nature
neuroscience

# Amusia is associated with deficits in spatial processing

Katie M Douglas & David K Bilkey

**Amusia (commonly referred to as tone-deafness) is a difficulty in discriminating pitch changes in melodies that affects around 4% of the human population. Amusia cannot be explained as a simple sensory impairment. Here we show that amusia is strongly related to a deficit in spatial processing in adults. Compared to two matched control groups (musicians and non-musicians), participants in the amusic group were significantly impaired on a visually presented mental rotation task. Amusic subjects were also less prone to interference in a spatial stimulus-response incompatibility task and performed significantly faster than controls in an interference task in which they were required to make simple pitch discriminations while concurrently performing a mental rotation task. This indicates that the processing of pitch in music normally depends on the cognitive mechanisms that are used to process spatial representations in other modalities.**

The perception, appreciation and production of music are achieved spontaneously, without conscious effort and in the absence of explicit training[1]. However, a few individuals show profound impairments in the musical domain without any other auditory problems[1]. These individuals are described as having amusia (more commonly known as tone-deafness). This condition was first reported in 1878 as note-deafness[2], but its systematic characterization has only recently begun[3–6]. We now know that amusia emerges in early life and persists throughout adulthood[3]. Individuals with this condition can show impaired performance on basic musical tasks including melodic discrimination and recognition[4] but detection of time changes in music can be unimpaired[7]. Although an inability to sing in tune is characteristic of this disorder, this on its own is not fully diagnostic[4,8]. 'True' amusia is estimated to occur in around 4% of the general population[4].

It has been proposed that a dysfunction in fine-grained pitch perception might be at the root of the melodic discrimination problems in amusic individuals[6]. Contrary to expectations, however, functional magnetic resonance imaging scans of amusic people do not show abnormal processing in the auditory cortex[4,6]. Furthermore, amusic individuals can use the small pitch changes that create inflections in spoken language to differentiate questions from statements[9,10]. When these sentences are stripped of their linguistic content, however, and the underlying sounds are played at their mean fundamental frequency, amusic subjects show considerable impairment in pitch discrimination[10]. This dissociation between spoken and musical pitch perception indicates that amusia might be due to a deficit at a level of processing that is higher than that involved in simple pitch detection[9,11].

One possible higher-order representational structure for musical pitch is spatial[12]. It is a Western cultural norm to represent pitch in a spatial configuration, with higher frequency sounds usually being represented as being higher in space than lower frequency sounds[13]. Until recently, however, it has been unclear whether the representation of sound pitch possesses intrinsic spatial characteristics above and beyond this norm. In a study that explored the spatial representation of pitch height through the pairing of pitch with different spatial response positions using an implicit stimulus-response compatibility (SRC) procedure[14], it was proposed that if spatial codes were assigned to pitch, performance would be better when pitch 'height' (frequency) corresponded to the response location than when it did not. The subjects' performance was faster and more accurate when they responded to high-frequency tones with a key placed higher in space and to low-frequency tones with a lower key than when the relationship was reversed[14]. These findings indicate that the representation of melodic pitch might have an intrinsically spatial basis. The aim of our study was to test the hypothesis that amusia is related to problems in spatial processing or the spatial representation of pitch. For this purpose, we divided 34 right-handed subjects into two control groups of non-amusic subjects who were either musicians or non-musicians, and an amusic group. Amusia was determined using a subtest of the Montreal Battery of Evaluation of Amusia (MBEA)[3]. We compared performance on a spatial task (mental rotation) and a non-spatial control task (animal matching). All subjects then performed an SRC-type test. Finally, we assessed performance on a pitch discrimination task while subjects simultaneously conducted mental rotation or animal matching interference tasks.

## RESULTS

### Montreal Battery of Evaluation of Amusia (MBEA)

When we plotted the distribution of scores on the contour-violated subtest of the MBEA for all subjects it was apparent that there was a bimodal distribution with clear separation between groups occurring at the score of 22–23 (**Fig. 1**). As an MBEA score of 22 has previously been used to discriminate between amusic and non-amusic individuals[3], we also used this criterion. As a result, we classified as amusic the eight

---

Department of Psychology, 95 Union St, University of Otago, PO Box 56, Dunedin 9054, New Zealand. Correspondence should be addressed to D.K.B. (dbilkey@psy.otago.ac.nz).

Received 4 April; accepted 18 May; published online 24 June 2007; doi:10.1038/nn1925

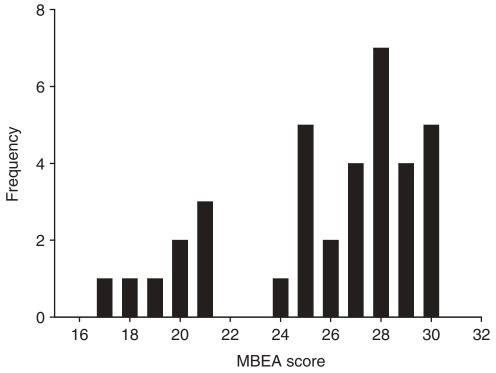

**Figure 1** Distribution of scores on the contour-violated subtest of the MBEA across all right-handed subjects. Individuals scoring less than 22 were classified as amusic.

