

## Transforming valences through transitive inference: How are faces emotionally dissonant?

Micah Amd <sup>a,b</sup> and Bryan Roche<sup>a</sup>

<sup>a</sup>Department of Psychology, Maynooth University Maynooth, Maynooth, Co. Kildare, Ireland; <sup>b</sup>Department of Psychology, Federal University of São Carlos, São Carlos, SP, Brazil

### ABSTRACT

Information that is emotionally incongruous with self-concepts can produce feelings of unease. This implies that embedding incongruous information in newly formed relational structures would have little effect on their previous emotive properties. Alternatively, Relational Frame Theory highlights the importance of contextualized stimulus-stimulus relations, where the structure of a relational series is key in determining the function of its elements. To see whether series membership can mitigate ‘dissonance’ when a salient element is employed, the present investigation trained and tested a seven-term relational series ( $X>A>B>C>D>E>Y$ ) using blurred faces as stimuli. Specifically, Stimuli X, A, B, D, E and Y were blurred unfamiliar faces and Stimulus C constituted of the participant’s own blurred face, assumed to be more salient than the former. To assess how the valences of the related stimuli were transformed by relational series membership, self-report ratings and electroencephalographic (EEG) recordings were collected before and after training the  $X>A>B>C>D>E>Y$  series. These pre vs. post contrasts revealed that, for unfamiliar faces, stimulus valence transformed as a function of relational structure. Conversely, the lack of difference in pre vs. post contrasts of Stimulus C, which maintained a high valence, suggest that relational series membership may not suffice to mitigate emotionally dissonant information.

### ARTICLE HISTORY

Received 9 September 2015  
Accepted 23 September 2016  
First Published Online 3  
November 2016

### KEYWORDS

Emotional dissonance;  
Frontal alpha asymmetry;  
Relational frame theory;  
Transitive inference; Valence

Our perceptions of emotionality from others’ faces are typically determined by the various observable characteristics of the latter, even when we are unaware of which characteristics we may be attending to per se (e.g., Dimberg, Thunberg, & Elmehed, 2000; Esteves & Öhman, 1993). Indeed, even faces normatively rated as exhibiting “neutral” expressions have been shown to evoke some degree of differential responding (e.g., Mignault & Chaudhuri, 2003). If no emotional information can be gleaned from one’s face, however, its emotional properties can be derived based on the relationships it shares with other faces within a relational structure (Amd & Roche, 2015, 2016; Silveira et al., 2015). For instance, if provided the relational information “Adam is happier than Bob”, or Adam  $>_{\text{happy}}$  Bob (where “Adam” and “Bob” are emotionally masked faces) then, holding all other factors constant, it is likely

that an individual will respond to Adam as happier than Bob (e.g., Amd, 2014; Amd & Roche, 2015; for similar effects with non-face stimuli, see Dougher, Hamilton, Fink, & Harrington, 2007; Dymond, Schlund, Roche, & Whelan, 2014; Roche & Dymond, 2008). In such cases, the “happiness” functions (valences) of Adam and Bob have transformed in accordance with the newly established relational structure Adam  $>_{\text{happy}}$  Bob. According to one account of symbolic stimulus relations (relational frame theory, or RFT; Hayes, Barnes-holmes, & Roche, 2001), the determining feature in such “transformations” is the context in which the face stimuli were related (e.g., the presence of the “is happier than” phrase), summarized as a transformation of function (TOF) effect (Dymond & Rehfeldt, 2000). There appears to be some evidence that embedding face stimuli in context-specific relational structures can

transform their valences predictably (cf. Amd, 2014; Amd & Roche, 2016).

This line of investigation is important for at least two reasons. First, research on the perception of emotion on faces has typically employed face stimuli that exhibit natural (Haxby & Gobbini, 2007; Wheatley, Weinberg, Looser, Moran, & Hajcak, 2011) or artificial (Seyama & Nagayama, 2007) emotional expressions. Such studies typically explore the emotion-eliciting functions of various facial characteristics (e.g., Leopold & Rhodes, 2010). By employing faces whose emotional-eliciting features have been concealed (as per the recommendations of Bassili, 1979), however, it becomes possible to understand the effect of relational structure membership on their valences, rather than the perception of facial features per se (Amd & Roche, 2015). Continuing along the line of some recent findings (e.g., Amd & Roche, 2015; Silveira et al., 2015), emotional TOF studies using faces have highlighted the importance of membership within newly established relational structures (versus pre-existing ones; see Lakens, Semin, & Foroni, 2012) in manipulating stimulus valence (Amd & Roche, 2016). This is why blurred faces were employed as stimuli in the present study, given previous findings showing happiness/valence TOF effects as most pronounced when emotionally masked faces are embedded in relational structures (Amd, 2014; Amd & Roche, 2015; but see Amd, Barnes-Holmes, & Ivanoff, 2013, for a non-face-related exception). This may have to do with the conspecific nature of faces, which humans can discriminate accurately from early in life (e.g., Bonatti, Frot, Zangl, & Mehler, 2002). This may also explain why “happiness” may be most readily perceived across a facial dimension (e.g., Ekman & Friesen, 1971). By manipulating the valence of otherwise homogeneous faces through relational categorization (cf. Dymond & Roche, 2013), we can better understand how doing so can affect emotional face perception when the faces involved are emotionally “neutral” (e.g., Todorov & Engell, 2008).

The current study also investigates a key prediction arising from research into “emotional dissonance” (Jansz & Timmers, 2002), a term that refers to the feeling of unease produced by information that is emotionally incongruous with notions related to one’s self. At a process level, such dissonance may arise when emotionally disparate elements participate in the same relational knowledge structure, yielding feelings that are conflicting. Some theorists (e.g., Barnes, 1994; Salzinger, 2008) have proposed that

membership within an appropriate relational structure should suffice to alter the emotional properties of relational elements, regardless of formal stimulus characteristics. The concept of “emotional dissonance” would suggest otherwise, however. Specifically, the literature on emotional dissonance suggests that any emotional incoherence obtained between an emotional stimulus and one’s own self creates a dissonance, and along with this come feelings of unease. Real-life occurrences of emotional dissonance can be found in a variety of situations—for instance, among call centre workers (Lewig & Dollard, 2003). It is not uncommon for such workers to encounter clients who are rude, with whom the former must nevertheless maintain an air of professionalism while inhibiting more natural tendencies (e.g., Greenberg & Alge, 1998). The underlying dissonance arises from not being able to respond unequivocally to an irate client, and is strongly related to high turnover rates and absenteeism in the industry (James, 1998). In plain terms, the negative emotionality of the external stimulus (a rude client) contradicts the generally positive emotionality of self-held notions (of competence) to yield feelings of dissonance.

The example highlights how diametrically valenced events/stimuli can participate in the same knowledge structure as that of related elements and still maintain their disparate functions (cf. Dymond, Roche, Forsyth, Whelan, & Rhoden, 2007). In such cases, TOF appears limited since the salience of the negative emotional event (client rudeness) does not seem to be mitigated by self-beliefs of competence, which are generally positive. This suggests that deriving stimulus–stimulus relations may not suffice to alleviate dissonance if the related stimuli were highly salient a priori. In other words, TOF may not be relevant when series consist of highly salient elements. On balance, several findings in the TOF literature (Amd & Roche, 2015; Dymond & Roche, 2013; Dymond et al., 2007; Hayes et al., 2001; Silveira et al., 2015) suggest that relational structure membership plays a determining role in altering the emotional properties of embedded elements, even when salient or familiar stimuli are used (e.g., Amd et al., 2013; Fields, Arntzen, Nartey, & Eilifsen, 2012). Indeed, if participating in relational knowledge structures was unimportant in mitigating valences of any of its elements, then the “meaning” of abstract stimuli could not be determined by semantic networks, which clearly do have the capacity to alter stimulus meaning (Harnad, 1990; but also see Stemmer, 1973).

In order to better understand the effect of introducing a salient stimulus amongst a set of topographically similar (but less salient) stimuli, the present study established a relational series with an already salient element (i.e., the participant's own face) as a member. This allowed us to determine whether the emotional valence of an already salient stimulus could be altered in accordance with the transformations for other members of a series (Amd & Roche, 2016). In turn, this will inform us as to whether valence TOF is affected more so by the highly salient emotional element (which was the participant's own face in the present experiment) through associative transfer processes (e.g., Bialer, 1961) or the structure of the relational series involved (Amd & Roche, 2015). Either finding will highlight the role of TOF in relation to emotional dissonance.

### **Assessing TOF through contextually controlled transitive inference**

One means of assessing valence TOF—namely, relational structures—is through contextually controlled transitive inference, or TI (e.g., Amd & Roche, 2016; Dougher et al., 2007). Briefly, TI refers to the derivation of relations between stimuli that have never been paired together formally but are nevertheless related by at least one mediating node (Fields, Adams, Verhave, & Newman, 1990; Munnely, Dymond, & Hinton, 2010; Vasconcelos, 2008). Following from our previous example, if one learns that “Adam is happier than Bob”, and “Bob is happier than Cindy”, deriving “Adam is happier than Cindy” would constitute a demonstration of TI, where the mediating stimulus is Bob. Additionally, if one were to then present the “Adam/Cindy” stimulus pair alongside a cue meaning “is unhappier than”, the inferred discrimination would reverse, as in “Cindy is unhappier than Adam”. Such TIs are “contextually controlled” as the correct discrimination (happier than [Adam +/Cindy–] or unhappier than [Cindy+/Adam–]) is contingent upon the specific contextual cue present (“happier than” or “unhappier than”). Deriving these transitive relations can transform the “happiness” functions of Adam, Bob, and Cindy in accordance with a  $A_{\text{Adam}} >_{\text{happy}} B_{\text{Bob}} >_{\text{happy}} C_{\text{Cindy}}$  relational structure. That is, given “Adam is happier than Bob”, and “Bob is happier than Cindy”, Adam should be responded to as the most positively valenced, Bob as intermediately valenced, and Cindy as the most negatively valenced (Amd & Roche, 2016). By

embedding emotionally masked face stimuli in a relational structure in such manner, it becomes possible to determine the extent to which term valences transform.

A recent study demonstrated the emergence of individual valence gradients in accordance with a five-term series using face-type stimuli (Amd & Roche, 2016). In the study, participants first rated five face silhouettes (categorized as A, B, C, D, and E) along a 5-point Likert scale (ranging from “happy” to “unhappy”) as their electroencephalographs (EEGs) were recorded. These were analysed to yield “baseline” valences of the five face stimuli. Next, a sequence of conditional discriminations established two abstract shapes as contextual cues functionally equivalent to the phrases “is happier than” and “is unhappier than”, respectively.<sup>1</sup> Next, participants were shown the A/B, B/C, C/D, and D/E pairs in the presence of the *happier than* cue, where the discriminations reinforced were A+/B–, B+/C–, C+/D–, and D+/E–. This was followed by tests for functional symmetry and transitivity, which involved presenting the pairs described above along with the A/C, B/D, and A/E pairs in the presence of both *happier than* and *unhappier than* cues without corrective feedback. Finally, participants had to once again rate the A, B, C, D, and E stimuli as their EEGs were collected. Among the participants who correctly derived the transitive relations, contrasting the baseline versus post ratings and EEGs revealed the valences of the A, B, C, D, and E stimuli to fall along gradients corresponding with the structure of the relational series established (i.e.,  $A >_{\text{happy}} B >_{\text{happy}} C >_{\text{happy}} D >_{\text{happy}} E$ ).

