Adaptive working memory training reveals a negligible effect of emotional stimuli over cognitive processing

Francisco J. Román [a], María J. García-Rubio [a], Jesús Privado [b], Dominique Kessel [a], Sara López-Martín [a], Kenia Martínez [a,c], Pei-Chun Shih [a], Manuel Tapia [a], Juan Manuel Serrano [a], Luis Carretié [a], Roberto Colom [a,*]

[a] Universidad Autónoma de Madrid, 28049 Madrid, Spain
[b] Universidad Complutense de Madrid, 28223 Madrid, Spain
[c] Hospital Gregorio Marañón, 28007 Madrid, Spain

ARTICLE INFO

ABSTRACT

Here we analyze how performance differences in an adaptive cognitive training regime based on the n-back task interact with emotional stimuli (scenes and faces) varying in their valence (negative, positive, and neutral). One hundred and three participants completed four training sessions across 2 weeks showing remarkable improvements from time to time. Results revealed similar results for faces and scenes regarding accuracy levels across increased complexity levels. However, reaction times (RTs) were sensitive to emotional conditions to some extent. Observed faster RTs to negative faces (disgust) were consistent with the negativity bias phenomenon, but this effect vanished for the highest levels of processing complexity. It is suggested that emotional information contents fail to interact with cognition when there are no cognitive resources left after the primary task is addressed.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

Cognitive training is capturing substantial scientific and media interest. Elderly people (e.g., Fandakova, Shing, & Lindenberger, 2012; Zelinski, 2009), ADHD patients (e.g., Beck, Hanson, Puffenberger, Benninger, & Benninger, 2010; Klingberg et al., 2005) or schizophrenic individuals (e.g., Subramaniam et al., 2012; Twamley, Jeste, & Bellack, 2003) are usual target populations. The training programs are based on different cognitive functions, but given its theoretical relevance (Ackerman, Beier, & Boyle, 2005; Colom, Abad, Quiroga, Shih, & Flores-Mendoza, 2008; Cowan et al., 2005; Martínez et al., 2011), working memory (WM) is frequently addressed. Two key issues have been discussed with respect to WM training. Firstly, are improvements in WM related with increased scores in far-transfer measures such as fluid intelligence tests? (Chooi & Thompson, 2012; Colom et al., 2010; Colom et al., 2013; Jaeggi, Buschkuehl, Jonides & Perrig, 2008; Jaeggi, Buschkuehl, Jonides, & Shah, 2011; Jaeggi et al., 2010; Redick et al., 2013; Shipstead, Redick, & Engle, 2012). Secondly, are these WM changes associated with variations in brain structure and function? (e.g., Buschkuehl, Hernandez-Garcia, Jaeggi, Bernard, &

Jonides, 2014; Jaušovec & Jaušovec, 2012; Jolles, Grol, Van Buchem, Rombouts, & Crone, 2010; McKendrick, Ayaz, Olmstead, & Parasuraman, 2014; Olesen, Westerberg, & Klingberg, 2004; Takeuchi et al., 2011). However, little attention has been devoted to the nature (neutral or emotional) of the information to be processed.

The n-back task has been used for designing WM training programs (Colom et al., 2013; Jaeggi, Buschkuehl, Jonides, & Perrig, 2008; Jaeggi, Buschkuehl, Jonides, & Shah, 2011; Stephenson & Halpern, 2013). However, to our knowledge, there are not studies addressing the interaction between cognitive performance observed in these training programs and emotional stimuli. For filling this gap here we study two types of stimuli, scenes and faces (Coan & Allen, 2007), because (a) they are important in evolutionary terms (Carretié et al., 2013), and (b) they are known to interact with cognitive requirements (e.g., Eastwood, Smilek, & Merikle, 2003). Furthermore, the most frequently administered visual stimuli in affective neuroscience are based on faces depicting different emotional expressions (e.g., POFA, Ebner, Riediger, & Lindenberger, 2010; KDEF, Lundqvist and Litton, 1998) or scenes showing positive, neutral and negative displays (e.g., IAPS, Lang, Bradley, & Cuthbert 2005).

Scenes are associated with affective reactions, such as phobias, and they have a direct explicit affective meaning (Carretié et al.,

* Corresponding author. Tel.: +34 91 497 41 14.
E-mail address: roberto.colom@uam.es (R. Colom).

2013). They correlate with psychophysiological responses providing validity to subjectively reported emotion induction (Lang & Bradley, 2007), and with neural activation (e.g. Britton, Taylor, Sudheimer, & Liberzon, 2006; Carretié, Hinojosa, Martín-Loeches, Mercado, & Tapia, 2004; Carretié et al., 2013; Olofsson, Nordin, Sequeira, & Polich, 2008). Faces are important in social interactions (Frith, 2007). Psychophysiological and brain responses to faces have been extensively studied (e.g. Achaibou, Pourtois, Schwartz, & Vuilleumier, 2008; Aguado et al., 2012, 2013; Britton et al., 2006). Furthermore, scenes and faces allow using negative, neutral and positive emotional valences.