**Table 1 General characteristics of the experimental groups**

| | Musicians | Non-musicians | Amusic |
|---|---|---|---|
| Age (years) | 22.1 ± 0.8 | 21.8 ± 0.8 | 22.0 ± 1.1 |
| Education (years) | 16.4 ± 0.4 | 16.0 ± 0.2 | 16.7 ± 0.7 |
| MBEA subtest score | 28.1 ± 0.5 | 26.6 ± 0.5 | 19.8 ± 0.5*** |
| Musical memory score | 95.4 ± 1.5% | 70.2 ± 8.0% | 42.9 ± 9.4%* |
| Musicality belief (1 = not musical, 8 = musical) | 7.1 ± 0.2 | 3.3 ± 0.5 | 2.1 ± 0.4** |
| Years of musical training | 15.1 ± 0.9 | 1.6 ± 0.6 | 1.9 ± 0.8** |
| Hours per day spent listening to music | 1.9 ± 0.4 | 2.2 ± 0.6 | 2.6 ± 0.6 |

Data presented as mean ± s.e.m. For comparisons between amusic and control groups: *significantly different from both control groups ($P < 0.01$); **significantly different from the musician group ($P < 0.0001$); ***significantly different from both control groups ($P < 0.0001$).

subjects who scored below 22. The non-amusic group was then divided into two control groups—musicians ($n = 14$) and non-musicians ($n = 12$)—on the basis of their responses to the Musicality Questionnaire. Males and females were distributed equally in the musician and non-musician groups, but there were fewer males ($n = 2$) in the amusic group. Individuals from the musician and non-musician groups achieved similar scores on the MBEA contour-violated subtest.

There were no significant differences between the three groups when they were compared for age and years of formal education (**Table 1**). Responses on the musical memory task for tunes rated as familiar (rating of 3 or less) were different for the three groups ($F = 15.9$, $P < 0.0001$). A post-hoc Newman-Keuls test revealed that the amusic group was significantly ($P < 0.01$) worse at determining whether the melody moved to a higher or lower note on the second note of the familiar tune compared to both of the control groups. This did not seem to be due to differences in tune familiarity, as the ratings for the non-musicians and amusics were not significantly different (Mann-Whitney U test, $P > 0.05$) on this measure, although musicians regarded the tunes as more familiar than either of the other groups. Scores on the musicality-belief question differed between the three groups (Kruskal-Wallace $\chi^2 = 24.4$, $P < 0.0001$). Post-hoc Mann-Whitney tests showed that the amusic group rated themselves as being less musical than the musicians ($P < 0.0001$), but the difference between amusics and non-musicians only approached significance ($P = 0.07$). As expected, the musician group had significantly more years of musical training than the other two groups ($t_{20} = 9.5$, $P < 0.0001$). However, non-musicians and amusic subjects had similar mean amounts of musical training ($t_{18} = 0.3$, $P > 0.05$). In addition, all groups spent a similar amount of time listening to music each day ($F_{2,31} = 0.37$, $P > 0.05$).

## Mental rotation and animal matching tasks

When we compared baseline performance on the mental rotation and animal matching tasks, we found that the error rates differed significantly between groups and across tasks (**Fig. 2**). The amusic group performed markedly more poorly than controls on the mental rotation task ($t_{32} = 6.08$, $P < 0.0001$) but not on the animal matching task ($t_{32} = 0.89$, $P > 0.1$). When the control group was broken into musician and non-musician subgroups, an ANOVA revealed a significant group ($F_{2,39} = 16.2$, $P < 0.0001$) and task ($F_{1,39} = 173.5$, $P < 0.0001$) effect and a significant group × task interaction ($F_{2,39} = 21.5$, $P < 0.0001$). Subsequent post-hoc analysis (Newman-Keuls) showed that musicians and non-musicians performed similarly on the mental rotation task ($P > 0.05$), whereas amusic subjects made significantly more errors than either of these control groups ($P < 0.01$ for both comparisons). For the animal matching control task, however, there were no significant differences between groups (all $P > 0.05$). When we analyzed the mean time taken to complete the mental rotation and animal matching tasks, we found no significant differences in comparisons (*t*-tests) between the control and amusic groups. A group × task ANOVA revealed a significant effect of task ($F_{1,39} = 4.1$, $P < 0.05$), but no group effect or interaction. This indicated that the group differences in error performance were not due to a trade-off between time and accuracy.

## Correlation between MBEA and mental rotation performance

When the MBEA subtest score was plotted against mental rotation performance, a striking relationship emerged (**Fig. 3**). Most amusic subjects performed more poorly than the control group subjects, with little overlap between groups. A simple regression analysis of these data revealed a correlation coefficient of $r = -0.73$ ($P < 0.0001$). As females perform more poorly than males on spatial tasks such as mental rotation[15–17], and as there was a female sex bias in the amusic group, we used a hierarchical regression procedure to evaluate whether mental rotation performance explained unique variance in MBEA performance over and above the variance explained by sex. The results showed that whereas sex explained 9% of the variance in MBEA score (not significant), mental rotation accounted for 47% of the

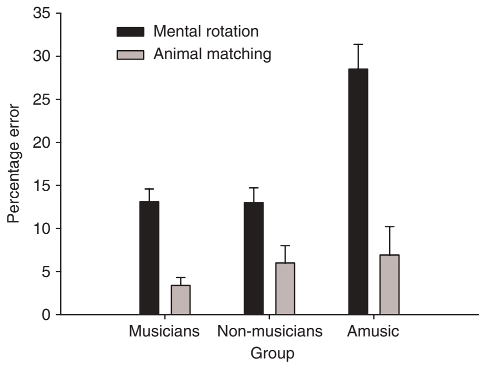

**Figure 2** Errors made during the mental rotation and animal matching control task. Data shown as mean ± s.e.m., $n = 20$ trials. The performance of the amusic group was significantly poorer than of that of the two control groups in the mental rotation task but not in the animal matching task.