One finding worth noting was the robust outcomes along the EEG metric used, which involved measuring hemispheric differences across individualized alpha bands over anterior regions (frontal alpha asymmetry, or FAA; see Bazanova & Vernon, 2014; Klimesch, 1999). Briefly, FAA describes the power difference in individually determined “alpha” bandwidths between the left and right anterior regions of the brain, where increased levels of alpha activation over the left lateral prefrontal cortex (LPFC; relative to right LPFC) is associated with positive/approach-related affect, and increased activation over the right LPFC (relative to left LPFC) is associated with negative/avoidance-related affect (Briesemeister, Tamm, Heine, & Jacobs, 2013; Gable & Harmon-Jones, 2008; Hagemann, Naumann, Becker, Maier, & Bartussek, 1998; Papousek, Reiser, Weber, Freudenthaler, & Schultze, 2012). In Amd and Roche's (2016) study, stimulus-elicited

FAAs appeared sensitive to the newly established valence gradients for the stimuli in the series. Given that FAAs can differentiate between valenced stimuli in humans from as early as two years of age (Davidson & Fox, 1989) and are arguably less broad in scope (Briesemeister et al., 2013, but see Coan & Allen, 2004) than other EEG metrics of valence (such as event-related potentials; see Amd et al., 2013; Patel & Azzam, 2005), the authors concluded that the observed FAA gradients might be reflective of derived valence TOFs.

The findings reported by Amd and Roche (2016) constituted the first published demonstration of individual valence gradients across face stimuli within a relational structure using explicit, implicit, and electrophysiological measures concurrently. Importantly, the changes in stimulus valence were *not* due to differential pairing histories between the faces and pre-experimentally salient (unconditioned) stimuli but the relational structure of the series established. This finding stands apart from the extensive literature on “semantic generalization” (cf. Feather, 1965) for at least two reasons. First, as there was no target stimulus that was emotionally conditioned a priori, emotional functions could not generalize along the basis of pre-existing stimulus associations (but see Boyle, Roche, Dymond, & Hermans, 2016). Second, when comparing the pre and post valences measured for the face stimuli, the valences for some stimuli (particularly Stimulus A) were shown to be enhanced in the post condition, countering common knowledge regarding generalization gradients (e.g., Cofer & Foley, 1942; Hermans, Baeyens, & Vervliet, 2013). On balance, the “enhancement” may have been driven at least partially by the “end-anchor” effect (Wynne, 1995), which simply means that the stimulus whose selection was always reinforced (Stimulus A) and the stimulus whose selection was never reinforced (Stimulus E) evoked differentially ratings/FAAs due to their differential reinforcement histories.

Two additional methodological limitations of the Amd and Roche (2016) study are worth noting. First, the sample employed was quite small ( $n = 10$ ), from whom only 6 participants were able to produce the correct transitive relations. Second, the ratings/EEGs acquired for Stimulus C (the intermediate term in the A–B–C–D–E series) were inconsistent across participants, even among those who demonstrated TI. While this may have been an artefact of the small sample size and/or the position of the stimulus within the series, an alternative preparation could

involve a stimulus that was emotionally salient a priori in order to definitively demonstrate whether the emergent valence gradient was a function of the relational series established. Given that the stimulus employed for this purpose in the current study was an image of the subject’s own face, this also allowed us to examine whether any dissonance in valence within the relational series would persist (given the high emotional valence of the self-face stimulus), or whether it could be transformed in accordance with the structure of the established series.

The current research addresses the above limitations while expanding upon the findings of the Amd and Roche (2016) study in two important ways. First, the relational series trained here involved seven stimuli (X, A, B, C, D, E, Y) rather than five, where only the intermediate A–B–C–D–E stimuli were assessed. This was done in order to mitigate for potential end anchor effects that may have confounded the results in Amd and Roche (2016). Second, the intermediate term in the series (Stimulus C) consisted of the participant’s own face, which was likely to be highly emotionally salient (Sifferlin, 2013). By incorporating such a stimulus, we could determine whether establishing a relational series in the presence of a valence-occasioning context (similar to that enacted by the phrase “is happier than”) would suffice to transform the valences of all stimuli within the structure, even a highly salient one. If so, we could conclude that TOF in relational knowledge structures is a key process in alleviating effects borne of emotional dissonance (by placing related elements along a similar valence gradient). Alternatively, it may be the case that the pre-existing emotional properties of a highly salient stimulus is *not* altered by membership in a relational series. Such an observation would highlight a boundary condition for the TOF effect, where contextual control over stimulus–stimulus relations may not suffice to transform the valences of stimuli that are already highly salient a priori. Furthermore, if membership within a relational structure does not suffice to alter the valence of its related elements in accordance with the series (cf. Amd & Roche, 2015), we will have a clearer understanding of how dissonant emotions can arise from elements that otherwise participate in the same relational structure.

The present study commenced with participants providing happiness ratings of the A, B, C, D, and E stimuli along a Likert scale while their EEGs were recorded in order to measure baseline valences for

the stimuli. Next, two abstract shapes were established as contextual cues equivalent to the phrases “is happier than” and “is unhappier than”, respectively. This was followed by a series of conditional discriminations designed to establish the relations  $X >_{\text{happy}} A$ ,  $A >_{\text{happy}} B$ ,  $B >_{\text{happy}} C$ ,  $C >_{\text{happy}} D$ ,  $D >_{\text{happy}} E$ , and  $E >_{\text{happy}} Y$ , followed by tests for symmetry and transitivity among the stimuli in the series. Finally, the emotionality of the face-stimuli were once again assessed using the Likert scale and EEG recordings (both measures were adapted from Amd & Roche, 2016).

The predictions made in accordance with the TOF effect were as follows. First, the A and B stimuli were predicted to evoke higher happiness ratings and a more positive FAA relative to baseline measurements, while stimuli D and E were predicted to evoke the opposite pattern. Second, the ratings/FAAs for the A, B, D, and E stimuli were predicted to correspond along the gradient suggested by the  $X >_{\text{happy}} A >_{\text{happy}} B >_{\text{happy}} C >_{\text{happy}} D >_{\text{happy}} E >_{\text{happy}} Y$  series during the post condition only. Satisfying these predictions would indicate that the valences of the blurred faces had transformed in accordance with the relational structure of the series trained. A third prediction involved the valences recorded for the C stimulus (own face). Specifically, if the valence of C post series establishment falls between those recorded for the B and D stimuli, we can conclude that relational categorization can overcome pre-existing stimulus saliences. Alternatively, observing no change in Stimulus C valence would suggest that salient elements maintain their valence within conflicting relational structures. Either finding would be informative in explaining the relation between TOF and emotional dissonance.

## Experimental study

### Method

#### Participants

Twenty-six participants were recruited through personal invitation from among the student body at Maynooth University. Participants were paid €10 (\$13 USD approx.) for participating in the study, which was provided upon their arrival to the laboratory. Two participants were excluded due to the presence of at least 30 seconds of artefact-laden EEG data while four participants withdrew at various stages of the study for personal reasons, leaving a final sample of 12 females and 8 males ( $M = 23.2$  years,  $SD = 1.8$ ). All participants were

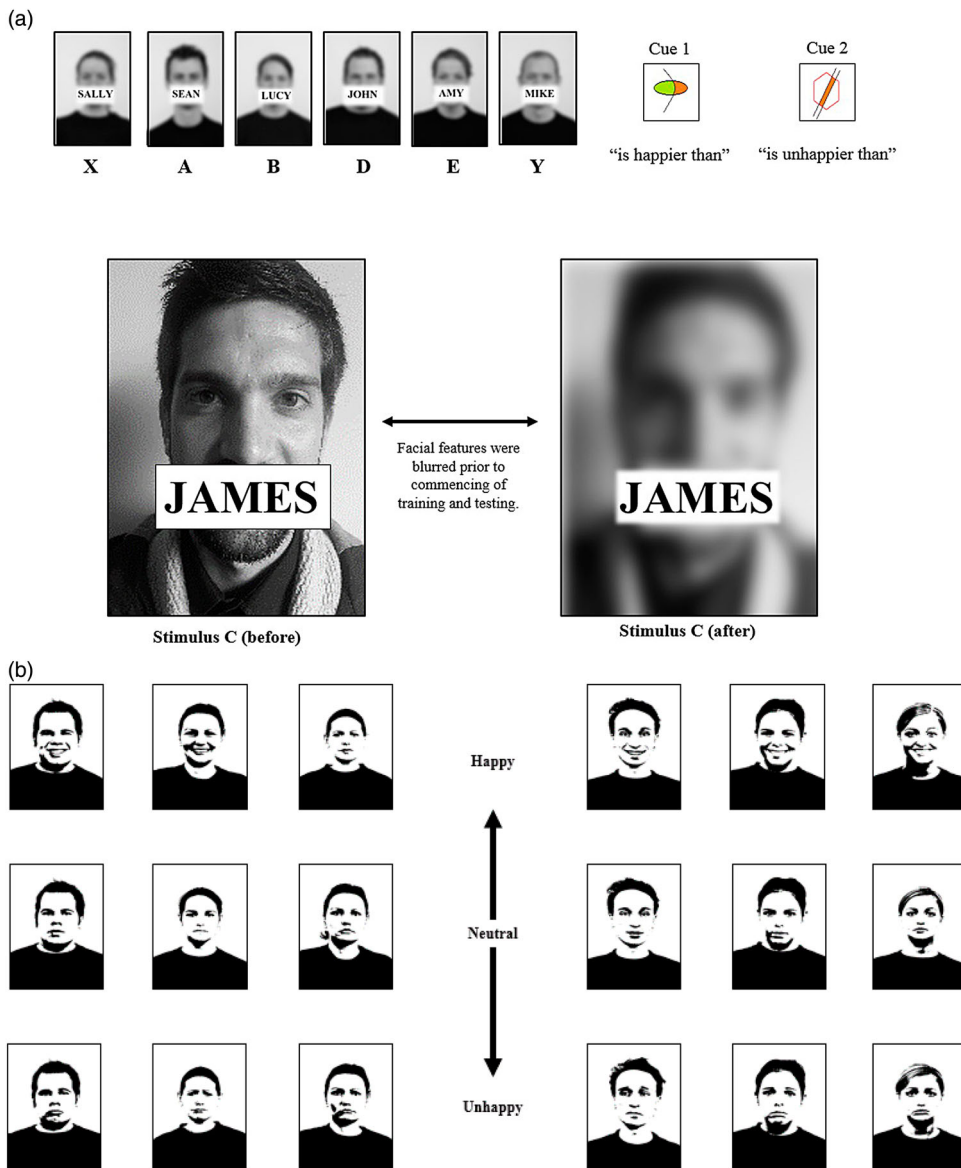
right-handed and had normal or corrected-to-normal vision with no confounding medical history (e.g., drug/medication use, seizures, depression, or schizophrenia). All participants provided written consent for participation. The study was approved by the Maynooth University bio-medical research ethics committee.

