

BRIEF REPORTS

Musical expertise modulates the effects of visual perceptual load

TONY RO

City College of the City University of New York, New York

ASHLEY FRIGGEL

University of Texas Southwestern Medical School, Dallas, Texas

AND

NILLI LAVIE

University College London, London, England

Several studies have demonstrated that faces are processed differently from other types of objects, implicating a special role that faces have within the human visual system. However, other studies have suggested that faces may be special only in that they constitute a highly familiar category of visual objects with which most humans have expertise. In this study, we tested a group of expert musicians with a musical instrument classification task during which irrelevant images of musical instruments were presented as visual distractors under varying conditions of perceptual load. Unlike nonmusicians (who had been tested in Lavie, Ro, & Russell, 2003, using the same paradigm as in the present study), the musicians processed these irrelevant images of musical instruments even under conditions of high perceptual load. These results suggest that musical instruments are processed automatically and without capacity limits in subjects with musical expertise and implicate a specialized processing mechanism for objects of high familiarity.

Whether faces are innately special or simply belong to a highly experienced category of visual objects has been the subject of considerable debate (see, e.g., Kanwisher, 2000). Many studies have shown that faces may be differentially processed in comparison with other object categories (for review, see Farah, Wilson, Drain, & Tanaka, 1998). For example, faces are affected disproportionately by stimulus inversion (Yin, 1969), are processed more holistically than are other types of objects (Tanaka & Farah, 1993), and may be processed within a dedicated brain region (De Renzi, Perani, Carlesimo, Silveri, & Fazio, 1994; Desimone, Albright, Gross, & Bruce, 1984; Kanwisher, McDermott, & Chun, 1997). We have also provided evidence that human faces may constitute a special stimulus for visual attention (Lavie, Ro, & Russell, 2003; Ro, Friggel, & Lavie, 2007; Ro, Russell, & Lavie, 2001). For example, whereas nonface distractors can be ignored successfully when the relevant task involves high perceptual load, interference by distractor faces is not modulated by the level of load in the task. These results support the claim that processing of faces may be mediated by a specialized face-processing system (a "module") for which processing capacity is unaffected

by the level of load in a nonface search task (using written names as the search stimuli).

Although there has been overwhelming evidence suggesting that the processing of faces is special, research has shown that words and objects (MacLeod, 1991; Tipper & Driver, 1988) can also be processed automatically, and other studies have shown that similar processing mechanisms for faces can be acquired for other object categories with visual expertise (Bukach, Gauthier, & Tarr, 2006). For example, when recognizing dogs, dog show judges demonstrated inversion costs that are similar to the costs measured with faces (Diamond & Carey, 1986). More recently, subjects who were trained on subordinate classification of novel objects demonstrated holistic processing for novel objects that are similar to faces (Gauthier & Tarr, 1997). Furthermore, neuroimaging studies of bird and car experts have demonstrated that the same brain area involved with processing faces in the fusiform gyrus is also involved with processing the objects of their expertise (i.e., birds for bird experts and cars for car experts; Gauthier, Skudlarski, Gore, & Anderson, 2000; Xu, 2005).

In the present study, we examined whether expertise with nonface objects can also render those objects into special

T. Ro, tro@ccny.cuny.edu

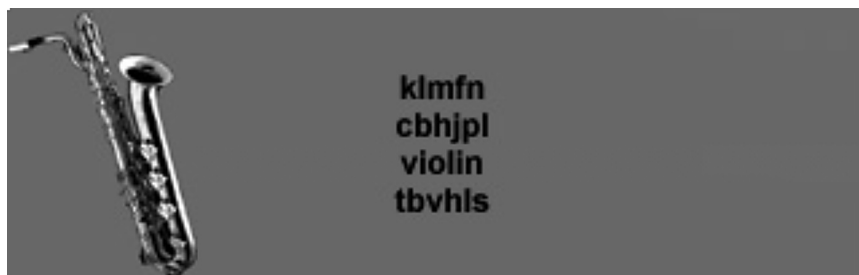


Figure 1. An example of the stimulus displays.