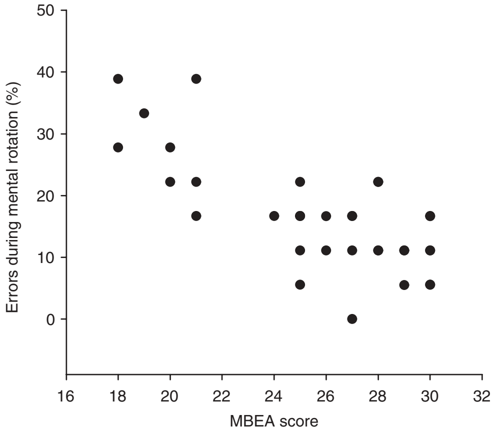

**Figure 3** Relationship between score on the contour-violated subtest of the MBEA and the percentage of errors made during mental rotation testing for all subjects.

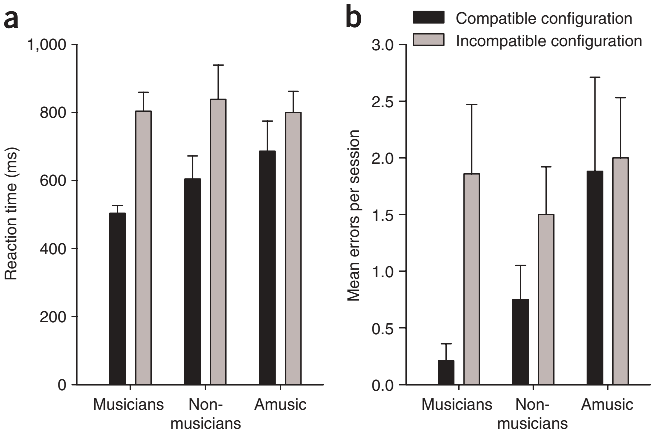

**Figure 4** Performance during the SRC pitch discrimination task for the compatible and incompatible response configurations. (**a**) Reaction time; (**b**) errors out of a maximum of 24. Both panels show mean ± s.e.m.

variance ($P < 0.0001$) with sex partialed out. To determine whether there was any underlying relationship between mental rotation and MBEA performance over and above that generated by group, we ran a further hierarchical regression with sex, group and mental rotation performance entered sequentially. Whereas sex and group accounted for 71% ($P < 0.0001$) of the variance in MBEA score, mental rotation performance accounted for an additional 8% of the variance ($P < 0.01$) with sex and group partialed out.

## SRC pitch discrimination task

All groups were slower when performing the SRC pitch discrimination task in the incompatible response configuration than in the compatible configuration (**Fig. 4a**). An initial ANOVA comparing the control and amusic groups across the two configurations revealed a significant effect of configuration ($F_{1,50} = 92.5$, $P < 0.0001$). Furthermore, the control group was slowed markedly more than the amusic group in the incompatible response configuration, when compared to the compatible response configuration, as revealed by a significant group × configuration interaction ($F_{1,50} = 15.3$), $P < 0.001$). When the ANOVA was repeated with the control group split into the two subgroups, there was again a significant effect of response configuration ($F_{1,39} = 76.1$, $P < 0.0001$) and a significant interaction between response configuration and group ($F_{2,39} = 4.86$, $P < 0.05$). Post-hoc comparisons revealed a significant configuration effect ($F_{1,26} = 57.5$, $P < 0.0001$) and interaction ($F_{1,26} = 11.6$, $P < 0.001$) when the musician and amusic groups were compared and, although a similar trend existed for the comparison between amusic and non-musician subjects, this interaction only approached statistical significance ($P = 0.08$).

All groups made more errors when performing the SRC pitch discrimination task in the incompatible response configuration than in the compatible configuration (**Fig. 4b**). An initial ANOVA comparing control and amusic groups across the two configurations revealed a significant effect of configuration ($F_{1,50} = 15.8$, $P < 0.001$). Furthermore, the control group made more errors than the amusic group in the incompatible response configuration, when compared to the compatible response configuration, as revealed by a significant group × configuration interaction ($F_{1,50} = 10.5$), $P < 0.01$). When the ANOVA was repeated with the control group split into two, we found a significant effect of response configuration ($F_{1,39} = 16.4$, $P < 0.001$) and a significant interaction between response configuration and group ($F_{2,39} = 4.51$, $P < 0.05$). ANOVAs comparing the individual groups revealed a significant effect of configuration ($F_{1,26} = 9.4$, $P < 0.01$) and a group × configuration effect across the musician and amusic groups ($F = 6.98$, $P < 0.05$). That is, musicians were significantly more inaccurate in the pitch discrimination task during the incompatible compared to compatible response configuration, relative to amusic subjects. Although a similar trend existed for the comparison between amusic and non-musician subjects, this interaction was not statistically significant.