### Materials

Face stimuli for the relational training and testing phase were taken from the Radboud Faces Database (Langner et al., 2010). Two sets of face stimuli were employed during the present task. The first set involved 8 faces with happy, 8 faces with neutral, and 8 faces with unhappy expressions, respectively, all of which were employed during the contextual cue training and testing phase only (see Figure 1b). The second set comprised six faces, labelled as A, B, D, E, X, and Y, as well as the participant’s own face, labelled as C—these stimuli were to be used during the transitive inference training and testing stage (see Figure 1a). For the second set of faces, the zygomatic major and orbicularis oculi muscles were blurred in order to make the faces as emotionally homogeneous as possible (Bassili, 1979; Ekman, Davidson, & Friesen, 1990). Additionally, the mouth-regions of all face stimuli in the second set were masked by gender-specific names.

### Electroencephalographic set-up and analysis

EEG activity was recorded using 26 silver/silver chloride electrodes on an amplifier supplied by BrainVision (QuickAmp, BrainProducts, Germany). The electrodes were mounted in an elastic cap and were fastened with a chin strap (Easy-Cap, Herrsching, Germany). Data were collected from electrode sites corresponding with the International 10-20 system for electrode placement. Similar to the set-up in Amd and Roche (2016), EEG activity was recorded over frontal (Fp1, Fp2, F3, F4, F5, F6, F7, F8, Fz), central (C1, C2, C3, C4, Cz), temporal (T1, T2), parietal (P1, P2, P3, P4, P5, P6, Pz), and occipital (O1, O2, Oz) regions. The reference electrode was located at the tip of the nose during recording and was re-referenced to Cz during analysis.<sup>2</sup> Both vertical and horizontal eye movements were detected from four electrodes located above and below the eyes. Blink and drift artefacts were removed off-line manually. To minimize “noise” generated by ocular and muscular artefacts (Hagemann, 2004), participants were instructed to refrain from



**Figure 1.** (a) Stimuli used during Part 2 of the transitive inference (TI) task (see Procedure). On the top left are stimuli employed during training and testing of stimulus relations (e.g., X is happier than A). On the top right are the shapes trained as contextual cues. The functions of the cues were established through a series of rewarded and non-rewarded conditional discriminations in the presence of alternative emotional faces (see Figure 1b). On the bottom panel is a self-image stimulus before (bottom-left) and after (bottom-right) image processing. Note how all regions of the face were blurred to minimize unintended stimulus control (McIlvane & Dube, 2003). (b) Stimuli used for establishing contextual cue functions. On the left side are some of the exemplars employed during training trials (i.e., with corrective feedback); on the right side are some of the exemplars employed during non-rewarded test trials. Note that the trained and tested discriminations constituted the following: In the presence of Cue 1 (see Figure 1a), the rewarded patterns were happy+/neutral-, neutral+/unhappy-, happy+/unhappy-. In the presence of Cue 2, the rewarded patterns were unhappy+/happy-, neutral+/happy-, unhappy+/neutral-.

blinking/moving their eyes/head during EEG recording phases.

Data were recorded at 256,000 bits per second, with all impedances being kept within a range of 5–9 k $\Omega$ . The voltage differences between the recording

and reference electrodes were extracted and amplified along a bandpass of 0.16–60 Hz with a gain of 1000 and were notch-filtered at 50 Hz. The conversion rate was set at 2000 Hz per channel within a 150-V range. EEG data were analysed offline using BESA

software. Specifically, epochs were defined from 0 ms to 2048 ms post stimulus onset. Fast Fourier transforms (FFT) were conducted over individualized 5-Hz bins within the 8–13-Hz frequency band to calculate power values, which were subsequently normalized (see Klimesch, 1999, for details regarding calculation of individualized alpha bands). Frontal alpha asymmetry (FAA) values were calculated as the normalized difference score between right (R) and left (L) hemispheric sites, utilizing the formula  $(R - L)/(R + L)$  (Allen, Urry, Hitt, & Coan, 2004). According to Klimesch (1999), greater alpha power over right, relative to left, anterior regions indicate greater left frontal activation ( $L > R$ ) whereas greater alpha power over left, relative to right, anterior regions, indicate greater right frontal activation ( $R > L$ ). The  $L > R$  and  $R > L$  activation patterns are indicated by positive and negative FAA values, respectively (Allen, Coan, & Nazarian, 2004; Harmon-Jones, Gable, & Peterson, 2010). The F3, F4, F7, and F8 electrode were selected to calculate FAA, in accordance with previous conventions (Harmon-Jones et al., 2010) and for the sake of consistency (Amd & Roche, 2016).

### Procedure

Upon receiving consent, each participant was photographed, and a head and shoulders image was created. These images were then matched for luminosity, contrast, resolution, and saturation levels in accordance with the six other face-stimuli employed (Figure 1a). This was to minimize unintended stimulus control during training (cf. McIlvane & Dube, 2003). The participant's own name was placed over the bottom half of the modified face stimulus and was included as Stimulus C in subsequent phases. Participants were not provided with any verbal information as to why their pictures were taken beyond the statement "It is a necessary part of the present study. You will be de-briefed completely when we finish".

**Phase 1: Baseline.** *Self-report/frontal alpha asymmetry.* Participants were seated in a Faraday shielded room in front of a 14" Dell computer screen. After the electrodes had been attached, and the impedance levels had been checked to be appropriate (i.e.,  $<10$  k $\Omega$ ), participants were instructed to relax and orient towards a white fixation point on a blank, black screen for 3 minutes, followed by an automated prompt to close their eyes for a further 3 minutes, during which their EEG was recorded. The eyes-open (EO) followed by eyes-closed (EC) procedure enabled identification of

individual alpha bandwidths (IABs). The procedure for calculating IABs was devised by Klimesch (1999) and consists of the following steps: first, calculate the power spectra for the EO and EC conditions separately. Next, define the "transition frequency" (TF) as the point at which "theta" (enhanced spectra during task performance) and "alpha" (suppressed spectra during task performance) intersect. An individual's IAB may be calculated as the range between TF and  $TF + 5$  (Klimesch, 1999, p. 171). Following EO–EC recording, the following instructions appeared on screen:

Welcome. In this part of the experiment, you will first see an image come up on screen. Please pay attention to the image. After a few moments, five boxes will appear near the bottom of the image. Please indicate how "happy" or "unhappy" the image makes you feel by clicking on the appropriate box (1 = happy, 2 = somewhat happy, 3 = neutral, 4 = somewhat unhappy, 5 = unhappy) with the mouse. Please respond using the index finger on your right hand ONLY during all stages of the experiment. IT IS IMPORTANT that you refrain from moving your head and blinking/moving your eyes when the image first appears—you may blink/move your head after the boxes have appeared. Please ask the experimenter if you have any further questions—otherwise press any key to begin. . . .

When a key was pressed, a blank screen with a fixation cross was displayed for 300 ms, followed by a face stimulus (A, B, C, D, or E) presented for 2000 ms in the centre of the screen. After 2000 ms, a mouse cursor and five white boxes with the numbers 1, 2, 3, 4, and 5 printed in black inside each box, respectively, appeared near the bottom half of the screen. Participants were required to move the cursor to one of the boxes and click on it in order to continue (clicking anywhere else on the screen did not have any effect). This was followed by the re-appearance of the fixation cross on a blank screen for 300 ms, after which the next trial commenced. Each of the five face stimuli appeared 15 times, leading to a block total of 75 trials. Completion of 75 trials terminated Phase 1.

**Phase 2: Transitive inference task.** For ease of comprehension, the tasks involved in Phase 2 have been described in two parts.

*Part 1: Establishing contextual cues.* The goal here was to establish two abstract shapes as functionally equivalent to the phrases "is happier than" and "is unhappier than", respectively. This was accomplished by presenting one of three emotional face pairs (i.e.,

happy/neutral, neutral/unhappy, happy/unhappy pairs—see Figure 1b) in the presence of either Cue 1 or Cue 2. If a pair appeared in the presence of Cue 1, selection of the relatively happier face was rewarded—that is, happy+/neutral–, neutral+/unhappy–, or happy+/unhappy–. Alternatively, if the pair appeared in the presence of Cue 2, the selection of the relatively unhappier face was rewarded—that is, neutral+/happy–, unhappy+/neutral–, or unhappy+/happy–. A correct discrimination was followed by the message “Correct” in green font, while an incorrect discrimination was followed by the message “Wrong” in a red font. Participants had an unrestricted time window to respond during any given trial. All participants responded across all trials within 10 seconds of stimulus presentation (see Figure 2a).

The number of training trials were fixed at 48 trials per participant distributed across six trial blocks (see Table 1, rows 1 and 2). Following the completion of 36 training trials across three blocks, test trials began to appear where participants were shown novel emotional face pairs in the presence of either Cue 1 or Cue 2. Participants had to produce the appropriate discriminations (e.g., select happier face when Cue 1 is presented, or select unhappier face when Cue 2 is presented) without any corrective feedback. Each participant had to undergo at least 15 test trials (from Blocks 4 to 6) within which to reach an accuracy criterion of 80%—not meeting this criterion led to a re-exposure to the test trial block (no more training trials were presented). If the 80% criterion was not met within 45 trials, the participant would be thanked and was excused from the study. All participants reached criterion before the end of 45 trials, however.

**Part 2.** Participants were trained in the B+/C– and C+/D– discriminations across 50 trials from Blocks 1–6, followed by the X+/A–, A+/B–, D+/E–, and E+/Y– discriminations across 80 trials from Blocks 2–6 (see Table 1). Note that, similar to Part 1, the number of training trials were fixed for all participants. Unlike Part 1, discriminations were rewarded in the presence of the *happier than* cue only. Corrective feedback for the trained discriminations was gradually reduced across trials, with any trial for which no feedback was provided deemed as a “test trial” (see Table 1 for the reinforcement leaning schedule across blocks). Briefly, participants were exposed to the discriminations reinforced during training (e.g., B/C in the presence of *happier than*) without feedback for at least 78 trials from Blocks 4–6. As before, participants had to respond accurately at least 80% of the

time or be re-exposed to the testing trials for a maximum of three cycles.

From Block 4, the X/A, A/B, B/C, C/D, D/E, and E/Y pairs began to appear in the presence of the *unhappier than* cue in trials without feedback (recall that training with feedback took place in the presence of the *happier than* cue only). These pairs appeared at least 45 times from Blocks 5–6, with an accuracy threshold set at 80%—not meeting criterion led to re-exposure to the test trials for a maximum of three times, as before. These particular trials were deemed tests for functional symmetry. For example, if a participant learned to discriminate A+/B– in the presence of the *happier than* cue, then discriminating B+/A– in the presence of the *unhappier than* cue would indicate functional symmetry (i.e., after learning A is happier than B, responding B is unhappier than A).

Finally, trials testing for functional transitivity began to appear in Block 6, which involved presenting face-stimuli that had not been paired together in previous trials (e.g., A/C, B/E, C/Y) in the presence of either *happier than* or *unhappier than* (i.e., Cue 1 or Cue 2) across a minimum of 75 trials. A successful demonstration of functional transitivity involved responding in accordance with previously established relations—for instance, upon learning A+/B– and B+/C– in the presence of *happier than*, subsequently responding A+/C– in the presence of *happier than*, or C+/A– in the presence of *unhappier than*, would constitute successful, transitive responses given that the relations trained were A is *happier than* B, and B is *happier than* C. An accuracy criterion of 80% was set for these test trials as well, with the maximum number of re-exposures to test trials set at three. Completion of test trials signalled the end of Part 2. Note that all participants underwent a fixed number of training trials, while the proportion of test trial types were contingent on individual participant performance. All participants completed the test trial requirements within three exposures to the test block.

