate, W., & Onghena, P. (2006). Analysing repeated measures data in cognitive research: A comment on regression coefficient analyses. *The European Journal of Cognitive Psychology, 18,* 937–952. http://dx.doi.org/10.1080/09541440500451526

Viinikainen, M., Jääskeläinen, I. P., Alexandrov, Y., Balk, M. H., Autti, T., & Sams, M. (2010). Nonlinear relationship between emotional valence and brain activity: Evidence of separate negative and positive valence dimensions. *Human Brain Mapping, 31,* 1030–1040. http://dx.doi.org/10.1002/hbm.20915

Williams, P., & Aaker, J. L. (2002). Can mixed emotions peacefully coexist? *The Journal of Consumer Research, 28,* 636–649. http://dx.doi.org/10.1086/338206

Yik, M. S., Russell, J. A., & Barrett, L. F. (1999). Structure of self-reported current affect: Integration and beyond. *Journal of Personality and Social Psychology, 77,* 600–619. http://dx.doi.org/10.1037/0022-3514.77.3.600

Appendix

Instructions for Ratings Scales

The instructions, given here, are written in everyday language to improve communication with participants. Each experimenter learned the instructions by heart and all experimenters had 2 days of practice to make sure the instructions were delivered fluently (a Power Point presentation, with all slides and an audio recording, is available by request from A. Kron).

Introduction to Instructions

- "Welcome to the experiment!"
- "In this experiment you will see a set of pictures and I will ask you to report about your feelings, about what you feel while viewing each picture. Now, reporting about your feelings is not always easy. Feelings can be vague. Sometimes it is not clear what to report about when asked to report about your feelings. In these instructions I will explain to you what I mean when I ask you to report about your feelings."
- "Let's begin by looking at these two pictures [participants see a neutral image followed by a picture that elicits strong unpleasant feeling]."
- "Did you notice any difference in your feelings?"
- "I want to be more specific. I am asking you to report about what you **actually felt**. I am not asking **you if the picture** was pleasant or unpleasant, your opinion about the picture, what you expected yourself to feel or what you thought we expected you to feel, but only the actual feeling that you experienced, while viewing the picture."

Overall Feelings Rating Scale

- "Each picture will be followed by the rating scales. This is the first rating scale: the intensity of your feelings" (see Figure A1).
- "Here, you are asked to report about any type of feelings. The questions here is weather you have **detected any type of feeling,** for example, pleasure, displeasure, arousal, or any other type of feelings. You can think of this scale as a volume knob that indicates the intensity of your feelings. If you have not detected any feelings inside you, press here [experimenter points on the screen to 0]. If you are not sure if you have detected any feelings, also press here [experimenter points on the screen to 0]. If you have detected some feelings, you can rate their intensity here [the rest of the scale]. If you experienced ambivalent feelings or several types of feelings at the same time, rate the intensity of the strongest type of feeling you experienced."
- "However, make sure that you move from here [pointing to 0] to here [pointing to 1] only when you are **certain** that you have detected some feeling inside you."

Figure A1. Rating scale for overall feelings. Ratings scales for other emotions where identical except for the title. We used the titled "pleasure", "displeasure," and "arousal" for the other three types of feelings.

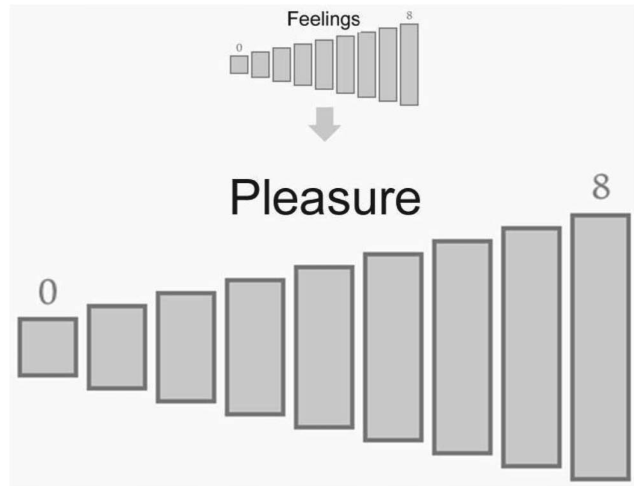

Figure A2. Scale for pleasant feelings. This scale emphasized that participants were to report about their feelings and not semantic knowledge.

Pleasure Rating Scale

- "The first scale asked whether you detected any feelings inside you. Now we move to the next scale; here I ask you about specific feelings."
- "This is the scale for pleasure, happiness, and other positive feelings" (pointing at Figure A2).
- "If you have not detected any real pleasure, happiness or other positive feelings inside you, press here [pointing to 0]. If you are not sure if you have detected any positive feeling, also press here [pointing to 0]. If the pleasant feeling was vague press here [pointing to 0]. You can move from here [pointing to 0] to here [pointing to 1] only if you are sure that you have

experienced some pleasant feeling. Once you move here [pointing to 1] you can indicate the intensity of the pleasant feelings on the rest of the scale. "

Displeasure Rating Scale

- "This is the scale for displeasure, sadness, and other negative feelings (same scale as Figure 2 except with the title 'displeasure' instead of pleasure). If you have not detected any real displeasure, sadness, or other negative feelings inside you, press here [pointing to 0]. If you are not sure if you have detected any negative feeling, also press here [pointing to 0]. If the unpleasant feeling was vague press here [pointing to 0]. You can move from here [pointing to 0] to here [pointing to 1] only if you are sure that you have experienced some unpleasant feeling. Once you moved here [pointing to 1] you can indicate the intensity of the unpleasant feelings on the rest of the scale."

Arousal Ratings Scale

- "This is the scale for feelings or sensation of arousal (by arousal I mean how excited, wide awake, frenzied, jittery you feel while viewing a picture [same scale as Figure 2 only with the title 'arousal' instead of 'pleasure']. If you have not detected any real sensation of arousal, inside you, press here [pointing to 0]. If you are not sure if you have detected any feeling or sensation of arousal, also press here [pointing to 0]. If you have detected actual feeling of arousal you can rate its intensity here [the rest of the scale]. Make sure that you move from here [pointing to 0] to here [pointing to 1] only when you are certain that you have detected real sensation or feeling of arousal inside you."
- [Begin practice.]

Received August 25, 2013
Revision received August 12, 2014
Accepted September 22, 2014 ■