stimuli for visual attention in the same way as has been shown for faces. We thus tested whether the perceptual processing of objects of expertise may also be unaffected by the level of perceptual load in a name search task. To this end, we recruited highly skilled musicians as subjects in an experiment employing the same paradigm as in one of our previous experiments (Lavie et al., 2003). The musicians were asked to classify names of musical instruments into wind versus string instruments, and we tested whether the level of interference by response-congruent or -incongruent musical instrument distractors (see Figure 1) would be modulated by the level of perceptual load in the name search task. We modulated perceptual load by varying the number of search items to produce a relevant search set size of one, two, or four items. This manipulation of perceptual load through the relevant search set size is well established (see Lavie & Tsai, 1994, for a review) and has been shown to modulate nonface distractor processing (e.g., letters or other meaningful nonface objects) in many previous studies (e.g., Lavie & Cox, 1997; see Forster & Lavie, 2007, 2008, for more recent examples).

METHOD

Subjects

Eight music students (2 males, 6 females; 18–21 years old; *M* = 19.4) from the Shepherd School of Music at Rice University participated in this experiment after giving informed consent. All subjects had at least 3 years of intensive experience playing a musical instrument (*M* = 9.9 years; range, 3–16 years). They therefore had extensive exposure to musical instruments on a regular basis.

The data from these musicians were compared with the data from Lavie et al.'s (2003) Experiment 4. These latter data were collected from a group of 12 undergraduate subjects (5 males, 7 females; 19–24 years old; *M* = 19.6),¹ who had not been selected on the basis of musical expertise. Although we cannot rule out that some of those nonmusicians had some musical expertise, because they had not been screened explicitly for this possibility, our present data would be conservatively biased against finding any differences between

our two groups of subjects if some of our nonmusician subjects actually had musical expertise.

Stimuli and Procedure

On each trial, subjects were asked to search for a name of a musical instrument among zero, one, or three nonword consonant letter strings (set sizes of one, two, or four, respectively) and to make speeded classification responses as to whether it was a string or a wind instrument, while ignoring a peripheral distractor (presented 5° from fixation either on the left or right with equal probability). A total of six string (banjo, guitar, harp, piano, sitar, violin) and six wind (bagpipe, flute, saxophone, trombone, trumpet, tuba) instruments were used. Both word and nonword letter strings were presented in black, lowercase Arial font on a gray background. The size of the letter strings was approximately 0.5° in height, with a width ranging from 1.5° to 3.0°, depending on the number of letters in the string. The nonword consonant letter strings had from 5 to 10 letters in each string, so as to roughly match the number of letters in the musical instrument names. Because the subjects had to first localize the word from any nonwords, extensive perceptual processing in the form of a visual search was required before the categorization response could be made, and it is this perceptual search stage for which load was varied through the increase in the search set size. The distractor was a grayscale image of a musical instrument, one from the opposite category (incongruent condition) or from the same category (congruent condition) as that indicated by the name. The distractors were 5° in height and 2.5°–5° in width, depending on the geometry of the musical instrument. The distractor image and the target word, along with the nonwords, were presented simultaneously on each trial and remained present until the participant responded or 2 sec had elapsed. The intertrial interval was 500 msec. Figure 1 shows a typical trial. Each subject ran through a practice block of 48 trials followed by four experimental blocks of 96 trials. Within each block, all conditions were intermixed randomly. Before the experiment, all subjects were shown a sheet of paper on which all of the labeled musical instruments used in this experiment were depicted.

RESULTS

Table 1 presents the mean reaction times (RTs) and percentage errors across the musicians and nonmusicians for

Table 1
Mean Reaction Times (RTs, in Milliseconds) and Error Rates for Each Condition

Set Size	Nonmusicians								Musicians							
	Congruent				Incongruent				Congruent				Incongruent			
	RT		Error Rate		RT		Error Rate		RT		Error Rate		RT		Error Rate	
	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>
1	732	30	1.9	0.7	781	26	5.5	0.6	668	29	3.3	1.0	739	42	4.1	1.1
2	807	24	2.3	0.7	828	28	4.1	0.8	749	32	3.1	1.1	780	28	3.7	1.1
4	1,032	32	3.6	0.9	1,012	26	5.9	1.3	913	34	4.7	2.3	952	35	4.9	1.8