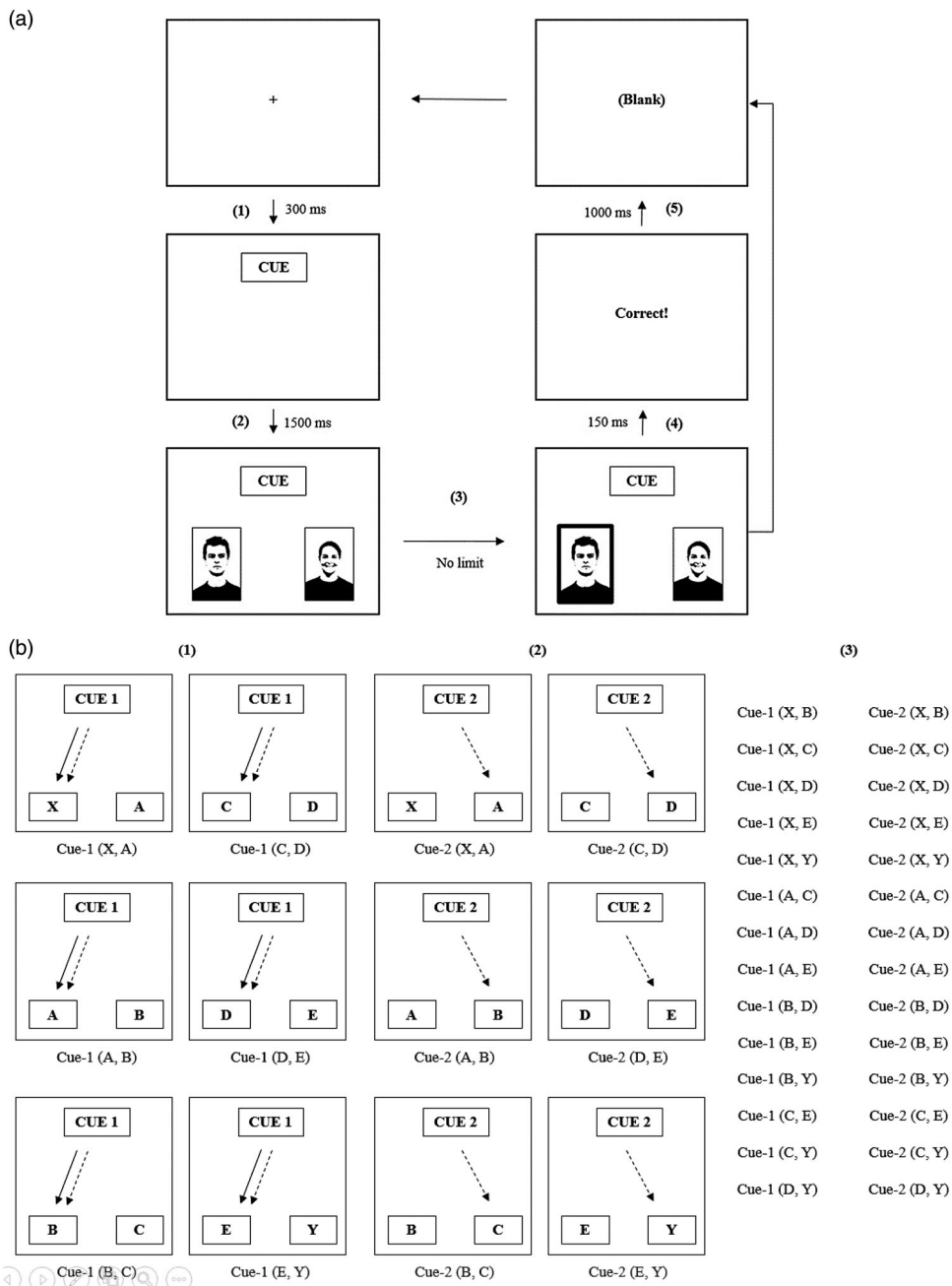
**Phase 3: Post assessment.** All tasks from Phase 1 were re-administered. That is, participants had to rate the A, B, C, D, and E stimuli as their EEG was recorded. Completion of 75 trials signalled the end of the experiment.

## Results

### Self-report

Mean ratings for the A, B, C, D, and E stimuli are presented in Table 2 and Figure 3. The values were





**Figure 2.** (a) (Left side) trial sequence during the transitive inference (TI) task. Each trial commenced with a plain fixation cross on a blank screen for 300 ms, followed by the presentation of a cue near the top half of the screen (1). After 1500 ms, two face stimuli would appear near the bottom half of the trial screen, from which the participant was required to make a selection in order to continue (2). Note that during Part 1 of the TI task (see Procedure), comparison pairs constituted of happy–unhappy, neutral–happy, and neutral–unhappy face pairs; during Part 2 of the TI task, the pairs constituted of pairs comprising the X, A, B, C, D, E, and Y stimuli. When a face was selected, its borders would be highlighted for 150 ms (3). Next, if the presented trial type appeared during training, a feedback message appeared for 1000 ms, followed by an inter-trial interval (ITI) of 200 ms (4). Otherwise, if the trial appeared during testing, the program immediately progressed to a 1200-ms ITI (5). This was followed by the onset of a fixation cross, signalling a new trial sequence. (b) Trial types during relational training and tests for TI. The selections of X + A–, A + B–, C + D–, D + E–, and E + Y– were rewarded (solid arrow) in the presence of Cue 1 only—note that the trained discriminations were presented to participants in trials without corrective feedback (dashed arrow) as well (1). Tests for functional symmetry, where the dashed arrow indicates the correct selection, involved presenting comparison pairs with Cue 2 (2). Tests for functional transitivity involved presenting all combinations of the X, A, B, C, D, E, and Y stimuli with Cue 1 and Cue 2 (3). The faces were taken from the Radboud Face database, which is publicly available research database.

**Table 1.** Trial sequence across transitive inference task.

Trial type	Block sequence					
	Block 1	Block 2	Block 3	Block 4	Block 5	Block 6
Cue train	20	10	6	4	4	4
Cue test	—	—	—	5(4)	5(4)	5(4)
X vs. A	—	—	10	6	2	2
A vs. B	—	—	10	6	2	2
B vs. C	—	10	5	6	2	2
C vs. D	—	10	5	6	2	2
D vs. E	—	—	10	6	2	2
E vs. Y	—	—	10	6	2	2
X vs. A	—	—	—	5(4)	4(3)	4(3)
A vs. B	—	—	—	5(4)	4(3)	4(3)
B vs. C	—	—	—	5(4)	4(3)	4(3)
C vs. D	—	—	—	5(4)	4(3)	4(3)
D vs. E	—	—	—	5(4)	4(3)	4(3)
E vs. Y	—	—	—	5(4)	4(3)	4(3)
A vs. X	—	—	—	—	5(4)	4(3)
B vs. A	—	—	—	—	5(4)	4(3)
C vs. B	—	—	—	—	5(4)	4(3)
D vs. C	—	—	—	—	5(4)	4(3)
E vs. D	—	—	—	—	5(4)	4(3)
X vs. B	—	—	—	—	—	5(4)
X vs. C	—	—	—	—	—	5(4)
X vs. D	—	—	—	—	—	5(4)
X vs. E	—	—	—	—	—	5(4)
X vs. Y	—	—	—	—	—	5(4)
A vs. C	—	—	—	—	—	5(4)
A vs. D	—	—	—	—	—	5(4)
A vs. E	—	—	—	—	—	5(4)
A vs. Y	—	—	—	—	—	5(4)
B vs. D	—	—	—	—	—	5(4)
B vs. E	—	—	—	—	—	5(4)
B vs. Y	—	—	—	—	—	5(4)
C vs. E	—	—	—	—	—	5(4)
C vs. Y	—	—	—	—	—	5(4)
D vs. Y	—	—	—	—	—	5(4)
Min. trials	20	30	56	75	70	135

Note: Trial type: trial types that participants were trained/tested on. For the cue training and testing trial types, participants were shown happy/neutral, neutral/unhappy, and happy/unhappy face pairs in the presence of Cues 1 and 2, where they were reinforced to select the happier face in the presence of Cue 1 and the unhappier face in the presence of Cue 2. The remaining trial types (e.g., X vs. A, B vs. C, etc.) describe the specific comparison pairs that were presented over the course of the task. Trial numbers in italics indicate the fixed number of rewarded, training trials. For test trial numbers, the numbers within parentheses indicate the success criterion for a specific trial type; for instance, "5(4)" indicates that for 5 presentations of a specific trial type, 4 correct discriminations in the absence of feedback was the success criterion. Min. trials: the minimum number of trials per block. Note that, for Blocks 1 to 3, the number of training trials were fixed. For Blocks 4 to 6, the minimum number of trials are presented, although the number could increase since participants could be re-exposed to the non-reinforced test-trial types for a maximum of 3 times if the success criterion was not met.

averaged from the 15 ratings made per stimulus (so 75 ratings per participant) for each condition (baseline vs. post). Recall that three predictions were made during study onset. First, we predicted that the A and B

stimuli would be rated "happier" than at baseline, while the ratings for the D and E stimuli would be "unhappier" than at baseline. We assessed this with four repeated measure *t* tests using a Bonferroni alpha of .0025 (.01/4) to control for familywise error rate. The  $\alpha$  criterion was set at .01 per  $H_0$  (e.g., A-post > A-base, D-post < D-base) instead of the conventional .05 given that several null hypotheses were being tested (Cohen, 1992, p. 156). In accordance with predictions, the ratings for Stimulus A,  $t(299) = 23.1$ ,  $p < .0001$ ,  $d = 1.89$ , and Stimulus B,  $t = 8.69$ ,  $p < .0001$ ,  $d = 0.71$ , were significantly higher than those recorded at baseline. Similarly, the ratings for Stimulus D,  $t = 11.62$ ,  $p < .0001$ ,  $d = 0.95$ , and Stimulus E,  $t = 26.3$ ,  $p < .0001$ ,  $d = 2.15$ , were significantly lower than those recorded at baseline. A similar contrast for Stimulus C yielded no significant differences between baseline and post ratings,  $t = 1.61$ ,  $p > .1$ ,  $d = 0.13$ . The effect sizes (Dunlap, Cortina, Vaslow, & Burke, 1996, p. 171) for the A/E and B/D pre-post contrasts are considered very large and large, respectively, for paired *t* tests (Cohen, 1988).