## Dual task: pitch discrimination

To assess performance on the dual-task procedure, we analyzed the pitch discrimination and mental rotation/animal matching component tasks separately. For the purpose of display, data are presented with reference to the baseline (no interference) version of the task. Performance on this baseline task was subtracted from performance on the dual-task version so that a positive value indicates poorer performance (interference) compared with the baseline condition.

There was little difference between the control and amusic groups in terms of the increase in pitch-discrimination reaction time (compared to baseline) generated by concurrent performance of the animal

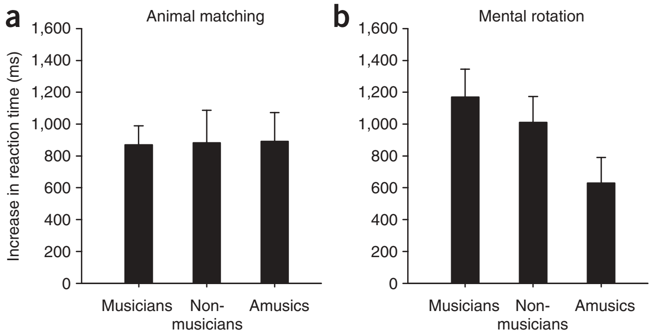

**Figure 5** Increase in reaction time on the pitch discrimination task compared to baseline when this task was performed concurrently with other tasks. Mean ± s.e.m. of increase during concurrent performance of (**a**) the animal matching task and (**b**) the mental rotation task. Amusic subjects were significantly less impaired when performing a concurrent mental rotation task, but not an animal matching task, compared with the two control groups.

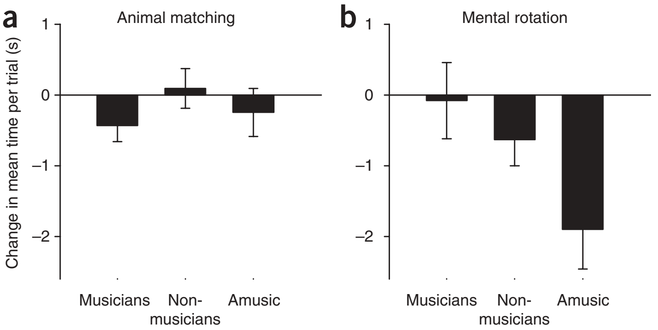

**Figure 6** Changes in performance speed on animal matching and mental rotation tasks when performed concurrently with pitch discrimination. Mean ± s.e.m. change in mean time per trial (compared to baseline). (**a**) Animal matching; (**b**) mental rotation. A negative change indicates improved performance in the dual-task paradigm compared to the no-interference baseline condition (and vice versa). Note that an underlying practice effect has shifted all data toward the negative end of the scale in **Figures 6** and **7**. Amusic subjects were significantly less impaired than controls on the mental rotation task, but not the animal matching task, when performing the concurrent pitch discrimination task.

matching task (**Fig. 5a**). By contrast, amusic subjects showed less interference (at borderline significance) than controls (**Fig. 5b**) when concurrently performing the mental rotation task ($t_{32} = 1.99$, $P = 0.054$). A group × task ANOVA of these data revealed a highly significant interaction ($F_{1,50} = 25.9$, $P < 0.0001$). This indicated that the control group was significantly more impaired on the pitch discrimination procedure by mental rotation interference than by animal matching interference. A subsequent ANOVA with separate control groups revealed a significant interaction between group and type of interference ($F_{2,39} = 8.00$, $P < 0.01$). To determine the source of the interaction, we used an ANOVA to determine that there were significant group × interference interactions between the non-musician and amusic groups ($F_{1,26} = 13.3$, $P < 0.01$), and the musician and amusic groups ($F_{1,26} = 14.6$, $P < 0.001$).

It was possible that the relatively fast reaction times observed in the amusic subjects while they performed the pitch discrimination task concurrently with mental rotation were made at the expense of an increase in errors. On average, however, the combined control group made 2.3 ± 0.5 pitch discrimination errors, whereas the amusic group made only 2.1 ± 0.5 errors ($t_{32} = 0.19$, not significant). When performing the animal matching task, the control and amusic groups made 2.0 ± 0.5 and 2.9 ± 0.9 pitch discrimination errors, respectively ($t_{32} = 0.83$, not significant). These data indicate that the amusic subjects were not making a trade-off between speed and accuracy during the pitch discrimination task. Rather, there is a trend in the opposite direction as the amusic subjects were both faster and slightly more accurate than the control subjects under mental rotation interference conditions.

### Dual task: mental rotation and animal matching

Control subjects completed an average of 16.1 animal matching trials during the dual task procedure, whereas amusic subjects completed 17.0 (not significant). During the mental rotation procedure, however, control subjects completed an average of 16.6 trials, whereas amusic subjects completed 21.7 ($t_{32} = 2.71$, $P < 0.05$). The number of trials completed by each participant was determined not only by their speed on these cognitive tasks, but also by their ability to perform the pitch discrimination task, as the test ceased when the 24 pitch discrimination trials were complete. The mean time per trial was calculated to allow for this fact.