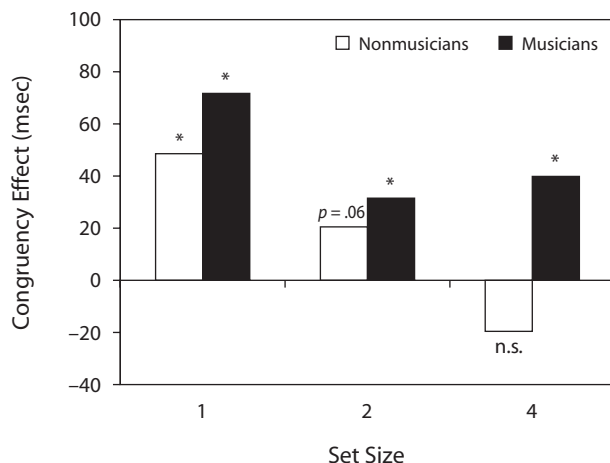


Figure 2. The mean congruency effects (incongruent minus congruent distractor conditions) for each of the set sizes in the nonmusicians (light bars) and the musicians (dark bars). * $p < .05$; n.s., not significant.

each of the congruency \times load conditions. A two-way, within-subjects ANOVA on the musician group RTs, with factors of congruency (congruent, incongruent) and relevant set size (one, two, four), showed a main effect of congruency [$F(1,7) = 21.46$, $MS_e = 1,272$, $p = .002$]: Target RTs were faster when the peripheral distractor was congruent with the target word as opposed to when the distractor was incongruent. The main effect of set size was also significant [$F(2,14) = 61.69$, $MS_e = 3,649$, $p < .001$], reflecting slower responses with increased set size (704, 765, and 933 msec for the one, two, and four set size conditions, respectively). Most importantly, and unlike the effects measured with nonmusicians (see Figure 2), there was no interaction of congruency and load [$F(3,21) = 1.84$, $MS_e = 980$, $p = .20$]. RTs in the congruent conditions were significantly faster than those in the incongruent conditions for all set sizes ($p < .005$, $p < .03$, and $p < .02$, with one-tailed t tests for the one, two, and four set size conditions, respectively) for these musicians.

This difference in the effect of set size on distraction by irrelevant musical instrument distractors between the musicians and nonmusicians was further confirmed by a linear regression analysis on the magnitude of the congruency effects across the different set sizes for the two groups. Whereas the congruency effects were linearly modulated by increasing set size for the nonmusicians ($r^2 = .99$, $p < .005$), there was no such effect for the musicians ($r^2 = .56$, $p > .10$). This difference in the extent to which congruency effects could be accounted for by an increased level of set size between the two subject groups was highly significant ($z = 3.61$, $p < .001$). In addition, the musicians showed larger congruency effects ($M = 48$ msec) than did the nonmusicians ($M = 17$ msec) [$F(1,18) = 6.97$, $MS_e = 1,005$, $p < .02$], and this difference was found across all set sizes. In other words, there was no interaction of set size, congruency, and group [$F(2,36) = 1.67$, $MS_e = 910$, $p = .20$]. Thus, musicians cannot ignore musical instruments, even under conditions of higher load in the relevant

task that allows for successful ignoring of such distractors in a nonmusician group.

Finally, as can be seen in Table 1, the musicians were faster than the nonmusicians, but this effect of group did not achieve significance [$F(1,18) = 2.62$, $MS_e = 46,480$, $p = .12$]. Perhaps more importantly, set size had a similar effect on the search RTs for the musicians and nonmusicians; the interaction between group and set size was not significant [$F(2,36) = 1.54$, $MS_e = 2,805$, $p = .23$]. These results indicate that the two groups differed specifically in the visual processing of the musical instrument distractors but not in the name search task itself.

DISCUSSION

The present findings show that expertise with nonface objects can result in a special status of these objects for attention, similar to that previously found for faces. Our musician subjects, who practice and perform music extensively and are around musical instruments regularly, showed specialized processing of musical instruments in comparison with the nonmusician subjects. Whereas for the nonmusicians interference by musical instrument distractors was modulated by the level of load in a name search task, the interference by the same musical instrument distractor for the musicians was not modulated by the level of load in the name search task.

Thus, as with faces, the processing of objects of expertise does not appear to be subjected to attentional capacity limits. This result suggests that prioritization of objects of expertise may provide a mechanism for the efficient processing of objects of expertise so that these can be tagged and processed in an automatized manner that does not depend on attentional capacity.