A second prediction was that the ratings provided for the A, B, D, and E stimuli would demonstrate a trend corresponding to the A > B > D > E series after the series was established. To assess for a linear trend across the means, we performed univariate analyses for baseline and post conditions, respectively, for the A, B, D, and E stimuli with coefficients set to +2, +1, -1, and -2, respectively (given the predicted A > B > D > E gradient). During baseline, the averaged ratings for the A ( $M = 3.373$ ,  $SD = 0.684$ ), B ( $M = 2.906$ ,  $SD = 0.658$ ), D ( $M = 3.01$ ,  $SD = 0.631$ ), and E ( $M = 3.21$ ,  $SD = 0.606$ ) stimuli did not show a statistically significant trend with equal variances assumed,  $F(3, 1196) = 3.591$ ,  $p = .058$ ,  $\eta^2 = .003$ . Levene's test indicated that the assumption of homogeneity was violated, however, which required the degrees of freedom to be adjusted from 1196 to 866,  $F = 3.893$ ,  $p = .059$ . During post, the averaged ratings for the A ( $M = 4.46$ ,  $SD = 0.613$ ), B ( $M = 3.387$ ,  $SD = 0.757$ ), D ( $M = 2.44$ ,  $SD = 0.51$ ), and E ( $M = 1.813$ ,  $SD = 0.707$ ) stimuli showed a highly significant trend with equal variances assumed,  $F(3, 1196) = 2733.301$ ,  $p < .0001$ ,  $\eta^2 = .701$ . To adjust for the violation of the homogeneity assumption, the degrees of freedom were adjusted from 1196 to 845,  $F = 2692.261$ ,  $p < .0001$ . The trend observed in the post condition conforms with the predicted A > B > D > E gradient (Table 2).

A visual analysis of the averaged ratings for the A, B, C, D, and E stimuli clearly indicates that the C stimulus

**Table 2.** Repeated measure *t* tests contrasting baseline versus post happiness ratings for Stimuli A, B, C, D, and E.

Stimuli	Baseline		Post		<i>t</i> (299)	<i>p</i> (1-tail) <sup>a</sup>	95% CI	Cohen's <i>d</i> <sup>b</sup>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>				
A	3.373	0.684	4.461	0.612	23.144	<.0001	[-1.18, -0.99]	1.89
B	2.907	0.657	3.387	0.755	8.703	<.0001	[-0.59, -0.37]	0.71
C	4.203	0.801	4.297	0.896	1.610	.1084	[-0.29, 0.02]	0.13
D	3.010	0.630	2.442	0.510	11.575	<.0001	[0.47, 0.67]	0.95
E	3.210	0.605	1.813	0.706	26.257	<.0001	[1.29, 1.50]	2.15

Note: CI = confidence interval.

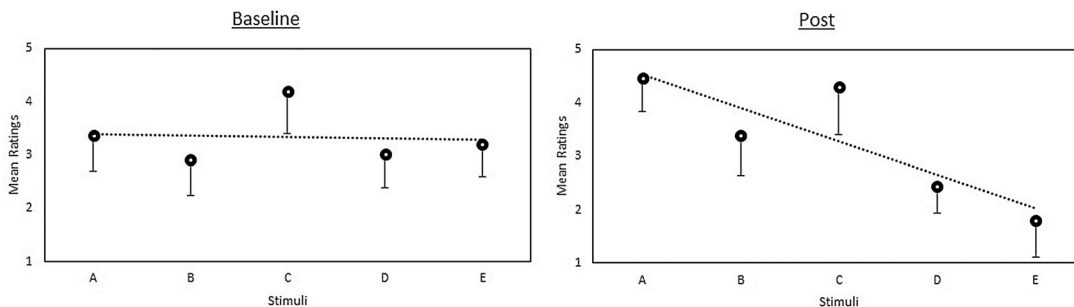
<sup>a</sup>The Bonferroni adjusted alpha value was .0025. A one-directional contrast was conducted as the ratings were predicted to go in specific directions for each stimulus. <sup>b</sup>Indicates magnitude of the effect.

was rated as happier than the B and D stimuli across both the baseline and post conditions (Figure 3). To determine whether the ratings for the C stimulus functioned as an outlier in the present study, we conducted univariate analyses for the A, B, D, E, and C stimuli with coefficients set to -1, -1, -1, -1, and 4, respectively, for the baseline and post conditions. The ratings for Stimulus C were significantly different from the averaged ratings for the A, B, D, and E stimuli during baseline,  $F(1, 1495) = 603.861$ ,  $p < .0001$ ,  $\eta^2 = .288$ , as well as during post,  $F(1, 1495) = 771.857$ ,  $p < .0001$ ,  $\eta^2 = .340$ . The results suggest that relational categorization did not suffice to alter the valence of the already salient C stimulus.

### Frontal alpha asymmetry

As the predictions for the FAA scores were similar to the self-report ratings, a similar battery of statistical tests were conducted. First, the post and baseline FAAs of stimuli A/B and D/E using four repeated measure *t*-tests with an adjusted Bonferroni alpha of .0025 were obtained. The FAAs for Stimulus A in the

post condition was significant compared to baseline,  $t(19) = 3.030$ ,  $p < .004$ ,  $d = 0.89$ . The pre-post FAA contrasts for the remaining stimuli B,  $t = 1.44$ ,  $p > .08$ ,  $d = 0.39$ , C,  $t = 0.34$ ,  $p > .16$ ,  $d = 0.05$ , D,  $t = 0.32$ ,  $p > .37$ ,  $d = 0.08$ , and E,  $t = 1.04$ ,  $p > .15$ ,  $d = 0.23$ , were not significantly different between the two conditions (see Table 3). Across the five contrasts, only the pre-post A contrast produced a large effect ( $d = 0.78$ ). Using the reported effect sizes, powers for the pre-post FAAs for the A, B, C, D, and E stimuli were calculated to be 96%, 63%, 12%, 11%, and 41%, respectively, with  $\alpha$  set at .01 and a sample size of 20 (using GPower\*3; Faul, Erdfelder, Lang, & Buchner, 2007). This suggests the inconclusive FAA results were due to a lack of power. A sensitivity analysis (Faul et al., 2007, p. 177) revealed that the minimum effect size for a sufficiently sensitive test would be 0.76, which only the pre-post FAA contrast for Stimulus A yielded. A prospective power analysis with  $\alpha = .01$ ,  $d = 0.8$ , and power at 80%, showed that a sample size of 19 should have sufficed to detect large effects ( $d = 0.8$ ). Note that, while this may have been the case



**Figure 3.** Baseline (left) and post (right) mean ratings of "happiness" along a 5-point Likert scale (along the y-axis), where the ratings range from "1 = very unhappy" to "5 = very happy". Stimuli A, B, C, D, and E are presented along the x-axis. Note that for Stimuli A and B, post ratings were significantly higher ( $p < .0001$ ) than those observed at baseline. Alternatively, for Stimuli D and E, post ratings were significantly lower ( $p < .0001$ ) than those at baseline. Ratings for the C (self) stimulus did not significantly vary between baseline and post ratings. Error bars indicate standard deviations and are one-sided in order to conserve space. Note that the trend for baseline was non-significant,  $F = 3.893$ ,  $p = .059$ , whereas the trend observed at post corresponded significantly,  $F = 2692.261$ ,  $p < .0001$ , with the predicted  $A > B > D > E$  gradient.

**Table 3.** Repeated measure *t* tests contrasting baseline versus post FAAs for Stimuli A, B, C, D, and E.

Stimuli	Baseline		Post		<i>t</i> (19)	<i>p</i> (1-tail)	95% CI	Cohen's <i>d</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>				
A	0.023	0.143	0.159	0.162	3.030	<.004	[-0.23, -0.042]	-0.89
B	0.015	0.176	0.089	0.204	1.444	>.08	[-0.59, -0.37]	-0.39
C	0.226	0.237	0.238	0.226	0.342	>.36	[-0.29, 0.02]	-0.05
D	0.031	0.202	0.016	0.171	-0.318	>.37	[0.47, 0.67]	0.08
E	0.070	0.185	0.017	0.264	-1.040	>.15	[1.29, 1.50]	0.23

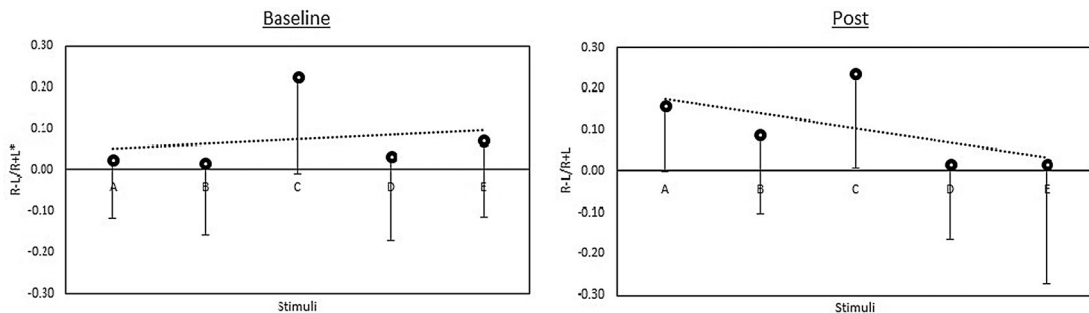
Note: FAA = frontal alpha asymmetry. CI = confidence interval.

for the self-report data provided above, it may not be applicable to the FAA data (cf. Hagemann, 2004). That is, given the low signal-to-noise ratio typical of electrophysiological measures (Teplan, 2002), at least when compared to a self-report rating scale, it is likely that a larger sample would be required for significant FAA contrasts to be detected. This is further corroborated by the design of the study, which would predict smaller effect sizes for stimuli that were closer to the intermediate term in the series (i.e., Stimuli B and D) than for those that were closer to the end anchors (Stimuli A and E). A lack of significant differences across the board does not necessarily indicate that the FAA metric was insensitive to the manipulation, however, as a visual inspection of the trends across members of the series in Figure 4 reveals. Specifically, the trends in FAA resemble, at least superficially, those that were observed for the self-reported ratings (Figure 3), indicating some sensitivity to the emergent valence gradient.

To assess whether the FAA shifts were significant, we assessed for linear trends among the FAAs for baseline and post conditions for the A, B, D, and E stimuli with coefficients set to +2, +1, -1, and -2,

respectively, as before. During baseline, the FAAs for the A ( $M = 0.023$ ,  $SD = 0.143$ ), B ( $M = 0.015$ ,  $SD = 0.175$ ), D ( $M = 0.031$ ,  $SD = 0.208$ ), and E ( $M = 0.07$ ,  $SD = 0.19$ ) stimuli did not show a statistically significant trend,  $F(3, 76) = 0.739$ ,  $p = .393$ ,  $\eta^2 = .010$ . During post, the FAAs for the A ( $M = 0.159$ ,  $SD = 0.162$ ), B ( $M = 0.089$ ,  $SD = 0.204$ ), D ( $M = 0.177$ ,  $SD = 0.040$ ), and E ( $M = 0.017$ ,  $SD = 0.212$ ) stimuli demonstrated the predicted trend,  $F(3, 76) = 5.927$ ,  $p = .017$ ,  $\eta^2 = .072$ . As with the ratings data, the post FAAs showed a statistically significant linear trend in accordance with a  $A > B > D > E$  structure (Figure 4).

Finally, we employed a univariate analysis to contrast the FAAs of the A, B, D, E, and C stimuli utilizing coefficients of -1, -1, -1, -1, and 4, respectively, for the baseline and post conditions. Similar to the outcome reported for the ratings data, the FAA for Stimulus C was significantly greater than the FAAs for Stimuli A, B, D, and E during baseline,  $F(1, 95) = 15.371$ ,  $p < .0001$ ,  $\eta^2 = .139$ , and during post,  $F(1, 95) = 9.902$ ,  $p = .002$ ,  $\eta^2 = .094$ . Similar to the results for the ratings data, the current analysis indicates that relational categorization did not have a significant effect on the FAAs for Stimulus C.