Here we investigate if there are behavioral differences between negative, neutral and positive stimuli (faces and scenes) comprised in an adaptive cognitive training based on the single *n*-back task. The interaction between emotion and memory has been investigated within several frameworks: episodic memory (e.g., Pillemer, Goldsmith, Panter, & White, 1988), long-term storage (e.g., Buchanan & Adolphs, 2003; Charles, Mather, & Carstensen, 2003) or working memory (e.g., Gray, 2001; Gray, Braver, & Raichle, 2002) but findings are far from consistent when the goal of the study is the nature of the stimuli (Gotoh, 2008; Holmes, Nielsen, Tipper, & Green, 2009; Kensinger & Corkin, 2003; Levens & Gotlib, 2010; Levens & Gotlib, 2012; Lindström & Bohlin, 2011).

The theoretical framework for the present study is based on previous evidence showing that individual differences in working memory can be attributed to capacity limitations for keeping a reliable mental representation of the relevant information (Colom, Shih, Flores-Mendoza, & Quiroga, 2006; Colom et al., 2013; Martínez et al., 2011). Therefore, we predict that emotions evoked by scenes and faces will interact with cognitive performance in the completed adaptive training program for the simplest processing levels. However, emotion will loose their evocative role at increased levels of cognitive complexity.

2. Method

2.1. Participants

Participants were recruited from the Faculty of Psychology (*N* = 76, 72.19%) and the Faculty of Computer Science (*N* = 27, 27.81%) at Universidad Autónoma de Madrid (*N* = 103, 61.20% were females). Their mean of age was 19.86 (SD = 3.85). Participants were randomly assigned to two groups: Face training (*N* = 51) and Scene training (*N* = 52 with 20.02). This study followed the Declaration of Helsinki.

2.2. Procedure and stimuli

The training regime consisted of four sessions completed across 2 weeks. In each session, participants performed the adaptive *n*-back task for approx. 30 min within individual cabins and under strict supervision. Participants were instructed to respond – as accurately and quickly as possible – each time the current stimulus was identical to that presented n positions back in the sequence, pressing the space bar for targets only. In the first session, participants started at level 1 (1-back) for each emotional condition. The difficulty level increased or decreased according to their performance at each emotional block. *n*-Back level was increased when participants had less than three omission or commission errors. *n*-Back level was reduced when participants committed more than five omission and commission errors.

For successive sessions, each participant began at the level achieved in the previous session for all stimuli. For example, if a participant achieved level 4 for negative stimuli, level 3 for positive stimuli, and level 2 for neutral stimuli in session one, then session

two begins in level 4 for negative stimuli, level 3 for positive stimuli, and level 2 for neutral stimuli.

Faces and scenes (negative, positive and neutral) were employed for all sessions. Specifically, each session included 12 blocks (4 blocks per emotional condition: neutral, negative and positive) with 20 + n stimuli for each block. All stimuli were displayed at a rate of 3 s (stimulus length, 500 ms interstimulus interval, 2500 ms; see Fig. 1 for faces and Fig. 2 for scenes). The order of emotional conditions was randomized for each participant within sessions. The four blocks for each emotional condition were successively administered. For example, the sequence in one session for a participant in the face group was: blocks 1–4: neutral faces, blocks 5–8: positive faces, and blocks 9–12: negative faces.

In the training group using faces, happy expressions were employed as positive stimuli, since expressions of positive valence other than happiness are problematic with respect to recognition rate (Tracy & Robins, 2008). Moreover, the expression used as negative stimulus was also single: disgust faces were selected, since it is better recognized (in terms of both reaction times and accuracy) than other negative expressions, such as fear or sadness (Tracy & Robins, 2008). Neutral, negative and positive emotional faces were selected from the FACES database (Ebner et al., 2010).

In the training group using scenes, stimuli were selected from the International Affective Picture System (IAPS) (Lang et al., 2005) and from EmoMadrid (http://www.uam.es/carretie/EmoMadrid.htm). All emotional scenes were chosen according to valence and arousal average assessments provided by these databases (see Fig. 1 for an example of the *n*-back task with emotional stimuli).

2.3. Analyses

First, we computed the percentage of improvement for each condition according to this formula (Chooi & Thompson, 2012):

$$\% \, Improvement$$
$$= \frac{Avg. \, Highest \, Training \, score - Avg. \, First \, Training \, score}{Avg. \, Highest \, score} \times 100$$

$$(1)$$

A one-way ANOVA was computed to check if improvement across training sessions was different for each valence (negative, neutral and positive). This analysis was done separately for faces and scenes.

Second, repeated measures (4 × 3) ANOVAs – Sessions (S1, S2, S3, S4) × Valence (negative, positive and neutral) – were computed for both accuracy (*n*-back level achieved in each session) and reaction times (RTs) (correct trials only). Again, these analyses were done separately for faces and scenes. Finally, post hoc analyses were computed using the Bonferroni correction.

3. Results

3.1. Improvement

The percentage of improvement for faces was 52.08% for negative, 50.71% for neutral and 53.18% for positive stimuli. For scenes, those percentages were 55.89% for negative, 54.71% for neutral and 55.08% for positive stimuli. Therefore, values were very similar for all emotional conditions in both types of training. The computed ANOVAs failed to show significant differences for all comparisons (*p* > .05) in both training conditions.

3.1.1. n-Back level

A 4 × 3 ANOVA (Session × Valence) was computed for the training conditions. The only significant result was the main effect of