It is important to note that the difference between musicians and nonmusicians in our study was confined to the processing of visual images of 3-D musical instruments. Performance of the name search task and the effects of set size on the search RTs were comparable between the groups. This pattern of findings can be illuminating with respect to the locus of the effects reported. On one hand, it suggests that the processing advantage with musical expertise is rather specific to the visual, rather than the verbal (cf. Franklin et al., 2008), processing of the objects of expertise. On the other hand, the heterogeneous range of musical instruments that we used suggests that the expertise effects measured in this study are unlikely to be based on the presence of specific local visual features, but rather indicate a more general and global level of perceptual expertise for musical instruments of various kinds. This suggestion is in line with previous findings that, relative to nonmusicians, musicians generally are faster and more accurate in tests of visual spatial perception and mental imagery (Brochard, Dufour, & Després, 2004; Patston, Hogg, & Tippett, 2007), have faster inter-hemispheric visual transfer times (Patston, Kirk, Rolfe, Corballis, & Tippett, 2007), and have larger cortical motor representations (Elbert, Pantev, Wienbruch, Rockstroh, & Taub, 1995). Note that these findings were not confined to the representations of musical instruments. However,

some of the tests in which the musicians showed performance advantages involved spatial-relational processing of visual stimuli (Brochard et al., 2004; Patston, Hogg, & Tippett, 2007). It is thus possible that the specificity of the advantage found in our study to the images of musical instruments, rather than to their names, reveals a more fundamental difference between the spatial and object recognition processing of visual 2- and 3-D objects as compared with verbal information (such as the letter strings in the name search task). The effects of musical expertise might therefore be different from other forms of expertise, such as expertise with faces, cars, and birds, that may be a consequence of extensive experience with the perceptual categorization of specific features and objects of expertise (Mondloch, Maurer, & Ahola, 2006; Tanaka, Curran, & Sheinberg, 2005).

Our results may also suggest that, at least in the case of attentional load, faces may be special, not because of some special innate feature but because of the extensive experience people have in classifying them (as with words in the Stroop effect). This conclusion is in line with other demonstrations that the objects of high expertise acquire processing characteristics similar to those for faces (Gauthier et al., 2000; Gauthier & Tarr, 1997), including similar neural correlates (Gauthier et al., 2000; Xu, 2005). Finally, a recent study (Thoma & Lavie, 2009) has demonstrated that distractor face processing has face-specific capacity limits. Interference by response-incongruent versus -congruent distractor faces in the task used by Lavie et al. (2003) was modulated by the level of load on faces. As soon as the names in the search plus classification task were replaced with faces, the effects of load on interference by the distractor faces were restored. Whether objects of expertise also show similar object-specific capacity limits would be an interesting topic for future research.

AUTHOR NOTE

This research was supported in part by grants from the National Institute of Health and National Science Foundation to T.R. and from the Human Frontiers Science Program, Wellcome Trust, and Medical Research Council to N.L. Address correspondence to T. Ro, City College of the City University of New York, Department of Psychology, NAC 7/120, 160 Convent Ave., New York, NY 10031 (e-mail: tro@ccny.cuny.edu).