**Figure 4.** Frontal alpha asymmetries (FAAs) acquired during baseline (left panel) and post (right panel) conditions. R = right; L = left. Positive values indicate greater activation over the left prefrontal cortex (PFC) than over the right PFC; negative values indicate greater activation over the right PFC than over the left PFC. The baseline versus post change in elicited FAAs observed for Stimulus A was significantly different ( $p < .007$ ). The remaining contrasts were non-significant. Similar to the ratings data, a significant,  $F(3, 76) = 5.927$ ,  $p = .017$ ,  $\eta^2 = .072$ , linear trend was observed during the post condition but not in the baseline condition,  $F(3, 76) = 0.739$ ,  $p = .393$ ,  $\eta^2 = .010$ , indicating the presence of a gradient corresponding to  $A > B > D > E$ . \*Calculation of the normalized FAA metric (cf. Allen, Urry, et al., 2004).

It is worth noting the obvious correspondence between the ratings and FAAs across both conditions for Stimulus C (Figures 3 and 4). Whether this is exclusively related to stimulus valence, familiarity, and/or motivation (cf. Berkman & Lieberman, 2010; Harmon-Jones, 2003) remains unknown, though it is not unwarranted to assume that both measures reflect stimulus valence to some degree. In any case, holding the effect for Stimulus C constant, we can conclude that the valences of the A, B, D, and E stimuli transformed in accordance with the structure of the relational series.

### Discussion

The goal of the current study was to assess whether the valences of blurred face stimuli could be transformed by virtue of membership in a seven-element series, and whether such a transformation could affect a stimulus that was emotionally salient a priori.

Performance across the self-report and FAA measures indicate that the functions of all stimuli, other than Stimulus C, had transformed in accordance with the relational structure of the series established. That is, the valences of the unfamiliar blurred faces fell along a gradient predicted by a  $X >_{\text{happy}} A >_{\text{happy}} B >_{\text{happy}} C >_{\text{happy}} D >_{\text{happy}} E >_{\text{happy}} Y$  series, sans Stimulus C. This was most apparent when viewing the trends for the self-reported ratings (Figure 3) and the FAAs (Figure 4) during the post condition. Critically, the A and B stimuli elicited higher happiness ratings and greater right prefrontal alpha activation than the D and E stimuli. Given that neither the A nor the E stimuli constituted “end anchors” of the series, the pre versus post differences are unlikely to have been driven by associative value transfer processes (e.g., Wynne, 1995). The high happiness ratings and FAAs reported for Stimulus C during baseline and post measurements suggest a relation between FAA and valence, motivation and/or familiarity (cf. Coan & Allen, 2003). The FAA shifts recorded for Stimuli B, D, and E, while individually non-significant (Table 2), nevertheless showed the predicted trend (Figure 4). The lack of significant FAA differences across the individual stimulus contrasts may be due to a lack in the sensitivity of the FAA measure—namely, transformed/derived valence effects (cf. Harmon-Jones et al., 2010), or it may be that FAAs were confounded by value transfer from the highly salient stimulus C (e.g., Greene, Spellman, Levy, Dusek, & Eichenbaum, 2001), although both interpretations are speculative

(more on FAA limitations in a moment). This suggests that, while useful, FAA measures may be best supplemented by additional dependent measures, such as self-reports or implicit measures (e.g., Amd & Barnes-Holmes, 2014; Amd & Roche, 2015, 2016). All in all, the performances observed here suggest that stimulus valence is at least partially a function of the relational structure of which it is an element, supporting earlier findings (e.g., Amd & Roche, 2015, 2016).

Alternatively, if a stimulus is already highly salient and/or familiar, relational categorization may have little effect on its baseline valence level, demonstrating a boundary condition for the TOF effect, at least using current procedures. The results indicate that under certain conditions (e.g., when the member of a structure is highly valenced a priori, such as the self-images used here), the derivation of the predicted relations may not suffice for TOF to be observed. For instance, note how Stimulus C was rated as “happier” than Stimulus B even though participants learned the  $B >_{\text{happy}} C$  relation during TI training and testing. One reason for this difference may be that, while the TI procedure involved training a  $B >_{\text{happy}} C$  relation directly, Stimulus C was already “different” from other face stimuli along a salience/valence dimension. It is also possible that Stimulus C may have participated in a “same-as” and/or hierarchical relation with the participant (as it was his/her own face). If so, the hierarchical relational structure may have interfered with TOF in the current series (cf. Slatery & Stewart, 2014). In order to determine whether this was the case, future researchers can attempt to strengthen stimulus relations to supersede pre-existing ones (e.g., through overtraining—cf. Bortoloti, Rodrigues, Cortez, Pimentel, & de Rose, 2013) and to note whether such interventions can transform the functions of a highly salient (e.g., self-image) stimulus.

In any case, our findings indicate that while TOF can be readily observed in experimental preparations employing relatively “neutral” elements (e.g., Amd & Roche, 2016), the valences of non-neutral elements may not be as susceptible to TOF. This suggests that emotionally salient elements can maintain their valence even while constituting a member of a relational structure/series that provides countering information (i.e., rating  $C >_{\text{happy}} B$  after learning  $B >_{\text{happy}} C$ ). In turn, these findings explain how instructions to suppress dissonant emotions may not always succeed (cf. Spokas, Luterek, & Heimberg, 2009). Future researchers could explore whether the valence shifts (in ratings and FAAs) can be affected

by the presentation of one's name with *any* blurred face stimulus. This would separate the effects of stimulus salience and relational series position on the valence of purely verbal and/or purely topographical (e.g., facial), personally salient, stimuli.

The current findings indicate that humans appraise their own faces as more positively valenced than non-familiar faces. While the non-significant FAA shifts observed in the presence of the B, D, and E stimuli challenge the utility of FAA as a reliable metric of emotional valence, it is worth bearing in mind that the evoked FAAs may have been confounded by factors not controlled for (e.g., menstruation cycles, sleeping habits; see Bazanova & Vernon, 2014). Nevertheless, the explicitly reported ratings suggest that it is at least plausible that the FAAs represent stimulus valence to some degree (e.g., Harmon-Jones et al., 2010; Schöne, Schomberg, Gruber, & Quirin, 2015), although further research is required to explicate the ways in which these two measures are functionally related. As noted previously, while some researchers have found FAA to be sensitive to valences established via conditioning (e.g., Cunningham, Espinet, DeYoung, & Zelazo, 2005), others have contested this claim and suggested FAA as more representative of approach tendencies than stimulus valence per se (Harmon-Jones et al., 2010). While the details of this debate are beyond the scope of the present article, the key point is that FAA may be sensitive to valences that are inferred/derived as a consequence of their position within a relational series (e.g., Amd & Roche, 2015, 2016) as opposed to purely on the basis of direct pairing (i.e., classical conditioning; see Cunningham et al., 2005). Maintaining this distinction between "conditioned" and "derived" valence change effects may be important in explaining the establishment of stimulus valences (Amd & Roche, 2015, p. 538). On balance, while identifying the neural substrates of these differing processes would be welcome, it is not necessary (cf. Miller, 2010; Miller, Crocker, Spielberg, Infantolino, & Heller, 2015; Staddon, 2014) for the purposes of experimental control.

Some potential limitations of the present study are worth noting. First, some may criticize our use of a normalized difference score [i.e.,  $(R - L)/(R + L)$ ] to compute FAA differences rather than calculating the difference between naturally log-transformed values (i.e.,  $\ln R - \ln L$ ), given the issues of reliability with the former (Allen, Coan, et al., 2004). On balance, Overall and Woodward (1975) demonstrated that reduced reliability for difference scores can yield

higher power for detecting significant differences. Given that the present study, along with Amd and Roche (2016), constitutes among the first known investigations of stimulus-elicited FAAs in relation to *derived* emotional effects, we believe that a powerful measure for detecting statistical differences is presently more valuable than prematurely relating the effects to the self-report measures. Additionally, the normalized difference score has shown a correlation exceeding .99 in relation to the natural log asymmetry measure (Allen, Urry, et al., 2004), indicating both to be robust measures of the asymmetry metric.

A second matter is the lack of topographical maps of pre and post FAA activation, which is related to our use of a normalized score. The reason for excluding them is that it is difficult to ascribe the observation of a positive FAA value to increased alpha activity over the right LPFC or decreased alpha activity over the left LPFC when using such scores (Bazanova & Vernon, 2014; Hagemann, 2004). Future researchers can employ more rigorous techniques, such as residualizing power at a specific electrode site with a hierarchical regression, and then correlating the residualized values with the variables of interest, in order to determine regions of interest (cf. Wheeler, Davidson, & Tomarken, 1993). Given the apparent incommensurability of a "standard" alpha generator across individuals (e.g., Bazanova & Vernon, 2014), it is more constructive at present to focus on the functional aspects underlying alpha asymmetry effects over anterior regions (cf. Klimesch, 2012).

A third matter involves the ecological validity of extrapolating from experimental findings with artificially blurred human faces, which are not typical to natural circumstances. As noted previously, the faces here were blurred in order to prevent participants from inferring emotional states that could have otherwise counteracted the information provided through series membership. Consequently, the present preparation allowed determining the influence of relational structure on stimulus valence gradients while minimizing counter-control by stimulus features *prima facie* (e.g., McIlvane & Dube, 2003). Subsequent preparations could employ natural faces and/or emotional contexts other than happiness, given that the latter can be inferred even when a face has been normatively rated as emotionally neutral.

Fourth, one could question *whose* happiness a participant was responding to as they provided their image ratings. That is, given the instructions "Please indicate how 'happy' or 'unhappy' the image makes

you feel”, participants may have responded to each stimulus in relation to the trained and derived relations within the series (e.g., A is happier than B) or to each stimulus in relation to themselves (e.g., A is happier relative to *me*). Given the TOFs observed for the A, B, D, and E stimuli, one could argue that the participants were providing ratings based on the relations trained. On balance, the consistently high ratings for Stimulus C, even after learning relations such as  $B >_{\text{happy}} C$ , suggest that participants were responding to the C stimulus in relation to themselves. This raises the question of whether rating Stimulus A as “happier than” C, for instance, means that the participant is happier in the presence of A, or whether she is simply describing the relationally relative valence of A. While the inclusion of additional dependent measures might clarify this issue somewhat (e.g., Amd & Roche, 2016; Dougher et al., 2007), future researchers can delve into state differences to see whether they would affect TOF for a priori salient stimuli. For example, one could measure the extent to which current mood indicators effect the valences of a self-stimulus, and whether this could influence TOF in a systematic way.

Finally, one may well ask what “happy” meant in the context of the present experiment. The reader should note that “happiness” was not employed here as a technical term but simply to maintain the reader’s orientation towards the phenomenon of interest (TOF). The emotion-eliciting functions of any stimulus, including “happiness” functions, can be analysed in terms of various hypothetical sub-components (Gerber et al., 2008), the most common being the bifurcated dimensions of valence and arousal. Briefly, valence indicates the degree to which a stimulus is appetitive or aversive for an organism, and arousal refers to the elicited degree of physiological excitability (Kensinger, 2004). Although it is likely that “happiness” refers to *some* combination of valence and arousal, the reader should note that the “happier than” and “unhappier than” cues used presently can be read as “more positively valenced” and “more negatively valenced”, respectively. Such cues function as “imperative environmental facts (determining) the direction of the behavior” (Lewin, 1935, p. 77), where the “happiest” and “unhappiest” members of the series can be conceptualized as opposite ends along a valence continuum (Tolman, 1932, p. 81). In such cases, the size of a relational series should not affect the primary prediction of the direction of valence gradients across series (Amd &

Roche, 2015). Nevertheless, future researchers wishing to assess whether such a continuum is applicable along other emotional dimensions could replicate the present work using cues specifying alternate relational emotions (e.g., angrier than, calmer than) to see whether similar TOF gradients emerge.