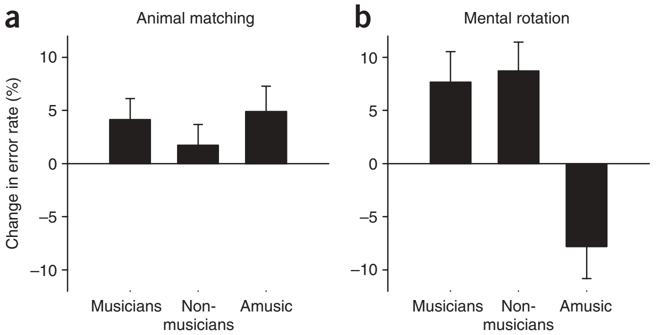

**Figure 7** Changes in error rate on animal matching and mental rotation tasks when performed concurrently with pitch discrimination. Mean ± s.e.m. percentage change in error rate. (**a**) Animal matching; (**b**) mental rotation. Positive levels of interference indicate impaired performance in the dual-task paradigm compared to baseline (and vice versa). Amusic subjects were significantly less impaired than controls on the mental rotation task, but not the animal matching task, when performing the concurrent pitch discrimination task.

When we assessed performance on the animal matching task, we found that there were no significant between-group differences in time taken to perform each trial regardless of the interference condition (**Fig. 6a**). In the mental rotation condition, however, the control group was significantly slower than the amusic group ($t_{32} = 2.43$, $P < 0.05$). An ANOVA analyzing both mental rotation and animal matching data and with controls split into subgroups revealed a significant effect of group ($F_{2,39} = 7.6$, $P < 0.01$) and a significant group × interference condition interaction ($F_{2,39} = 3.34$, $P < 0.05$) (**Fig. 6b**).

When we analyzed the error rates (% error) for the animal matching task with a comparison between baseline and dual task performance, there was no group effect or group × interference interaction. However, there was a significant interference effect ($F_{1,39} = 13.75$, $P < 0.01$), indicating that all groups had been affected similarly by the dual-task procedure (**Fig. 7a**).

The results of a similar ANOVA conducted on the mental rotation interference condition showed that there was a significant effect of group ($F_{2,39} = 6.66$, $P < 0.01$) and a borderline effect of interference ($F_{1,39} = 3.66$, $P = 0.06$). There was also a significant group × interference effect ($F_{2,39} = 12.88$, $P < 0.001$). Post-hoc ANOVAs conducted to reveal the source of these interactions showed significant group × interference interactions between musicians and amusics ($F_{1,26} = 16.3$, $P < 0.001$), and between non-musicians and amusics ($F_{1,26} = 34.0$, $P < 0.0001$), but not between musicians and non-musicians. Amusic subjects, therefore, showed significantly less interference on this task than subjects in either of the control groups (**Fig. 7b**).

## DISCUSSION

We identified a group of subjects who were impaired on the contour-violated subtest of the MBEA, a measure of amusia. Scores on this subtest have been shown to be consistent with a composite MBEA score that characterizes amusia[3]. We verified that this group had amusia by showing that they also had difficulty in identifying the relative pitch relationship between the first two notes of a melody that they had rated as being familiar. When we tested amusic subjects in a mental rotation task, their performance was significantly worse than that of non-amusic control subjects. The difficulty that amusic subjects have with melodic discrimination is therefore related to poor performance on a test of spatial processing.

The results from two further tests were consistent with this initial finding. The amusic group showed a reduced SRC effect compared with the control group, indicating that amusic subjects have a weaker link between spatial representation and pitch discrimination. Furthermore, in the dual-task procedures, amusic subjects' reaction times on the pitch discrimination task were significantly faster than those of non-amusic subjects when they were concurrently performing a mental rotation task. This effect seemed to be specific to the spatial task as it was not apparent when the subjects were performing a control animal matching task. Although performance was high for all groups in this latter task, raising the possibility that differences in error rates were lost in a floor effect, it should be noted that when animal matching errors were increased in the dual-task condition there was no evidence that the amusic group was differentially affected compared to the control group, as would be predicted by this hypothesis. We found the same pattern of results when we compared mean time and error rates on the mental rotation and animal matching tasks.

These results cannot be attributed to differences in the age, level of education or general hearing ability of the subjects, which were similar across groups. Similarly, years of musical training or time spent listening to music could not explain the differences, as these were also similar in the amusic and non-musician control groups. However, the amusic group comprised two males and six females, whereas the control groups had equal numbers of each sex. As females perform more poorly than males on spatial tasks such as mental rotation[15–17], it is possible that sex bias in the amusic sample contributed to the magnitude of the observed effects. However, closer examination of the results indicates that this cannot account for the full effect. First, there was no clear grouping of males at the upper end of the mental rotation performance range, and when we used a hierarchical regression procedure to partial out the effects of sex before assessing mental rotation influence on MBEA score, mental rotation still accounted for a large proportion of the variance. It is of interest to consider whether the sex bias in the amusic group reflects the pattern that would be obtained in a random sampling of a larger population of amusic individuals. To our knowledge these data are not available, although the sample of amusic individuals described in a previous study[3] was also biased toward females.