REFERENCES

- BROCHARD, R., DUFOUR, A., & DESPRÉS, O. (2004). Effect of musical expertise on visuospatial abilities: Evidence from reaction times and mental imagery. *Brain & Cognition*, *54*, 103-109.
- BUKACH, C. M., GAUTHIER, I., & TARR, M. J. (2006). Beyond faces and modularity: The power of an expertise framework. *Trends in Cognitive Sciences*, *10*, 159-166.
- DE RENZI, E., PERANI, D., CARLESIMO, G. A., SILVERI, M. C., & FAZIO, F. (1994). Prosopagnosia can be associated with damage confined to the right hemisphere: An MRI and PET study and a review of the literature. *Neuropsychologia*, *32*, 893-902.
- DESIMONE, R., ALBRIGHT, T. D., GROSS, C. G., & BRUCE, C. (1984). Stimulus-selective properties of inferior temporal neurons in the macaque. *Journal of Neuroscience*, *4*, 2051-2062.
- DIAMOND, R., & CAREY, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology: General*, *115*, 107-117.
- ELBERT, T., PANTEV, C., WIENBRUCH, C., ROCKSTROH, B., & TAUB, E. (1995). Increased cortical representation of the fingers of the left hand in string players. *Science*, *270*, 305-307.
- FARAH, M. J., WILSON, K. D., DRAIN, M., & TANAKA, J. N. (1998). What is "special" about face perception? *Psychological Review*, *105*, 482-498.
- FORSTER, S., & LAVIE, N. (2007). High perceptual load makes everybody equal: Eliminating individual differences in distractibility with load. *Psychological Science*, *18*, 377-381.
- FORSTER, S., & LAVIE, N. (2008). Failures to ignore entirely irrelevant distractors: The role of load. *Journal of Experimental Psychology: Applied*, *14*, 73-83.
- FRANKLIN, M. S., MOORE, K. S., YIP, C.-Y., JONIDES, J., RATTRAY, K., & MOHER, J. (2008). The effects of musical training on verbal memory. *Psychology of Music*, *36*, 353-365.
- GAUTHIER, I., SKUDLARSKI, P., GORE, J. C., & ANDERSON, A. W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neuroscience*, *3*, 191-197.
- GAUTHIER, I., & TARR, M. J. (1997). Becoming a "greeble" expert: Exploring mechanisms for face recognition. *Vision Research*, *37*, 1673-1682.
- KANWISHER, N. (2000). Domain specificity in face perception. *Nature Neuroscience*, *3*, 759-763.
- KANWISHER, N., McDERMOTT, J., & CHUN, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, *17*, 4302-4311.
- LAVIE, N., & COX, S. (1997). On the efficiency of visual selective attention: Efficient visual search leads to inefficient distractor rejection. *Psychological Science*, *8*, 395-398.
- LAVIE, N., RO, T., & RUSSELL, C. (2003). The role of perceptual load in processing distractor faces. *Psychological Science*, *14*, 510-515.
- LAVIE, N., & TSAL, Y. (1994). Perceptual load as a major determinant of the locus of selection in visual attention. *Perception & Psychophysics*, *56*, 183-197.
- MACLEOD, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, *109*, 163-203.
- MONDLOCH, C. J., MAURER, D., & AHOLA, S. (2006). Becoming a face expert. *Psychological Science*, *17*, 930-934.
- PATSTON, L. L. M., HOGG, S. L., & TIPPETT, L. J. (2007). Attention in musicians is more bilateral than in non-musicians. *Laterality*, *12*, 262-272.
- PATSTON, L. L. M., KIRK, I. J., ROLFE, M. H. S., CORBALLIS, M. C., & TIPPETT, L. J. (2007). The unusual symmetry of musicians: Musicians have equilateral interhemispheric transfer for visual information. *Neuropsychologia*, *45*, 2059-2065.
- RO, T., FRIGGEL, A., & LAVIE, N. (2007). Attentional biases for faces and body parts. *Visual Cognition*, *15*, 322-348.
- RO, T., RUSSELL, C., & LAVIE, N. (2001). Changing faces: A detection advantage in the flicker paradigm. *Psychological Science*, *12*, 94-99.
- TANAKA, J. W., CURRAN, T., & SHEINBERG, D. L. (2005). The training and transfer of real-world perceptual expertise. *Psychological Science*, *16*, 145-151.
- TANAKA, J. W., & FARAH, M. J. (1993). Parts and wholes in face recognition. *Quarterly Journal of Experimental Psychology*, *46A*, 225-245.
- THOMA, V., & LAVIE, N. (2009). *The role of perceptual load in processing distractor faces: Evidence for face-specific capacity limits*. Manuscript submitted for publication.
- TIPPER, S. P., & DRIVER, J. (1988). Negative priming between pictures and words in a selective attention task: Evidence for semantic processing of ignored stimuli. *Memory & Cognition*, *16*, 64-70.
- XU, Y. (2005). Revisiting the role of the fusiform face area in visual expertise. *Cerebral Cortex*, *15*, 1234-1242.
- YIN, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology*, *81*, 141-145.

NOTE

1. The gender, mean age, and error rates were not available for 2 of the nonmusician subjects.

(Manuscript received June 30, 2008;
revision accepted for publication November 26, 2008.)