Important theoretical implications arising from the present findings relate to the observation that responses towards an emotionally masked face may at least partly be a function of the relational structure(s) of which it is a member (Amd & Roche, 2015). First, given the clear TOF effects observed for the A and E stimuli, it appears that constituting end-terms in a relational series was not a significant confound in TOF yields here or elsewhere (e.g., Amd & Roche, 2015). Secondly, TOF can take place when the stimuli related within a valence-occasioning context are emotionally homogeneous in relation to each other. If a highly salient/familiar stimulus is introduced, however, relational series membership may have little influence on the emotional properties of the salient element *prima facie*. Our results indicate that emotional reactions to faces need neither be directly learned nor instructed as a rule, but can be derived through the logical relations they share with other symbolic stimuli (Dymond & Roche, 2013; Dymond et al., 2014; Holt, 1914). Thus, TOF is more readily observed when the related stimuli are not salient, motivating and/or familiar a priori. Simply put, the relations between stimuli determine the direction of the valence gradient across them only if those stimuli were emotionally homogeneous and relatively neutral to begin with. This highlights a boundary condition of the TOF process for the first time, which has been garnering increasing attention from symbolic communication researchers (Dymond & Roche, 2013).

From a broader perspective, the current findings highlight how elements embedded within the same knowledge structure may give rise to feelings of “dissonance” (Jansz & Timmers, 2002), which in turn can produce feelings of unease. Although the present study did not assess “unease” per se, the pre versus post valence ratings indicate how such unease could arise if the elements being related are disparate enough. For instance, if a participant is presented with a stimulus that is already positively valenced towards her, and is then informed that some other, lesser valenced stimulus should be responded to more positively than the original

stimulus (e.g., B ><sub>happy</sub> C in the present study), it is easy to see how she might resist responding to the first stimulus as more negative than the latter (cf. McConnell & Leibold, 2001). Given enough of a disparity between the two stimuli (e.g., requiring a Christian to respond to the Islamic crescent as a positively valenced symbol; see Kraemer, 1997), unease is likely to emerge. Taken together, this suggests that while relational knowledge structures play a key role in higher cognition (cf. Halford, Wilson, & Phillips, 2010), such knowledge is less likely to have an influence on emotionally salient elements embedded within those structures, at least within the period of a single experimental session. Future research could attempt to expand upon the relational series established here (e.g., by using different exemplar types, real-world cues, combinations of relation-types) in order to assess the type of structure necessary to alter the valences of salient elements, short of direct pairings with biologically relevant stimuli (cf. Amd et al., 2013; Dougher et al., 2007).

In conclusion, the present study demonstrates a transformation of valence functions for masked face stimuli following membership within the same relational series. In doing so, a possible boundary condition for the TOF effect was identified. Specifically, newly established relational structures may fail to alter the valence of self-image stimuli across a single session, presumably based on their a priori salience. On a general level, these findings highlight how dissonance can persist after emotionally salient elements become incorporated into the same relational knowledge structure.

## Notes

1. Abstract shapes were used in lieu of naturally occurring phrases (such as “is happier than” and “is unhappier than”) given that the latter are used in everyday discourse and may not have been situationally equivalent across all the participants involved.
2. Note that while the validity of a Cz reference electrode for identifying the source of alpha activity has been contested (Hagemann, 2004), it nevertheless remains the most commonly used reference in FAA research (Amd & Roche, 2016; Coan & Allen, 2003; Huang et al., 2015) validating its employment for the present inquiry. For further details, see Footnote 2 in Amd and Roche (2016).

## Disclosure statement

No potential conflict of interest was reported by the authors.

## Funding

This work was supported by Fundação de Amparo à Pesquisa do Estado de São Paulo [grant number 2015/24159-4]; Irish Research Council for Science, Engineering and Technology.

## ORCID

Micah Amd  <http://orcid.org/0000-0003-4225-7026>

## References

- Allen, J. J. B., Coan, J. A., & Nazarian, M. (2004). Issues and assumptions on the road from raw signals to metrics of frontal EEG asymmetry in emotion. *Biological Psychology*, 67(1–2), 183–218. doi:10.1016/j.biopsycho.2004.03.007
- Allen, J. J. B., Urry, H. L., Hitt, S. K., & Coan, J. A. (2004). The stability of resting frontal electroencephalographic asymmetry in depression. *Psychophysiology*, 41(2), 269–280. doi:10.1111/j.1469-8986.2003.00149.x
- Amd, M. (2014, December). Relative shifts in frontal alpha asymmetry and the transformation of evaluative functions. *Canadian Journal of Experimental Psychology*, 68(4), 255–255.
- Amd, M. (2016). *A derived transformation of valence functions across complex relational networks*. PhD thesis. Maynooth University, Maynooth.
- Amd, M., & Barnes-Holmes, D. (2014). A derived transformation of evaluative preferences using implicit association tests. *The Psychological Record*, 64, 475–485. Retrieved from [http://download.springer.com/static/pdf/321/art253A10.1007252Fs40732-014-0033-z.pdf?auth66=1421915528\\_f6f535efbbd514b1cb38c48868ef8123&ext=.pdf](http://download.springer.com/static/pdf/321/art253A10.1007252Fs40732-014-0033-z.pdf?auth66=1421915528_f6f535efbbd514b1cb38c48868ef8123&ext=.pdf)
- Amd, M., Barnes-Holmes, D., & Ivanoff, J. (2013). A derived transfer of eliciting emotional functions using differences among electroencephalograms as a dependent measure. *Journal of the Experimental Analysis of Behavior*, 99(3), 318–334. doi:10.1002/jeab.19
- Amd, M., & Roche, B. (2015). A derived transformation of valence functions across two 8-member comparative relational networks. *The Psychological Record*, 65(3), 523–540.
- Amd, M., & Roche, B. (2016). A derived transformation of emotional functions using self-reports, implicit association tests, and frontal alpha asymmetries. *Learning & Behavior*, 44, 175–190.
- Barnes, D. (1994). Stimulus equivalence and relational frame theory. *The Psychological Record*, 44(1), 91–124.
- Bassili, J. N. (1979). Emotion recognition: The role of facial movement and the relative importance of upper and lower areas of the face. *Journal of Personality and Social Psychology*, 37, 2049–2058. doi:10.1037/0022-3514.37.11.2049
- Bazanov, O. M., & Vernon, D. (2014). Interpreting EEG alpha activity. *Neuroscience and Biobehavioral Reviews*, 44, 1–17. doi:10.1016/j.neubiorev.2013.05.007
- Berkman, E. T., & Lieberman, M. D. (2010). Approaching the bad and avoiding the good: Lateral prefrontal cortical asymmetry distinguishes between action and valence. *Journal of Cognitive Neuroscience*, 22(9), 1970–1979.
- Bialer, I. (1961). Primary and secondary stimulus generalization as related to intelligence level. *Journal of Experimental Psychology*, 62(4), 395–402.



- Bonatti, L., Frot, E., Zangl, R., & Mehler, J. (2002). The human first hypothesis: Identification of conspecifics and individuation of objects in the young infant. *Cognitive Psychology*, 44(4), 388–426.
- Bortoloti, R., Rodrigues, N. C., Cortez, M. D., Pimentel, N., & de Rose, J. C. (2013). Overtraining increases the strength of equivalence relations. *Psychology & Neuroscience*, 6(3), 357–364.
- Boyle, S., Roche, B., Dymond, S., & Hermans, D. (2016). Generalisation of fear and avoidance along a semantic continuum. *Cognition and Emotion*, 30(2), 340–352.
- Briesemeister, B., Tamm, S., Heine, A., & Jacobs, A. (2013). Approach the good, withdraw from the Bad—A review on frontal alpha asymmetry measures in applied psychological research. *Psychology*, 4(3), 261–267. Retrieved from <http://file.scirp.org/Html/29227.html>
- Coan, J. A., & Allen, J. J. (2003). Frontal EEG asymmetry and the behavioral activation and inhibition systems. *Psychophysiology*, 40(1), 106–114.
- Coan, J. A., & Allen, J. J. B. (2004). Frontal EEG asymmetry as a moderator and mediator of emotion. *Biological Psychology*, 67(1–2), 7–49. doi:10.1016/j.biopsycho.2004.03.002
- Cofer, C. N., & Foley Jr, J. P. (1942). Mediated generalization and the interpretation of verbal behavior: I. Prolegomena. *Psychological Review*, 49(6), 513–540.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155.
- Cunningham, W. A., Espinet, S. D., DeYoung, C. G., & Zelazo, P. D. (2005). Attitudes to the right-and left: Frontal ERP asymmetries associated with stimulus valence and processing goals. *NeuroImage*, 28(4), 827–834.
- Davidson, R. J., & Fox, N. A. (1989). Frontal brain asymmetry predicts infants' response to maternal separation. *Journal of Abnormal Psychology*, 98(2), 127–131.
- Dimberg, U., Thunberg, M., & Elmehed, K. (2000). Unconscious facial reactions to emotional facial expressions. *Psychological Science*, 11(1), 86–89.
- Dougher, M. J., Hamilton, D. A., Fink, B. C., & Harrington, J. (2007). Transformation of the discriminative and eliciting functions of generalized relational stimuli. *Journal of the Experimental Analysis of Behavior*, 88(2), 179–197. doi:10.1901/jeab.2007.88-179
- Dunlap, W. P., Cortina, J. M., Vaslow, J. B., & Burke, M. J. (1996). Meta-analysis of experiments with matched groups or repeated measures designs. *Psychological Methods*, 1(2), 170–177.
- Dymond, S., & Rehfeldt, R. A. (2000). Understanding complex behavior: The transformation of stimulus functions. *The Behavior Analyst*, 23(2), 239–254.
- Dymond, S., & Roche, B. (2013). In S. Dymond & B. Roche (Eds.), *Advances in relational frame theory*. Oakland, CA: New Harbinger. Retrieved from <https://www.newharbinger.com/advances-relational-frame-theory>
- Dymond, S., Roche, B., Forsyth, J. P., Whelan, R., & Rhoden, J. (2007). Transformation of avoidance response functions in accordance with same and opposite relational frames. *Journal of the Experimental Analysis of Behavior*, 88(2), 249–262.
- Dymond, S., Schlund, M. W., Roche, B., & Whelan, R. (2014). The spread of fear: Symbolic generalization mediates graded threat-avoidance in specific phobia. *Quarterly Journal of Experimental Psychology* (2006), 67(2), 247–259. doi:10.1080/17470218.2013.800124
- Ekman, P., Davidson, R. J., & Friesen, W. V. (1990). The Duchenne smile: Emotional expression and brain physiology. II. *Journal of Personality and Social Psychology*, 58(2), 342–353. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/2319446>
- Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology*, 17(2), 124–129.
- Esteves, F., & Öhman, A. (1993). Masking the face: Recognition of emotional facial expressions as a function of the parameters of backward masking. *Scandinavian Journal of Psychology*, 34(1), 1–18.
- Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G\* power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191.
- Feather, B. W. (1965). Semantic generalization of classically conditioned responses: A review. *Psychological Bulletin*, 63(6), 425–441.
- Fields, L., Adams, B. J., Verhave, T., & Newman, S. (1990). The effects of nodality on the formation of equivalence classes. *Journal of the Experimental Analysis of Behavior*, 53(3), 345–358. doi:10.1901/jeab.1990.53-345
- Fields, L., Arntzen, E., Nartey, R. K., & Eilifsen, C. (2012). Effects of a meaningful, a discriminative, and a meaningless stimulus on equivalence class formation. *Journal of the Experimental Analysis of Behavior*, 97(2), 163–181.
- Gable, P., & Harmon-Jones, E. (2008). Relative left frontal activation to appetitive stimuli: Considering the role of individual differences. *Psychophysiology*, 45(2), 275–278. doi:10.1111/j.1469-8986.2007.00627.x
- Gerber, A. J., Posner, J., Gorman, D., Colibazzi, T., Yu, S., Wang, Z., ... Peterson, B. S. (2008). An affective circumplex model of neural systems subserving valence, arousal, and cognitive overlay during the appraisal of emotional faces. *Neuropsychologia*, 46(8), 2129–2139. doi:10.1016/j.neuropsychologia.2008.02.032
- Greenberg, J., & Alge, B. J. (1998). Aggressive reactions to workplace injustice.
- Greene, J., Spellman, B. A., Levy, W. B., Dusek, J. A., & Eichenbaum, H. B. (2001). Relational learning with and without awareness: Transitive inference using nonverbal stimuli in humans. *Memory & Cognition*, 29(6), 893–902.
- Hagemann, D. (2004). Individual differences in anterior EEG asymmetry: Methodological problems and solutions. *Biological Psychology*, 67(1–2), 157–182. doi:10.1016/j.biopsycho.2004.03.006
- Hagemann, D., Naumann, E., Becker, G., Maier, S., & Bartussek, D. (1998). Frontal brain asymmetry and affective style: A conceptual replication. *Psychophysiology*, 35(4), 372–388. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9643052>
- Halford, G. S., Wilson, W. H., & Phillips, S. (2010). Relational knowledge: The foundation of higher cognition. *Trends in Cognitive Sciences*, 14(11), 497–505. doi:10.1016/j.tics.2010.08.005
- Harmon-Jones, E. (2003). Clarifying the emotive functions of asymmetrical frontal cortical activity. *Psychophysiology*, 40(6), 838–848.
- Harmon-Jones, E., Gable, P. A., & Peterson, C. K. (2010). The role of asymmetric frontal cortical activity in emotion-related