Our findings support the proposal that amusia is more than a simple deficit in fine-grained pitch discrimination[9]. Here we show that melodic amusia is highly correlated with poor performance on a task that requires the manipulation of objects in space. Furthermore, amusic individuals are less susceptible to the interference that can occur when control subjects perform a task that combines spatial and melodic components. However, the direction of causality between spatial and musical ability remains unclear. There might be a bidirectional relationship between spatial and musical ability, whereby an increase in spatial ability will lead to enhanced musical ability and vice versa, because the two processes share a common representational framework. In support of this proposal, children who are randomly selected to take part in music lessons subsequently outperform their peers, who receive either computer lessons or no intervention, on tests of spatial skills[18], and well-trained orchestral musicians show enhanced mental rotation abilities[19]. The 'Mozart effect', a transient, and controversial, enhancement of spatial abilities resulting from brief exposure to classical music, is also consistent with this proposal[20,21]. Interestingly, in two amusic participants whose data were excluded from our study because they were left-handed, mental rotation performance was higher than the mean for the right-handed, amusic subjects (**Supplementary Fig. 1** online). Although the sample size is low, this raises the possibility that musical and spatial ability might be dissociable under some circumstances, and that lateralization processes might influence their relationship[22].

An alternative possibility is that musical and spatial ability might be independent of each other but both are modulated by some common factor, such as sex hormone exposure. It has been proposed that fetal testosterone facilitates the development of the right hemisphere, leading to enhanced spatial and musical abilities[23]. Consistent with this, women who are exposed to high levels of androgens during prenatal development, either because they suffered from congenital adrenal hyperplasia, or because their mothers were given synthetic diethylstilbestrol, score better than controls (women who have not been exposed to high levels of androgens) on spatial tasks[24]. Mental rotation performance in women varies with the menstrual cycle and shows positive and negative correlations with testosterone and estradiol, respectively[25]. Mental rotation performance is also better in individuals with smaller 2D:4D digit ratios (relatively longer ring fingers than index fingers), a characteristic that has been putatively linked to prenatal testosterone levels[26]. Although it is known that in some species androgens can modulate melodic vocalization[27], at this stage there is little evidence that hormonal exposure has direct effects on musical ability in humans. However, there is some evidence that adrenocorticotropic hormone levels (which modulate androgen synthesis) are positively correlated with performance on spatial and musical tests[28], and a smaller 2D:4D ratio has been linked to superior musical abilities[29].

In summary, our findings provide evidence that amusia is strongly linked to a deficit in spatial representation or processing. This provides an important insight into the possible causes of amusia and, more generally, the cognitive underpinning of melodic representation. Furthermore, these findings have potential implications for the early targeting of individuals who might be prone to amusia. It would be of interest, for example, to determine whether amusic deficits can be ameliorated through spatial training.

## METHODS

**Subjects.** Thirty-four volunteers (19 females and 15 males; mean age: 22.0) took part in all aspects of the experiment. Informed consent and Otago Ethics approval was obtained for all subjects. The 'non-musician' (six females and six males; mean age: 21.8) and 'amusic' groups (six females and two males; mean age: 22.0) comprised students from Otago and Canterbury Universities. The 'musician' group (seven females and seven males; mean age: 22.1) was made up of students from these universities who were also present or past members of the New Zealand Youth Choir or who were classed as musicians according to the Musicality Questionnaire and musicianship criteria. Musicians were classified as such if they reached at least two of the three following requirements: two or more instruments learned for 5 years or more or grade five achieved in both; grade six music theory achieved or underwent six years or more of musical training; and current rehearsal and performance of an instrument or voice for more than eight hours per week.

All subjects were right-handed as measured by the Edinburgh Handedness Inventory[30]. For thirty-two of the subjects, English was their native language. Indonesian and Mandarin were the native languages of the two other subjects (both in the non-musician group), but they were both fluent English speakers and had spoken it throughout their lives. All subjects had received at least 14 years of formal education, and therefore were not cognitively restricted. None of the subjects had any psychiatric or neurological illness, or had hearing problems.

**Materials.** We created a Musicianship Questionnaire to gain general information concerning the subject's gender, age, educational level and musical abilities. The questionnaire was followed by questions concerning the subjects' beliefs about their musicality. We used the Edinburgh Handedness Questionnaire[30] to determine the handedness of the subjects and the strength of this handedness.

Mental rotation stimuli were presented on black cards (8 cm × 18 cm). These were line drawings of two three-dimensional objects created from cubes and each rotated to a different orientation, as used in ref. 31. Subjects were required to report whether the two objects were the same (requiring a 'yes' response) or different (requiring a 'no' response). During the animal matching task, subjects were presented simultaneously with pictures of 15 different animals (20 cm × 30 cm) and a further test sample of three different animal pictures (8 cm × 18 cm). Again, 'yes' or 'no' responses were required by the subjects according to whether all the three test animals appeared in the larger panel or not.