- phenomena: A review and update. *Biological Psychology*, 84 (3), 451–462. doi:10.1016/j.biopsycho.2009.08.010
- Harnad, S. (1990). The symbol grounding problem. *Physica D: Nonlinear Phenomena*, 42(1), 335–346.
- Haxby, J. V., & Gobbini, M. I. (2007, January 7). The perception of emotion and social cues in faces. *Neuropsychologia*, 45(1), 1. doi:10.1016/j.neuropsychologia.2006.11.001
- Hayes, S. C., Barnes-holmes, D., & Roche, B. (2001). *Relational frame theory: A post-skinnerian account of human language and cognition*. New York: Springer. Retrieved from <http://books.google.com/books?hl=en&lr=&id=n4RmapzrihAC&pgis=1>
- Hermans, D., Baeyens, F., & Vervliet, B. (2013). Generalization of acquired emotional responses. *Handbook of Cognition and Emotion*, 117–134.
- Holt, E. B. (1914). *The concept of consciousness*. New York: Macmillan.
- Huang, Y., Zhou, R., Cui, H., Wu, M., Wang, Q., Zhao, Y., & Liu, Y. (2015). Variations in resting frontal alpha asymmetry between high-and low-neuroticism females across the menstrual cycle. *Psychophysiology*, 52(2), 182–191.
- James, D. (1998). Stress takes its toll at call centers. *Business Review Weekly*, 20, 87–89.
- Jansz, J., & Timmers, M. (2002). Emotional dissonance when the experience of an emotion jeopardizes an individual's identity. *Theory & Psychology*, 12(1), 79–95.
- Kensinger, E. A. (2004). Remembering emotional experiences: The contribution of valence and arousal. *Reviews in the Neurosciences*, 15(4), 241–253. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/15526549>
- Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: A review and analysis. *Brain Research. Brain Research Reviews*, 29(2–3), 169–195. doi:10.1016/S0165-0173(98)00056-3
- Klimesch, W. (2012).  $\alpha$ -Band oscillations, attention, and controlled access to stored information. *Trends in Cognitive Sciences*, 16 (12), 606–617. doi:10.1016/j.tics.2012.10.007
- Kraemer, J. L. (1997). Comparing crescent and cross [Review of under crescent and cross: The Jews in the middle ages]. *The Journal of Religion*, 77(3), 449–454. Retrieved from <http://www.jstor.org/stable/1206999>
- Lakens, D., Semin, G. R., & Foroni, F. (2012). But for the bad, there would not be good: Grounding valence in brightness through shared relational structures. *Journal of Experimental Psychology: General*, 141(3), 584–594.
- Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H. J., Hawk, S. T., & van Knippenberg, A. (2010). Presentation and validation of the radboud faces database. *Cognition & Emotion*, 24(8), 1377–1388. doi:10.1080/02699930903485076
- Leopold, D. A., & Rhodes, G. (2010). A comparative view of face perception. *Journal of Comparative Psychology*, 124(3), 233–251.
- Lewig, K. A., & Dollard, M. F. (2003). Emotional dissonance, emotional exhaustion and job satisfaction in call centre workers. *European Journal of Work and Organizational Psychology*, 12(4), 366–392.
- Lewin, K. (1935). *A dynamic theory of personality: Selected papers*. (D. K. Adams & K. E. Zener, trans.). New York: McGraw-Hill.
- McConnell, A. R., & Leibold, J. M. (2001). Relations among the implicit association test, discriminatory behavior, and explicit measures of racial attitudes. *Journal of Experimental Social Psychology*, 37(5), 435–442.
- Mcllvane, W. J., & Dube, W. V. (2003). Stimulus control topography coherence theory: Foundations and extensions. *Behavior Analyst*, 2(2), 195–213.
- Mignault, A., & Chaudhuri, A. (2003). The many faces of a neutral face: Head tilt and perception of dominance and emotion. *Journal of Nonverbal Behavior*, 27(2), 111–132.
- Miller, G. A. (2010). Mistreating psychology in the decades of the brain. *Perspectives on Psychological Science*, 5(6), 716–743.
- Miller, G. A., Crocker, L. D., Spielberg, J. M., Infantolino, Z. P., & Heller, W. (2015). Issues in localization of brain function: The case of lateralized frontal cortex in cognition, emotion, and psychopathology. *Current Research and Emerging Directions in Emotion-Cognition Interactions*, 7, 370–379.
- Munnely, A., Dymond, S., & Hinton, E. C. (2010). Relational reasoning with derived comparative relations: A novel model of transitive inference. *Behavioural Processes*, 85(1), 8–17. doi:10.1016/j.beproc.2010.05.007
- Overall, J. E., & Woodward, J. A. (1975). Unreliability of difference scores: A paradox for measurement of change. *Psychological Bulletin*, 82(1), 85–86. Retrieved from <http://psycnet.apa.org/index.cfm?fa=buy.optionToBuy&id=1975-08623-001>
- Papousek, I., Reiser, E. M., Weber, B., Freudenthaler, H. H., & Schuler, G. (2012). Frontal brain asymmetry and affective flexibility in an emotional contagion paradigm. *Psychophysiology*, 49(4), 489–498. doi:10.1111/j.1469-8986.2011.01324.x
- Patel, S. H., & Azzam, P. N. (2005). Characterization of N200 and P300: Selected studies of the event-related potential. *International Journal of Medical Sciences*, 2(4), 147–154.
- Roche, B., & Dymond, S. (2008). A transformation of functions in accordance with the nonarbitrary relational properties of sexual stimuli. *The Psychological Record*, 58, 71–90.
- Salzinger, K. (2008). Skinner's verbal behavior. *International Journal of Psychology and Psychological Therapy*, 8(3), 287–294.
- Schöne, B., Schomberg, J., Gruber, T., & Quirin, M. (2015). Event-related frontal alpha asymmetries: Electrophysiological correlates of approach motivation. *Experimental Brain Research*, 2(234), 559–567.
- Seyama, J., & Nagayama, R. S. (2007). The uncanny valley: Effect of realism on the impression of artificial human faces. *Presence: Teleoperators and Virtual Environments*, 16(4), 337–351. doi:10.1162/pres.16.4.337
- Sifferlin, A. (2013). Why selfies matter. *TIME Magazine*. Retrieved February 7, 2015, from <http://healthland.time.com/2013/09/06/why-selfies-matter/>
- Silveira, M. V., Aggio, N. M., Cortez, M. D., Bortoloti, R., Rico, V. V., & de Rose, J. C. (2015). Maintenance of equivalence classes and transfer of functions: The role of the nature of Stimuli. *The Psychological Record*, 1(66), 65–74.
- Slattery, B., & Stewart, I. (2014). Hierarchical classification as relational framing. *Journal of the Experimental Analysis of Behavior*, 101(1), 61–75.
- Spokas, M., Luterek, J. A., & Heimberg, R. G. (2009). Social anxiety and emotional suppression: The mediating role of beliefs. *Journal of Behavior Therapy and Experimental Psychiatry*, 40 (2), 283–291.
- Staddon, J. E. R. (2014). *The new behaviorism*. New York: Psychology Press.
- Stemmer, N. (1973). Language acquisition and classical conditioning. *Language and Speech*, 16(3), 279–282.
- Teplan, M. (2002). Fundamentals of EEG measurement. *Measurement Science Review*, 2(2), 1–11.

- Todorov, A., & Engell, A. D. (2008). The role of the amygdala in implicit evaluation of emotionally neutral faces. *Social Cognitive and Affective Neuroscience*, 3(4), 303–312.
- Tolman, E. C. (1932). *Purposive behavior in animals and men*. Oakland, CA: University of California Press.
- Vasconcelos, M. (2008). Transitive inference in non-human animals: An empirical and theoretical analysis. *Behavioural Processes*, 78(3), 313–334. doi:10.1016/j.beproc.2008.02.017
- Wheatley, T., Weinberg, A., Looser, C., Moran, T., & Hajcak, G. (2011). Mind perception: Real but not artificial faces sustain neural activity beyond the N170/VPP. *PLoS One*, 6(3), e17960–e17968. doi:10.1371/journal.pone.0017960
- Wheeler, R. E., Davidson, R. J., & Tomarken, A. J. (1993). Frontal brain asymmetry and emotional reactivity: A biological substrate of affective style. *Psychophysiology*, 30(1), 82–89. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/8416065>
- Wynne, C. D. L. (1995). Reinforcement accounts for transitive inference performance. *Animal Learning & Behavior*, 23(2), 207–217. doi:10.3758/BF03199936