During the SRC pitch discrimination experiments, subjects were exposed to different pair combinations of pure tones produced using Audacity (Version 1.2.4, http://audacity.sourceforge.net). Eight tones were used as comparison tones at frequencies of 164.81, 185.00, 207.65, 233.08, 293.66, 329.63, 369.99 and 415.30 Hz (corresponding to E3, F3#, G3#, A3#, D4, E4, F4# and G4#, respectively). The fixed reference tone was at a frequency of 261.63 Hz, corresponding to C4 (middle C). Test tones were presented binaurally through headphones (Panasonic, RP–HT376) that were connected to a laptop computer. The tones were presented at a comfortable listening level and at constant amplitude. The program used to run the pitch discrimination tasks was written in the language DMDX (version 3.1.4.4; J. C. Forster, University of Arizona). The pitch discrimination task consisted of 24 pairs of pure tones in which the fixed reference tone was followed by one of the eight comparison tones. Each tone lasted for 1,000 ms with no delay between the fixed reference tone and the comparison tone. Once the subject had responded to the tones by pressing an appropriate key, the words 'RESPONSE RECEIVED' appeared on the screen for 500 ms.

We used the MBEA to determine whether the subjects were amusic. The battery usually contains six components that test key musical abilities. All six of the tests use the same pool of 30 musical phrases that were composed according to the rules of the Western tonal system by Irene Deliege. Owing to time constraints, we used only one of the six subtests. Previous studies have shown that mean performance on all of these tests is similar in control subjects (either 27 or 28 out of 30; ref. 3). We used the contour-violated melody subtest, the first of three melodic organization tests. This subtest contained 30 same-different trials in which a standard melody was followed by a comparison melody. For trials where the comparison melody differed from the standard, one note was altered, thereby changing the contour of the melody while preserving the scale.

The Musical Memory and Perception Questionnaire was created to measure the participants' musical imagery and memory of ten well known melodies, such as the New Zealand National Anthem. For each melody the participant was first asked to rate how familiar it was on a scale of one (very familiar) to eight (not familiar). They were then asked whether, for this melody, the second note of the tune was higher or lower than the first note. Four possible responses could be made. These were 'higher', 'lower', 'the same' or 'don't know'. The participants were told to avoid guessing and, if they were unsure of the answer, to circle the 'don't know' option.

**Procedure.** Before taking part in the 45-min experiment, subjects read an information sheet, signed an informed consent form, and completed the Musicality Questionnaire and the Edinburgh Handedness Questionnaire. Not all individuals who filled out the Musicality Questionnaires could participate in the experiment, as many of them could not clearly be classed as either musicians or non-musicians. A high standard of musicianship was required for classification as a musician and negligible musical experience was required for inclusion in the non-musician group. The Musical memory and perception questionnaire was then completed, followed by baseline mental rotation (spatial) and animal matching (control) tasks (the order of these was counterbalanced). During the mental rotation trials, subjects were sat in front of a laptop computer, presented with instructions, and given two practice trials. The experimenter then presented successive mental rotation cards above the laptop screen. Subjects responded verbally. This procedure continued until 20 cards had been presented. The total time taken was measured by stopwatch and the number of errors made during the task was recorded.

The procedure for the animal matching task was similar to the mental rotation task except for the stimuli used. Again, 20 sample cards were presented to compare to the panel of 15 animals, and the responses from the subject were either 'yes' or 'no'. Two different panels of 15 animals were used and swapped after every fifth trial.

The subjects then completed the SRC pitch discrimination tasks. Half of the subjects performed the pitch discrimination with the compatible configuration first and the remaining half completed the pitch discrimination with the incompatible configuration first. During the compatible configuration task, five practice trials were presented to the subject through headphones at a comfortable listening level. The subject then went on to the main experiment, pressing the number '6' on the keyboard with a finger on their right hand if they believed that the second tone was higher than the first tone and the letter 'b' if they believed that the second tone was lower than the first. After 24 trials the subjects were given a 1-min break before completing the next part of the experiment, in which the configuration of the responses was swapped.

The interference task involved performing two of the above tasks simultaneously: either the compatible component of the SRC task (pitch discrimination) with the mental rotation task, or pitch discrimination with the animal matching task. Again, the subject groups were split so that half would complete the mental rotation task first and the remainder would complete the animal matching task first. Task order was fully counterbalanced. Both tasks were run in the same way as the previous baseline trials with the exception that the spacebar did not need to be pressed to continue on to the next trial. Once the subject had completed the 24th pitch discrimination trial, the experiment ceased and the experimenter recorded the time taken, the number of cards completed and the number of errors made. The computer automatically recorded reaction time and error information for the pitch discrimination task.

All subjects completed the MBEA after the main experimental testing was complete. This meant that during the previous experiments, the experimenter was unaware of whether the subject would belong to the amusic or control group. Subjects placed their headphones on and initially listened to two example trials. The subjects then completed the full 30 trials.

*Note: Supplementary information is available on the Nature Neuroscience website.*

ACKNOWLEDGMENTS
Thanks to L. Franz and R. O'Shea for useful comments on the manuscript and the procedure.

COMPETING INTERESTS STATEMENT
The authors declare no competing financial interests.

1. Stewart, L., von Kriegstein, K., Warren, J.D. & Griffiths, T.D. Music and the brain: disorders of musical listening. *Brain* **129**, 2533–2553 (2006).
2. Allen, G. Note-deafness. *Mind* **3**, 157–167 (1878).
3. Peretz, I., Champod, A.S. & Hyde, K. Varieties of musical disorders. The Montreal Battery of Evaluation of Amusia. *Ann. NY Acad. Sci.* **999**, 58–75 (2003).
4. Sloboda, J.A., Wise, K.J. & Peretz, I. Quantifying tone deafness in the general population. *Ann. NY Acad. Sci.* **1060**, 255–261 (2005).
5. Peretz, I. Brain specialization for music. New evidence from congenital amusia. *Ann. NY Acad. Sci.* **930**, 153–165 (2001).
6. Peretz, I. *et al.* Congenital amusia: a disorder of fine-grained pitch discrimination. *Neuron* **33**, 185–191 (2002).
7. Hyde, K.L. & Peretz, I. Brains that are out of tune but in time. *Psychol. Sci.* **15**, 356–360 (2004).
8. Cuddy, L.L., Balkwill, L.L., Peretz, I. & Holden, R.R. Musical difficulties are rare: a study of "tone deafness" among university students. *Ann. NY Acad. Sci.* **1060**, 311–324 (2005).
9. Patel, A.D., Foxton, J.M. & Griffiths, T.D. Musically tone-deaf individuals have difficulty discriminating intonation contours extracted from speech. *Brain Cogn.* **59**, 310–313 (2005).
10. Ayotte, J., Peretz, I. & Hyde, K. Congenital amusia: a group study of adults afflicted with a music-specific disorder. *Brain* **125**, 238–251 (2002).
11. Foxton, J.M., Dean, J.L., Gee, R., Peretz, I. & Griffiths, T.D. Characterization of deficits in pitch perception underlying 'tone deafness'. *Brain* **127**, 801–810 (2004).
12. Griffiths, T.D. *et al.* Spatial and temporal auditory processing deficits following right hemisphere infarction. A psychophysical study. *Brain* **120**, 785–794 (1997).
13. Pratt, C.C. The spatial character of high and low tones. *J. Exp. Psychol.* **13**, 278–285 (1930).

14. Rusconi, E., Kwan, B., Giordano, B.L., Umilta, C. & Butterworth, B. Spatial representation of pitch height: the SMARC effect. *Cognition* **99**, 113–129 (2006).
15. Astur, R.S., Ortiz, M.L. & Sutherland, R.J. A characterization of performance by men and women in a virtual Morris water task: a large and reliable sex difference. *Behav. Brain Res.* **93**, 185–190 (1998).
16. Collins, D.W. & Kimura, D. A large sex difference on a two-dimensional mental rotation task. *Behav. Neurosci.* **111**, 845–849 (1997).
17. Gouchie, C. & Kimura, D. The relationship between testosterone levels and cognitive ability patterns. *Psychoneuroendocrinology* **16**, 323–334 (1991).
18. Rauscher, F.H. *et al.* Music training causes long-term enhancement of preschool children's spatial-temporal reasoning. *Neurol. Res.* **19**, 2–8 (1997).
19. Sluming, V., Brooks, J., Howard, M., Downes, J.J. & Roberts, N. Broca's area supports enhanced visuospatial cognition in orchestral musicians. *J. Neurosci.* **27**, 3799–3806 (2007).
20. Rauscher, F.H., Shaw, G.L. & Ky, K.N. Music and spatial task performance. *Nature* **365**, 611 (1993).
21. Schellenberg, E. Music and cognitive abilities. *Curr. Dir. Psychol. Sci.* **14**, 317–320 (2005).
22. Kopiez, R., Galley, N. & Lee, J.I. The advantage of a decreasing right-hand superiority: the influence of laterality on a selected musical skill (sight reading achievement). *Neuropsychologia* **44**, 1079–1087 (2006).
23. Geschwind, N. & Galaburda, A.M. Cerebral lateralization. Biological mechanisms, associations, and pathology: I. A hypothesis and a program for research. *Arch. Neurol.* **42**, 428–459 (1985).
24. Berenbaum, S.A. Cognitive function in congenital adrenal hyperplasia. *Endocrinol. Metab. Clin. North Am.* **30**, 173–192 (2001).
25. Hausmann, M., Slabbekoorn, D., Van Goozen, S.H., Cohen-Kettenis, P.T. & Gunturkun, O. Sex hormones affect spatial abilities during the menstrual cycle. *Behav. Neurosci.* **114**, 1245–1250 (2000).
26. Peters, M., Manning, J.T. & Reimers, S. The effects of sex, sexual orientation, and digit ratio (2D:4D) on mental rotation performance. *Arch. Sex. Behav.* **36**, 251–260 (2007).
27. Wade, J. & Arnold, A.P. Sexual differentiation of the zebra finch song system. *Ann. NY Acad. Sci.* **1016**, 540–559 (2004).
28. Hassler, M., Gupta, D. & Wollmann, H. Testosterone, estradiol, ACTH and musical, spatial and verbal performance. *Int. J. Neurosci.* **65**, 45–60 (1992).
29. Sluming, V.A. & Manning, J.T. Second to fourth digit ratio in elite musicians: Evidence for musical ability as an honest signal of male fitness. *Evol. Hum. Behav.* **21**, 1–9 (2000).
30. Oldfield, R.C. The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* **9**, 97–113 (1971).
31. Shepard, R.N. & Metzler, J. Mental rotation of three-dimensional objects. *Science* **171**, 701–703 (1971).